NORTHWESTERN UNIVERSITY

Evaluating Engagement with Behavior Change Apps for Groups Experiencing Disadvantages

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Media, Technology, and Society

By

Hannah Badal Getachew-Smith

EVANSTON, ILLINOIS

June 2022

Abstract

Digital behavior change interventions (e.g., mHealth, websites, behavior change apps) can be an effective way to engage groups who experience disadvantages in terms of social and economic attainment, with tailored health content and have potential to improve health outcomes and reduce health disparities. Given the importance and pervasiveness of behavior change apps, it is essential to evaluate their effectiveness to determine if they create change in behavior that improves health and well-being. Engagement, or how users interact with behavior change apps, is often used to quantify success. However, evaluating user engagement with behavior change apps is difficult. Within the field of behavioral sciences, little consensus exists on how to conceptualize and measure user engagement with behavior change apps. Effective measures to evaluate engagement are needed. This would allow for a better understanding of effectiveness and inform future interventions. Moreover, there is limited research on engagement with groups experiencing disadvantages.

Thus, the purpose of this dissertation was to conduct exploratory research to understand behavioral and psychosocial (e.g., cognitive and affective) engagement with behavior change apps for groups experiencing disadvantages. Using a behavioral scientist lens, I employed a mixed-methods approach to evaluate engagement with an existing intervention, the Centers for Disease Control and Prevention's (CDC) *Milestone Tracker*— a mobile app to help families identify developmental delays and disabilities in young children. Using a three article model approach and drawing on models of engagement that center engagement as a multidimensional phenomenon inclusive of behavioral (both engagement with the app and engagement with the health behavior), cognitive, and affective processes, this dissertation consisted of three studies: (1) a systematic review of the literature (*N* = 21 articles included) to identify existing conceptual definitions, measures, and methods to measure engagement with behavior change apps for groups experiencing disadvantages, (2) baseline and one-month follow-up surveys and app usage data to understand parent and caregiver ($N = 72$) engagement with the *CDC's Milestone Tracker* app, and (3) interviews to explore perceptions of engagement with *CDC's Milestone Tracker* app from the perspective of pediatric clinicians at Federally Qualified Health Centers ($N = 20$). The findings from these three studies fill a gap in the engagement literature specific to behavior change apps for groups experiencing disadvantages. The culmination of this research helps to consolidate our understanding of engagement and can guide future intervention and evaluation design to develop behavior change apps that are engaging for all.

Acknowledgment

This journey was only possible because of the unwavering love and support of so many people. First, an enormous thank you to my dissertation committee, Drs. Courtney L. Scherr, Megan Y. Roberts, and Madhu C. Reddy, who gave me the space and guidance I needed to conduct this research and provided me with thoughtful feedback on this dissertation. My deepest gratitude to my committee chair, Dr. Courtney Scherr, who also served as my advisor throughout the duration of my doctoral program. Thank you for your time and energy and for trusting me to help with your research projects. Conferencing together from Portugal to New Orleans to Washington DC will forever be a highlight of my time at Northwestern. I learned so much from you about conducting research, teaching, and writing during our six years working together, and will carry much of it with me as I take the next steps in my career.

Beyond my committee, I am grateful to a number of faculty who provided mentorship and advice throughout my doctoral program: Drs. Daniel J. O'Keefe, Nathan Walter, Jakob D. Jensen, Lauren S. Wakschlag, and Andy J. King. I am especially thankful to Andy for sharing your insight, playlists, and love of beer that made graduate school considerably more enjoyable.

I am forever grateful to everyone who worked in the Health Communication Interaction Design Lab, especially my lab mates, Sanjana Ramesh, Kerstin Kalke, Callie Sartain Kalny, and Hoda Fakhari. I could not have asked for a better group of women to work with, write with, and learn from.

To my MTSB cohort—Bri Hightower, Angela Smith, Ada Ng, Jabari Evans, Anne-Marie Singh, Renwen (Alice) Zhang, Kaitlyn Childs, Ellie Burgess, and Diego Gómez-Zará—I really lucked out going on this adventure with such an amazing group of people. A special thank you to Bri Hightower for being the best coworking buddy and cheerleader. You are a gem.

I was incredibly fortunate to have two amazing role models in Miya Williams Fayne and Amy Ross Arguedas. Thank you both for your advice and for helping me navigate graduate school from day one. A special thank you to Marlon Twyman for always checking in on me. Many thanks and cheers to my fellow Presidential Fellows—Heather McCambly, Lindsay Zimmerman, Lea Richardson, Ángel Escamilla Garcia, and Luis Fernando Amaya Muñoz—you are all brilliant and inspiring. To Ilya Gokhman, Broderick Turner, and the rest of the grown-up happy hour crew, thank you for all of the laughs and for helping me maintain perspective. I am so grateful our paths crossed.

Thank you to the many mentors who have helped me immensely throughout my career including Jo Ellen Stryker, who hired me for my first job after Emory and who was the first person to encourage me to apply to doctoral programs. Thank you to my former colleagues Jocelyn Coleman-Taylor, Donata Green, Kimberly Leeks, and Euna August, who taught me how to be a good researcher and evaluator. Many thanks to Garth Walker and Lew Owen for sharing your academic and parenting expertise. To Sam Macy and Sean Serluco, thank you for listening and validating my experiences and for coaching me through this process.

This research was made possible with the support of a team of people who helped with data collection and data analysis. Thank you to Nihmotallahi (Nihmot) Adebayo and Alexandra Brockner, who screened and coded more than 1900 articles for the systematic review; Northwestern University librarian Lauren McKeen who consulted on the systematic review plan; Sydney Moe who helped to code the pediatric clinician interviews; Nivedita Mohanty who recruited the pediatric clinicians; and Nina Stein who coordinated participant payment. Thank you also to all of the parents and caregivers (the real humans, not the bots) who helped to distribute and complete my surveys, and to the pediatric clinicians who took time out of their busy schedules to share their experiences and perspectives.

Thank you to my colleagues on the *Learn the Signs. Act Early.* team at the Centers for Disease Control and Prevention and the Association of University Centers on Disabilities for the opportunity to evaluate the *CDC's Milestone Tracke*r app, for sharing your app data, and for inviting me to the Act Early world: Rebecca Wolf, Katie Green, Amber Broughton, Samantha Budzyn, Karnesha Slaughter, Danielle Webber, and Harpur Schwartz.

My research was generously supported by several fellowships and grants: The Presidential Fellowship at Northwestern University, the Institute for Innovations in Developmental Sciences (DevSci), the Sexualities Project at Northwestern, The Graduate School, the Department of Communication Studies, the National Communication Association Health Communication Division, the National Communication Association Black Caucus, and the Society of Behavioral Medicine.

To my friends who kept me centered throughout this process with your empathy and encouragement. Ama Appenteng-Milam, Uchenna Moka-Solana, Sonia Atsegbua, Lauren Alexander, Natasha Dossa, and Jordan Curry, thank you for the countless book clubs, study breaks, check-ins, and goodies to keep me motivated. To my sisters, Anna Darby, Natalia Casella, Sarah Hernandez, Rivkah Cohen, Jennifer Purdy, and many others, thank you for being there in times of crisis and celebration, and for always hyping me up, especially when it felt like the weight of the world was on my shoulders. Thank you to my brothers, Niko Philpides, Evan Meredith, Dante Williams, Nick Perovich, and many others for making sure I did not strike out with the bases loaded.

I am eternally grateful to my parents, Jeanette Badal and Raymond Walters, who instilled in me the confidence and fight to persevere. Mama, your commitment to service is what drove me to the field of public health and encouraged me to view my work from a health equity lens. You have always been my number one fan and loved me unconditionally. Daddy, thank you for answering the phone at any hour to impart your wisdom and for always beaming with pride at every milestone. This degree is for you. Thank you to Kiddy Getachew-Smith and David Getachew-Smith, Sr. for your love and for helping with Kenny when I desperately needed to work.

To my incredibly compassionate and supportive brother Daniel Murray-Badal and his brilliant partner in life, Kara Murray-Badal, thank you for making time and space for me and for spreading your joy. Thank you to my other brother, Daniel Getachew-Smith, for being my champion and keeping me motivated. Thank you to the Brinson family for always having my back. Love and thanks to my grandparents through marriage, Pauline Smith, Getachew Gessesse, and Tsehainesh Negatu who welcomed me with open arms.

While there are too many to name, thank you to all of my aunts, uncles, and cousins who have always believed in me. Please know that your love and support are forever appreciated. Special thanks to my cousin Kyle Southern, who made getting a PhD look easy and who shared excellent memes and advice about graduate school at the most opportune times. While they are no longer physically here with us, my grandparents Alice Virginia Badal and Baba Joseph Badal, and dear uncle Robert Edward Badal, guided me through the darkest moments of this journey and gave me the strength to keep going.

To Kenny, who entered this world as I wrote my prospectus, you are the greatest blessing. You brought sunshine and clarity even on my busiest and hardest days and reminded me what is most important—you. Thank you for letting me work when you wanted to play, and for always knowing when I needed a hug. To Dre—the best pup and my road dog—thanks for staying up late with me so that I never felt alone.

And to David Getachew-Smith, Jr., my rock and best friend. None of this would have been possible without your enduring support and patience. Thank you for taking a risk and moving to Chicago with me and for all the time and energy you put into making our family whole. Even though we did not plan it this way, I am so appreciative I got to spend the two years that I worked on this dissertation with you as my officemate. Thank you for listening and for problem-solving with me when I needed it most. We did it.

List of Abbreviations

ADHD	Attention-deficit/hyperactivity disorder
AAP	American Academy of Pediatrics
ASD	Autism spectrum disorder
BCTs	Behavior change techniques
CDC	Centers for Disease Control and Prevention
DBCI	Digital behavior change intervention
EI	Early intervention
FQHC	Federally Qualified Health Center
HCI	Human-computer interaction
LTSAE	*Learn the Signs. Act Early.*
MoAs	Mechanisms of action
SES	Socioeconomic status
WIC	Special Supplemental Nutrition Program for Women, Infants, and Children

For Virginia and Bobby, who were there from the start, and for Kenny, who helped me cross the finish line.

Table of Contents

List of Tables and Figures

Tables

Figures

Chapter 1

Introduction to Research Program

Digital technologies have revolutionized how we consume information and hold great promise to improve individual and collective health and well-being (Devlin et al., 2015; Murray et al., 2016; Suggs, 2006). Particularly within the context of the COVID-19 pandemic and mitigation measures, digital tools were mobilized to support all aspects of life. These digital health interventions include websites, smartphone mobile applications (hereafter referred to as apps), wearables such as Fitbits, and social media designed to promote healthy behaviors, manage chronic conditions, and provide access to treatment (Aitken & Nass, 2021; Murray et al., 2016). Digital behavior change interventions (DBCI) are digital health interventions rooted in behavior change theories that use digital channels to promote and maintain health via primary and secondary prevention and management (Yardley, Choudhury, et al., 2016). DBCI often include behavior change techniques (BCTs) which are intervention components designed to enable behavior change (Michie et al., 2013). DBCI can engage groups experiencing disadvantages in terms of social and economic attainment, with tailored health content and have potential to improve health outcomes and reduce health disparities (Kreps & Neuhauser, 2010; Kreuter et al., 2013; Lustria et al., 2009).

Given the importance and pervasiveness of DBCI, it is essential to evaluate their effectiveness to determine if they create changes in behaviors that improve health and well-being. Engagement, or how users interact with DBCI, is often used to quantify success (Perski, Blandford, West, et al., 2017; Short et al., 2018; Yardley, Spring, et al., 2016). Existing methods to measure engagement include: (1) analysis of system usage data (e.g., time spent in an app and number of pages viewed), (2) qualitative methods, (3) ecological momentary assessments (real-time, self-reported user data), and (4) self-report questionnaires (few of which are validated and specific to health; Short et al., 2018). However, little consensus exists within behavioral sciences on how to conceptualize and measure user engagement with DBCI, and what data currently associated with engagement truly represent. Effective measures to evaluate engagement with DBCI are needed to understand how users interact with DBCI. This would allow for a better understanding of effectiveness and inform future interventions (Cole-Lewis et al., 2019). Moreover, there is limited research on engagement behaviors and experiences among groups experiencing disadvantages (Cole-Lewis et al., 2019).

This exploratory dissertation project sought to understand behavioral and psychosocial (e.g., cognitive and affective) engagement with DBCI for groups experiencing disadvantages from the perspective of the field of behavioral sciences. Building on behavioral science theories and practices, I conducted formative research to explore how engagement with DBCI (specifically behavior change apps) for groups experiencing disadvantages is conceptually defined, measured, and what methods are used to measure engagement, and then I explored engagement from the perspectives of parents and caregivers of young children and pediatric clinicians. I situated my dissertation in the context of developmental delays and disabilities among young children, a widespread issue that disproportionately affects groups experiencing disadvantage (e.g., people from minoritized racial and ethnic groups, people with lower levels of socioeconomic status, and people with disabilities), but can be improved through early identification and early interventions (Adams & Tapia, 2013; Zablotsky et al., 2019). Utilizing the Centers for Disease Control and Prevention's (CDC) *Milestone Tracker*—a mobile app to help families (specifically those from groups experiencing disadvantages) identify developmental delays and disabilities in young children—this project employed a mixed-methods approach to explore engagement measurement.

The purpose of this dissertation was to conduct formative research to understand how to better measure engagement with behavior change apps for groups experiencing disadvantages. The specific context in which engagement was studied for this dissertation was one behavior change app, the *CDC's Milestone Tracker* app. Using a three-article model, this dissertation includes three studies. The first study, reported in the second chapter, was a systematic review of the literature, which aimed to synthesize existing conceptual definitions, measures, and methods of engagement with behavior change apps for groups experiencing disadvantages. The second study (Chapter 3) used a baseline and one-month follow-up survey and app usage data to understand parent and caregiver engagement with the *CDC's Milestone Tracker* app. Finally, the third study (Chapter 4) used interviews to explore perceptions of engagement with *CDC's Milestone Tracker* app from the perspective of pediatric clinicians at Federally Qualified Health Centers (FQHC). Together these three studies fill a gap in our understanding of engagement specific to behavior change apps for those experiencing disadvantages and inform the future development of a behavior change app engagement measurement framework.

The remainder of this chapter details the research program and provides a roadmap for the three studies described in the chapters that follow. The next section provides an overview of the current DBCI landscape, a discussion of behavior change apps and health disparities, and an examination of engagement. Additionally, an overview of developmental delays and disabilities among young children and the *CDC's Milestone Tracker* app are presented. Finally, this chapter concludes with a statement on motivation and positionality, a summary of the dissertation study purpose, and a brief outline of each study included in this dissertation.

Background

Digital Health Landscape

In 2022, Internet use is near universal (Pew Research Center, 2021a). With 93% of adults in the United States using the Internet, digital channels are a convenient way to increase the reach of health interventions to improve health outcomes (Cugelman et al., 2011; Pew Research Center, 2021a). As a result of the COVID-19 pandemic, investments in digital health grew almost 80% to more than $57 billion (Donnelly, 2022) as medicine and health promotion shifted to remote and virtual service delivery. As such, there has been a proliferation of eHealth (electronic health) and mHealth (mobile health) interventions delivered via digital platforms such as websites, mobile apps, wearables for fitness tracking, parameter-specific biosensors, social media, telemedicine, virtual home assistants (e.g., Alexa and Google Home), personal health records, and text messaging and email (Aitken & Nass, 2021; Labrique et al., 2013; Murray et al., 2016; Sinnenberg et al., 2017). DBCI draw on behavior change theories and include BCTs via their corresponding mechanisms of action (MoAs; the processes through which BCTs have their effects; Carey et al., 2018) to promote healthy behaviors like physical activity (Bennett et al., 2014; Glasgow et al., 2007; Molina & Sundar, 2018; Pagoto et al., 2018; Sharpe et al., 2017; Short et al., 2014), healthy eating (Delaney et al., 2018; Vandelanotte et al., 2016), and smoking cessation (Blok et al., 2019; Businelle, Ma, Kendzor, Frank, Vidrine, et al., 2016; Businelle, Ma, Kendzor, Frank, Wetter, et al., 2016; Coa et al., 2019; Herbst et al., 2019), manage chronic conditions such as diabetes (Glasgow et al., 2011; Lie et al., 2017; Sepah et al., 2017) and heart disease (Anand et al., 2016; Beatty et al., 2013), support HIV testing, linkage to care, and treatment adherence (Cao et al., 2017; Muessig et al., 2017), and provide access to treatment for mental health (Firth et al., 2017; Fitzpatrick et al., 2017; Larsen et al., 2016; Mani et al., 2015; Mohr et al., 2017; Ng et al., 2018), and pain management (Escriva Boulley et al., 2018; Nicholl et al., 2017). Relevant to this dissertation project, DBCI are used to improve early identification of developmental delays and disabilities in young children, as discussed later in this chapter.

Benefits of DBCI

Compared to didactic health information found on posters and brochures, an advantage of DBCI is that they are intended to attract and engage users by requiring interactivity (e.g., user participation and input) to create two-way communication between the system and user (O'Brien & Toms, 2008; Sawesi et al., 2016). Further, DBCI have the potential to reach more people in their everyday lives, which allows for improved sustainability and reduced dissemination costs (Bernhardt et al., 2011; Devlin et al., 2015; Kreps & Neuhauser, 2010). Interventions can now reach millions of people at a fraction of the cost of delivering traditional interventions (e.g., face-to-face; Suggs, 2006). For example, while smoking cessation interventions via phone were estimated to cost $150 to $250 per smoker, and tailored print interventions ranged from $5 to $40 per smoker, tailored online smoking cessation interventions could cost less than $1 per smoker, depending on the population (Cugelman et al., 2011). Other analyses further demonstrate the magnitude of potential healthcare cost savings. Digital health apps designed for five patient populations (diabetes prevention, diabetes, asthma, cardiac rehabilitation, and pulmonary rehabilitation) could reduce acute care utilization, saving an estimated $7 billion annually (Aitken et al., 2017). If savings were extended to all diseases and health conditions, an estimated $46 billion could be saved (Aitken et al., 2017). Further, the COVID-19 pandemic highlighted the need for interventions that can be implemented remotely and asynchronously to improve health and well-being when in-person interventions are not an option.

DBCI also provide more tailored messaging than interventions delivered using traditional channels (e.g., in-person; Lustria et al., 2009). Tailored messages are more likely to be persuasive because they are highly personalized based on an individual's predispositions and abilities, making them relevant to priority audiences (Atkin & Rice, 2013; Kreuter et al., 2013; Kreuter & Wray, 2003). Utilizing targeted channels to reach intended audiences (particularly people from groups experiencing disadvantages; Badal et al., 2018) can boost the reach of messages that facilitate behavior change. As such, DBCI need to be evaluated to understand their reach and impact.

Behavior Change Apps

The expansive digital health landscape includes several types of DBCI, yet in recent years, mobile apps have become the most utilized digital channel for health intervention dissemination. Given that 85% of people in the United States own a smartphone, and 15% of those people use their smartphones as the primary way of accessing the Internet (Pew Research Center, 2021a, 2021b), it is not surprising that more than 250 health apps are added to app stores daily on average, resulting in more than 350,000 health apps available worldwide (Aitken & Nass, 2021). Additionally, a very small number of health apps account for the majority of downloads. Recent research indicates that only 110 apps comprise half of all downloads, and each of these have been downloaded over 10 million times. Conversely, the majority (83%) of health app have been downloaded less than 5,000 times (Aitken & Nass, 2021), potentially indicating a need for more appealing and engaging apps to truly capitalize on their affordances to change health behaviors and improve health and well-being.

Behavior Change Apps and Health Disparities

Health is influenced by social determinants such as poverty, lack of access to quality education, unemployment, unhealthy housing, unsafe neighborhoods, incarceration and police brutality, and other systemic or cultural barriers (Solar & Irwin, 2007). Health disparities, or differences in health among groups of people, disproportionately affect those from groups experiencing disadvantages (e.g., people from some racial and ethnic groups, people with disabilities, people with lower incomes), resulting in poorer health outcomes (Kawachi et al., 2005; Kawachi et al., 2002; Williams et al., 2016). Importantly, socioeconomic status (SES), a composite measure of income, education, occupation and social status (Carlson et al., 2011), is a key underlying risk factor for poor health (Adler & Newman, 2002). As such, innovative approaches to address health disparities are needed.

Behavior change apps have expanded opportunities to disseminate health messages and information to groups experiencing disadvantages, who are often disproportionately affected by many health conditions (Gibbons et al., 2011; Kreps & Neuhauser, 2010; Swindle et al., 2014). For example, place-based interventions are impractical for populations experiencing homelessness (Barman-Adhikari et al., 2016), however highly targeted digital channels offer alternative ways to reach groups experiencing disadvantages (Kreuter et al., 2013). Unfortunately, this has not translated into significant growth in socioculturally tailored behavior change apps (Brewer et al., 2020). Grimes Parker, Alcaraz, and others have called for culturally relevant research and design in the creation of behavior change apps for more than 15 years (Alcaraz et al., 2017; Grimes & Grinter, 2007; Stowell et al., 2018), yet there is a paucity of research on behavior change apps with groups experiencing disadvantages (Brewer et al., 2020). The limited research that exists suggests behavior change apps can be effective for weight management among Black or African American (hereafter referred to as Black) and Hispanic or Latinx (hereafter referred to as Latinx) persons (Bennett et al., 2014), parent training programs for parents experiencing poverty (Brager et al., 2019), and colorectal cancer screening among patients with lower incomes (Miller et al., 2018), however a systematic review found few improvements in health outcomes across studies (Stowell et al., 2018). Consequently, many have expressed concerns that if people from diverse groups are not included in the development and evaluation of behavior change apps (including the development of evaluation tools), there is potential to widen health disparities and perpetuate health inequalities (Brewer et al., 2020; DiMaggio et al., 2001; Lee et al., 2022; Lee & Viswanath, 2020). More (and purposeful) research is needed to capitalize on the high adoption and utilization rates of smartphones among groups experiencing disadvantages to reach these communities with interventions that facilitate engagement with behavior change components to improve health outcomes and achieve health equity (Gibbons, 2011; Gibbons et al., 2011; Lopez et al., 2011).

Engagement

Given the merits of behavior change apps, it is important to evaluate their effects and effectiveness. Program evaluation, which is a systematic method for determining overall success of an intervention, assesses whether or not an intervention was implemented as planned and if it improved health outcomes (Valente, 2002). Evaluation enables researchers and practitioners to measure the effectiveness and impact of an intervention to create intended behavior change (and antecedents to behavior change). Research on the effectiveness of behavior change apps is mixed, varying by health issue and intervention purpose (Hall et al., 2015). Evaluations often demonstrate short-term benefits, but evidence of long-term benefits and population-level impacts of behavior change apps is lacking (Free et al., 2013; Gordon et al., 2021; Hall et al., 2015).

Within the context of behavior change apps, some researchers argue evaluating behavior change is too complex and can be limiting for novel apps and other DBCI. Instead, they argue that focusing on evaluation to understand how and why behavior change apps are used would be more valuable to help researchers learn how and why a system works or does not work (Klasnja et al., 2011). Further, some behavior change app evaluations demonstrate limited effects, citing attrition, incorrect or suboptimal use of the app as the reason (Kelders et al., 2012; Lie et al., 2017; Sieverink et al., 2017). For example, up to 50% of participants stop using downloaded health apps, noting data entry burden, loss of interest, and costs as reasons for disengaging with apps (Krebs & Duncan, 2015; Lie et al., 2017; Pfammatter et al., 2017). The high number of participants who drop out of behavior change app trials and stop using interventions makes it difficult to establish evidence-based apps and evaluate effects (Eysenbach, 2005). As such, there is a need to study the extent to which people use interventions and reasons for app nonadherence in order to understand how to keep users engaged (Sieverink et al., 2017). Moreover, in behavioral sciences it is assumed that some level of engagement is necessary for behavior change apps to be effective (Coa et al., 2019; Yardley, Spring, et al., 2016). Highly interactive apps may generate the level of exposure necessary to facilitate behavior change (Crutzen et al., 2010). Examining user engagement, or how people interact with digital technology, may help reduce attrition (Couper et al., 2010; Krebs & Duncan, 2015) and is central to understanding how and why behavior change apps are used. Additionally, there is no universally accepted set of constructs or definitions for user engagement with behavior change apps. Measuring engagement is essential to understanding the uses and benefits of behavior change apps, but remains elusive and understudied within behavioral sciences (Short et al., 2018).

Conceptual Definitions

As mentioned, no single definition of engagement exists within behavioral sciences (Cole-Lewis et al., 2019; Yardley, Spring, et al., 2016), which makes it difficult to operationalize and measure. Often used interchangeably with other terms (e.g., usability and acceptability), increasingly behavioral scientists posit that engagement is multidimensional, comprised of psychosocial and behavioral constructs (Flaherty et al., 2021; Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017; Yardley, Spring, et al., 2016). Psychosocial constructs are primarily cognition and affect (Kelders, van Zyl, et al., 2020). Cognitive processes include user attention, interest, flow, cognitive absorption, and immersion in the app content and/or platform. Affective processes include user enjoyment and satisfaction with the app content and/or platform.

Behavioral processes focus on how individuals use the app. Behavioral engagement is believed to include two forms of engagement: (1) how users interact with features of the app designed to encourage use, and (2) how users interact with the behavior change components of the app such as goal-setting, tracking tools for self-monitoring behaviors, and other BCTs. Cole-Lewis et al (2019) argue that how users interact with the health behavior is dependent on how users interact with the technology because if users do not interact with or enjoy using the app, they will have limited exposure to important behavior change components and therefore less likely to change behaviors. As such, both forms of engagement are important. Further, some believe engagement is a process, with different stages or phases of engagement such as getting engaged, staying engaged, disengaging, and re-engaging that fluctuate over time (Kelders, van Zyl, et al., 2020; O'Brien & Toms, 2008).

In accordance with other behavioral scientists, the conceptualization of engagement used in this dissertation is that engagement is multidimensional that involves behavioral (both interaction with health behaviors and interaction with app features), cognitive, and affective processes. Combined, these three domains characterize the user experience with behavior change apps (Perski, Blandford, West, et al., 2017; Yardley, Spring, et al., 2016). While the presence of the three domains is generally agreed upon by behavioral scientists, little is known about how these three domains interact and which are most important to facilitate behavior change (notably, these research gaps are outside of the scope of this dissertation project but important to acknowledge nonetheless). Additionally, recent reviews highlight differences in how behavioral scientists weigh the behavioral, cognitive, and affective components of engagement (Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017; Short et al., 2018; Yardley, Spring, et al., 2016). As such, a recent review by Kelders et al (2020) concluded that context-specific definitions of engagement may be necessary, as long as conceptual definitions are clearly stated and accompanying measures are described.

Methods, Measures, and Frameworks

Existing methods and frameworks to measure digital engagement include: (1) analysis of system usage data (e.g., time spent in an app and number of pages viewed), (2) qualitative methods, (3) ecological momentary assessments (EMA; real-time, self-reported user data), and (4) self-report questionnaires (few of which are validated and specific to health; for a review see Short et al., 2018). These methods capture both objective and subjective measures of the three engagement domains. Each method is discussed briefly below.

System Usage Data. System usage data is backend or log data, sometimes referred to paradata (Couper et al., 2010). Per Couper and colleagues (2010), paradata “capture details about the *process* of interaction with an intervention” (p. 2). System usage data includes metrics such as number of visits, page views, timestamp, overall time and time spent in different aspects of a platform or intervention, and logins (Baltierra et al., 2016; Bauermeister et al., 2017; Morrison & Doherty, 2014; Short et al., 2018). These data can help us understand how users access and move through behavior change apps (i.e., user engagement behaviors), indicating which interventions components are used or viewed most often and what aspects needs to be improved (Bauermeister et al., 2017). System usage data are highly touted as an objective measure of engagement and are most often used to measure behavioral engagement. Combined with other data sources (e.g., self-report questionnaires and EMA), usage data can provide insight into the relationship between usage and psychosocial aspects of engagement (Short et al., 2018). While some measures passively collect user activity data (Perski, Blandford, West, et al., 2017; Short et al., 2018), other measures are active (i.e., answering a quiz or goal-setting), and at times, burdensome for users, and may attribute to the high attrition and nonuse of behavior change apps discussed above.

Qualitative Methods. Qualitative methods, such as semi-structured interviews, focus groups and think-aloud activities, capture subjective measures of engagement and allow for an in-depth description of user experience as well as user perceptions about how the behavior change app facilitates behavior change (Short et al., 2018). Qualitative methods are useful to generate engagement hypotheses and explore hypotheses to understand individual-level engagement. Qualitative methods are beneficial to capture cognitive and affective engagement, but can also be used to measure subjective behavioral engagement (via questions about app use). Further, Short and colleagues (2018) recommend qualitative methods to explore how the relationship between behavior change apps features and psychosocial features relate to use (or nonuse).

Ecological Momentary Assessments. EMA assesses user behaviors, perceptions and experiences in real-time by pinging users (often multiple times throughout a day or other time frame) to respond to short surveys. EMA can provide insight into the subjective behavioral, cognitive, and affective aspects of engagement with behavior change apps (Short et al., 2018).

Self-report Questionnaires. Over the past several years, several self-report questionnaires and scales have been developed to measure engagement with digital platforms. These instruments capture subjective behavioral, cognitive, and affective engagement with the behavior change app (Short et al., 2018). Examples of scales include the: *Immersion Experience Questionnaire* (Jennett et al., 2008); *eHealth Engagement Scale* (Lefebvre et al., 2010); *Health IT Usability Evaluation Scale* (Yen et al., 2010); *User Engagement Scale* (O'Brien & Toms, 2010); *Mobile Application Rating Scale* (Stoyanov et al., 2016); and more recently, the *Digital Behavior Change Intervention (DBCI) Engagement Scale* (Perski, Blandford, et al., 2019; Perski, Lumsden, et al., 2019); and *TWente Engagement with Ehealth Technologies Scale* (TWEETS; Kelders & Kip, 2019; Kelders, Kip, et al., 2020).

Many self-report scales attempt to capture different dimensions of engagement (e.g., involvement, affect, attention) but are somewhat limited in their application, particularly given the rapid advances in technology to request and collect data. With the exception of the *DBCI Engagement Scale* and TWEETS, many engagement instruments and tools were developed outside the context of health, fail to account for health behavior change theories that may explain and contextualize underlying mechanisms, and do not incorporate behavior change as a component of behavioral engagement. Moreover, research on the relationship between subjective measures of engagement and outcomes is also scarce (Graham et al., 2021). As the *DBCI*

Engagement Scale and TWEETS are rooted in the theory that engagement is multidimensional and are the most salient to this dissertation, both are described in more detail below.

Perski et al's *DBCI Engagement Scale*. The *DBCI Engagement Scale* is a 10 item scale that assesses engagement using two independent subscales: (1) *experiential subscale* which includes eight items measured on a 7-point scale ranging from 1 = *Not at all* to 3 = *Moderately* to 7 = *Extremely*, and (2) *behavioral subscale* which includes two items ("How much time [in minutes] do you roughly think that you spent on the app?" and "Which of the app's components do you remember visiting?"). The *DBCI Engagement Scale* posits that for a user to be engaged, both experiential engagement (i.e., attention, interest, and enjoyment or what previously was described as psychosocial factors that comprise engagement such as cognition and affect) and behavioral engagement (i.e., the amount and depth of use) are necessary. Accordingly, the authors define engagement as "a state-like construct which occurs each time a user interacts with a DBCI" along these experiential and behavioral dimensions (Perski, Blandford, et al., 2019, p. 798). Psychometric evaluations of the *DBCI Engagement Scale* show moderate internal reliability but low validity, likely a result of the combination of subjective and objective engagement which demonstrate weak associations between the behavioral subscale and behavioral engagement (Perski, Blandford, et al., 2019; Perski, Lumsden, et al., 2019).

Kelders et al's TWEETS. The TWEETS is a nine-item scale assessed on a 5-point scale ranging from 0 = *Strongly disagree* to 4 = *Strongly agree* that defines engagement as a combination of behavior, cognition, and affect, where behavior is engagement with the technology as well as the health behavior. TWEETS theorizes that behavioral engagement is routine use of a technology that requires low effort to use, but that use may fluctuate to meet current needs. Cognitive engagement is believed to be related to the technology supporting and motiving a user to reach their behavior change goals, driven by intrinsic motivation. Affective engagement involves user's emotions such as enjoyment that stem from making progress towards behavior change goals (or not) and using the technology. TWEETS also considers identity as an important construct as users need to identify with the technology or the goal of the technology (Kelders, Kip, et al., 2020). TWEETS demonstrated good reliability and validity in a study of 288 student users of a step counter behavior change app (Kelders & Kip, 2019; Kelders, Kip, et al., 2020). Conceptually, TWEETS appears to be the most aligned with the framework that theorizes engagement as multidimensional (comprised of two levels of behavior, cognition, and affect), and the most promising given the findings of the psychometric evaluation.

Historically, the generalizability of instruments has been called into question as there are notable challenges related to replication of instruments when applied to new or different samples (O'Brien & Toms, 2013). Relevant to this dissertation, it is worth stating that neither the *DBCI Engagement Scale* nor the TWEETS were developed with diverse populations. Engagement mechanisms are understudied and poorly understood for diverse populations, requiring further inquiry as to whether these measures are sufficient for other audiences (Brager et al., 2019; Cole-Lewis et al., 2019; Flaherty et al., 2021; Kelders, Kip, et al., 2020; Yardley, Spring, et al., 2016).

In summary, engagement with behavior change apps is not well understood and several conceptual definitions and measures exist to assess engagement, but the framework of engagement as multidimensional, characterized by behavioral (interaction with the technology and the health behavior), cognitive, and affective engagement, may provide guidance to help explain how and why users engage with a platform. The next section describes the health context for this dissertation research, developmental delays and disabilities among young children.

Developmental Delays and Disabilities among Young Children

In the United States, the prevalence of developmental delays (e.g., cognitive, speech and language, gross and fine motor, and social delays) and disabilities (e.g., attention-deficit/hyperactivity disorder [ADHD], autism spectrum disorder [ASD], cerebral palsy, blindness, hearing loss, learning and intellectual disability)—which was 15% in 2008—has significantly increased in the past decade (Boyle et al., 2011; Zablotsky et al., 2019). Today, approximately 17% of children in the United States are living with a developmental delay or disability (Zablotsky et al., 2019). A developmental delay is when a child is not developing (e.g., language or fine motor skills) within an anticipated time frame (Sices et al., 2003; Wellman & Davis, 2018) whereas developmental disabilities are mental and/or physical impairments that substantially limit functional daily activities (American Academy of Pediatrics, 2006). Despite the high prevalence of developmental delays and disabilities, most children are not identified as early as possible, with many children diagnosed after 24 months (Boyle et al., 2011; Maenner et al., 2021; Mann et al., 2008). Consequently, up to 90% of children with developmental delays and disabilities do not receive early intervention (EI) therapies or services known to improve functional outcomes and reduce secondary behavioral issues (Adams & Tapia, 2013; American Academy of Pediatrics, 2001b; Guralnick & Bricker, 1987; Rosenberg et al., 2008).

Early Intervention

Early identification of developmental delays and timely intervention can result in improved functional outcomes (Adams & Tapia, 2013; Bradshaw et al., 2015; Bruder, 2010), and can prevent or reduce developmental problems in children and improve long-term function and quality of life (American Academy of Pediatrics, 2001b; Guralnick & Bricker, 1987). Conversely, undetected and untreated developmental delays contribute to early school failure, social, and emotional problems (American Academy of Pediatrics, 2006; Guralnick, 1991). EI services have economic advantages by reducing the need for long-term treatment and the associated long-terms costs (Adams & Tapia, 2013; Bradshaw et al., 2015). As such, the Individuals With Disabilities Education Act (IDEA; Public Law 101-476), mandates that states develop and implement coordinated, family-centered, culturally competent, community-based systems of care to help families with children who have suspected or diagnosed developmental delays and disabilities (Adams & Tapia, 2013; American Academy of Pediatrics, 2001b; Dunst, 2007). IDEA Part C (Public Law 105-17) requires states to provide evaluations to determine if a child qualifies for EI and services such as occupational therapy, speech language pathology and audiology, physical therapy, developmental therapy, service coordination, social work, assistance with transportation and related costs, family training, counseling, and home visitation for children three years and under (American Academy of Pediatrics, 2001b, 2006). Despite this mandate to ensure the availability of EI services, developmental delays at early ages go largely undetected (Rosenberg et al., 2013).

Barriers to Early Identification

Delays in identification of developmental delays likely stem from several challenges in the identification and evaluation process. The spectrum of delays and cognitive issues—varying in type and severity—makes childhood development difficult to measure (American Academy of Pediatrics, 2001a). Thus, a single assessment is insufficient to gain insight into this dynamic developmental process. As such, periodic developmental and autism screening using validated screening instruments to detect delays is necessary (American Academy of Pediatrics, 2001a). The American Academy of Pediatrics (AAP) guidelines recommend developmental screening at 9-, 18-, and 30-month well-child visits and autism screening at 18- and 24-month well-child visits or when a parent or caregiver raises a concern about their child's development (American Academy of Pediatrics, 2006; Lipkin, Macias, Council On Children With Disabilities, et al., 2020). Despite these recommendations, only 63% of pediatricians reported using developmental screening tools, an increase of more than threefold since 2002, but still not nearly enough to facilitate early identification of developmental delays and disabilities (Lipkin, Macias, Baer Chen, et al., 2020).

To complement developmental screening, AAP guidelines also recommend developmental surveillance (also known as developmental monitoring) which is the process of routinely eliciting parent or caregiver concerns and observing a child at every well-child visit (American Academy of Pediatrics, 2006; Lipkin, Macias, Council On Children With Disabilities, et al., 2020). While children who receive a combination of developmental surveillance and screening are more likely to receive EI services (Barger et al., 2018), only 19% of young children received both surveillance and screening (Hirai et al., 2018). Parent and caregiver concerns are strong predictors of developmental delays (American Academy of Pediatrics, 2001a), yet few parents report that clinicians' elicit their concerns during visits (Adams & Tapia, 2013; Marshall et al., 2016) and only 30% of parents and caregivers of children under three years reported receiving developmental screening (Hirai et al., 2018). Parents and caregivers also reported feeling unheard if concerns were voiced (Woolfenden et al., 2015), resulting in delayed identification and intervention (American Academy of Pediatrics, 2001a).

Finally, AAP guidelines call for administration of and timely referral to EI (American Academy of Pediatrics, 2006), however research indicates that anywhere from 30-40% of children with developmental delays are not referred to EI (Jimenez et al., 2014; Lipkin, Macias, Baer Chen, et al., 2020). Even upon referral, many children are not connected to EI evaluation and fail to enroll in EI services, resulting in considerable loss to follow-up (Atkins et al., 2020; Jimenez et al., 2012; Jimenez et al., 2014; Little et al., 2015; Rosenberg et al., 2013). Moreover, differences in surveillance, screening, and referral to EI contribute to the significant disparities increasingly present in young children with developmental delays.

Existing Health Disparities

Racial and ethnic (Barger et al., 2022; Mandell et al., 2007; Mandell et al., 2009; Marshall et al., 2016; Valicenti-McDermott et al., 2012), language (Ratto et al., 2015; Stow & Dodd, 2003; Woolfenden et al., 2015), and socioeconomic (Bradley & Corwyn, 2002; Daniels & Mandell, 2013; Durkin et al., 2017; Fountain et al., 2011; Mandell et al., 2005; Marshall et al., 2016; Mazurek et al., 2014; Wittke & Spaulding, 2018) disparities exist in prevalence, developmental surveillance and screening, identification, and diagnosis of delays, timely referral to EI, and participation in EI (Bilaver et al., 2020; Donohue et al., 2017; Zuckerman et al., 2014). Children from disadvantaged racial, ethnic, and socioeconomic backgrounds are more likely to be identified at older ages and receive fewer EI services, reducing opportunities for treatments that can improve cognitive, physical, and emotional outcomes (Adams & Tapia, 2013; Mandell et al., 2009). For example, Black and Latinx children were 30% less likely to be diagnosed with ASD than White children (Shaw et al., 2021). In terms of utilization of services, Black children were five times less likely to participate in EI services as compared to White children at 24 months (Feinberg et al., 2011). Black and Latinx children were also more likely than White children to have unmet therapy needs (Magnusson & Mistry, 2017). With regard to disparities by language, Latinx children with Spanish-speaking mothers were diagnosed later and experienced longer delays between the time when mothers expressed concerns to when a child was diagnosed compared with White children whose mothers primarily speak English (Ratto et al., 2015). The prevalence of developmental delays was higher among children living below the federal poverty level (20% vs. 15% for those above the federal poverty level) and among children living with mothers with lower educational attainment (17% for those with less than a high school degree vs. 14% of college graduates; Zablotsky et al., 2019). A review examining factors associated with age at diagnosis for ASD found that children from higher income and higher education households were diagnosed with ASD at earlier ages (Daniels & Mandell, 2013). Further, infants from low income families were less likely to receive adequate EI services (Fefferman et al., 2017).

Multiple factors likely contribute to these racial, ethnic, and socioeconomic disparities including differences in trust in healthcare providers and the healthcare system, clinician surveillance, screening, and referral practices, parent and caregiver knowledge about developmental milestones, bias and discrimination, and structural racism (Bilaver et al., 2020; Rosenberg et al., 2008; Zuckerman et al., 2014). A retrospective chart review indicated those with less than a high school education may be less likely to verbally share concerns about their child's development, despite marking a concern on a screening instrument (Eremita et al., 2017). Similarly, there are racial differences in parents and caregivers reporting concerns about children's behavior prior to receiving an ASD diagnosis as Black parents reported fewer autism concerns compared to White parents (Donohue et al., 2017). Clinician biases also play a role as clinicians serving lower income families reported differences in screening (Arunyanart et al., 2012) and not feeling confident or comfortable discussing concerns identified through screening with families (Moore et al., 2017). Clinicians also overestimated how often they discuss screening with families and make referrals for EI services (Bright et al., 2019) contributing to existing disparities. Notably, research suggests that racial, ethnic, and socioeconomic related disparities might be attributed to differences in parent or caregiver knowledge about developmental milestones, as parents and caregivers of higher educational attainment and higher income have been found to be more knowledgeable of developmental milestones (Mandell et al., 2009; Marshall et al., 2016; Ratto et al., 2015). Given the extensive health disparities in developmental delays, there is a considerable need for interventions to help identify children with developmental delays at earlier ages.

Interventions to Improve Early Identification

Various structural, interpersonal, and individual-level interventions have been employed to improve timely identification of developmental delays, referral to and enrollment in EI services. Structural-level interventions include developing strategies to improve routine developmental surveillance and screening (Berry et al., 2014; Bright et al., 2019), involving early childhood care providers in the developmental surveillance and screening processes (Kiing et al., 2019), and parent or caregiver completed screening instruments in clinic waiting rooms (Rydz et al., 2006). Interpersonal and individual-level interventions tend to focus on training and raising awareness. Examples include trainings for childcare providers to increase knowledge of developmental milestones (Abercrombie et al., 2021; Chödrön, Barger, et al., 2021; Chödrön, Pizur-Barnekow, et al., 2021) health coaching to empower parents and caregivers of children with developmental delays (Majnemer et al., 2019), video-based decision aids to help parents and caregivers identify delays (Jimenez et al., 2017), and training parents and caregivers to provide interventions to improve EI outcomes (Roberts & Kaiser, 2015). There is also tremendous potential for utilizing digital platforms to disseminate interventions for developmental delays (McClure et al., 2018). For instance, online resources and peer-support groups offer parents and caregivers access to much needed information and social support

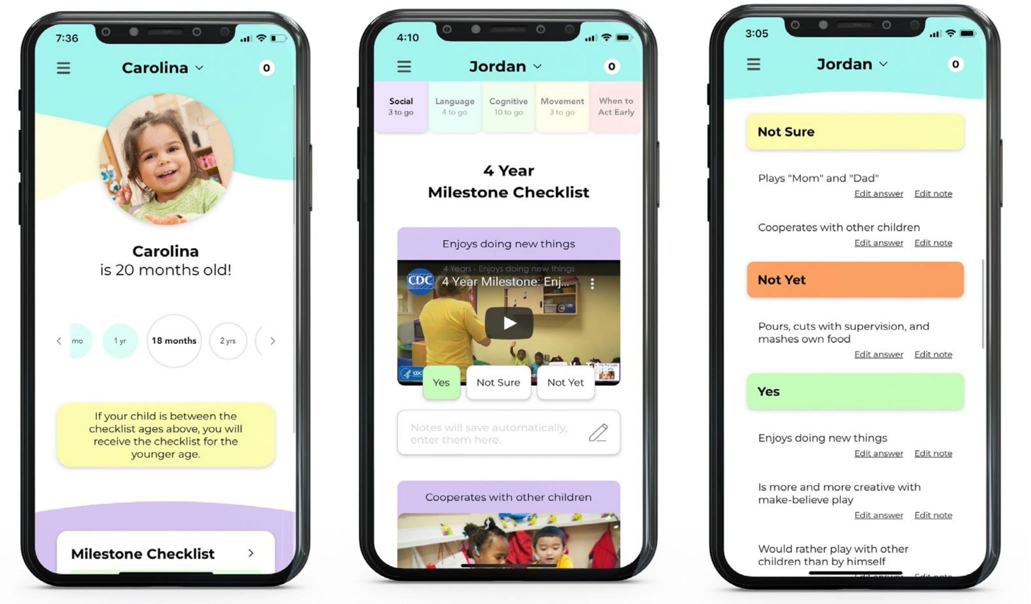
(Majnemer et al., 2019). However there is a dearth of research about which parents and caregivers would access and use behavior change apps for developmental delays, and therefore more research is needed to understand DBCI use among parents and caregivers from groups experiencing disadvantages(McClure et al., 2018).

***Learn the Signs. Act Early*. Program**

The CDC's *Learn the Signs. Act Early*. (LTSAE) program (www.cdc.gov/ActEarly) aims to improve early identification of developmental delays and disabilities, so children and their families can get the services they need as early as possible. The program provides free, high-quality, research-informed tools and resources to support parent-engaged developmental surveillance for all children (Centers for Disease Control and Prevention, 2021a). Priority program audiences are parents and caregivers (particularly those from groups experiencing disadvantages), early childhood educators and care providers, clinicians, and others who work with young children (Centers for Disease Control and Prevention, 2021a). Established in 2004, LTSAE offers free developmental milestone checklists, children's books, brochures, posters, trainings, and fact sheets to support developmental surveillance and early identification (Abercrombie et al., 2021; Zubler et al., 2022). The 12 age-appropriate developmental milestone checklists align with the AAP recommended well-child visit schedule (ages 2, 4, 6, 9, 12, 15, 18, 24, 30 months, 3, 4, and 5 years) and use milestones most children (at least 75%) would be expected to achieve by specific well-child visit ages. Initially developed based on AAP's *Caring for Your Baby and Young Child: Birth to Age 5* (fifth edition) and *Bright Futures: Guideline for Health Supervision of Infants, Children, and Adolescents* (third edition), the checklists use plain language, are organized by developmental domains, and are easy to observe in natural settings. The checklists are developmental surveillance tools intended to support (not replace) developmental screening (Zubler et al., 2022). In 2017, LTSAE turned the milestone checklists into an interactive mobile app called *CDC's Milestone Tracker* (Figure 1).

Figure 1

Images of CDC's Milestone Tracker App (Version January 2022)



If your child is between the checklist ages above, you will receive the checklist for the younger age.

Notes will save automatically, enter them here.

Plays "Mom" and "Dad"

Cooperates with other children

Pours, cuts with supervision, and mashes own food

Enjoys doing new things

Is more and more creative with make-believe play

Would rather play with other children than by himself

CDC's Milestone Tracker App

CDC's Milestone Tracker (www.cdc.gov/MilestoneTracker) is a behavior change app to help families (particularly those from groups experiencing disadvantages) learn about child development, track developmental milestones, and take early action if they have concerns about their child's development. The app was developed by the LTSAE program and the *Computing for Good* program at the Georgia Institute of Technology (Muñoz & Arriaga, 2015). Formative research to develop the app was conducted with a diverse group of parents and caregivers with lower incomes recruited from Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) settings (Muñoz & Arriaga, 2015). *CDC's Milestone Tracker* is free to download in English and Spanish from the Apple and Google Play app stores. As of January

2022, the app includes 10 interactive, age-appropriate developmental milestone checklists (for ages 2, 4, 6, 9, 12, 18 months, 2, 3, 4, and 5 years) with photos and videos for each milestone that can be observed in natural settings as well as space to enter notes. Users can create multiple, customized profiles for each child (by entering a child's birthdate and gender) and receive corrected checklists for children born prematurely based on a child's developmental age. The app generates a milestone summary that parents and caregivers can share with clinicians and other early childhood providers. There are also appointment reminders and tips and activities to support early development built into the app. The app sends notifications reminding users to complete milestone checklists, schedule appointments, and provides recommendations (e.g., discuss concerns with their child's doctor and any incomplete or missing milestones). There have been more than 1 million times downloads of *CDC's Milestone Tr*acker (as of January 2022).

The *CDC's Milestone Tracker* app facilitates behavior change through seven BCTs such as demonstration of the behavior and prompts or cues (Figure 2) and their 14 MoAs (see Table 1 for BCTs and their MoAs classifed using the BCT Taxonomy version 1; Carey et al., 2018; Michie et al., 2021; Michie et al., 2013). The BCTs are one of the processes through which engagement with the app occur, and the BCTs are operationalized via their corresponding MoAs. For example, the milestone photos and videos (app feature) use demonstration of the behavior and social comparison (BCTs) to influence beliefs about capabilities to conduct developmental surveillance, change norms and subjective norms about developmental surveillance, and enhance social learning/imitation related to developmental milestones (MoAs). As such, behavioral engagement with the app is possible due to the BCTs and their MoAs.

Prior research and evaluations of the LTSAE materials (primarily paper versions of the milestone checklists) in various settings demonstrate positive outcomes including increased

Figure 2

Behavior Change Techniques (BCTs) within CDC's Milestone Tracker App

Note. App version January 2022. BCTs are classified using Michie et al (2013) BCT Taxonomy (v1).
* BCTs 6.1 and 6.2. are embedded within the app feature Milestone Photos and Videos but since there is no menu option for Milestone Photos and Videos, arrow points to the Milestone Overview menu option instead.

Table 1

CDC's Milestone Tracker App, BCTS, and Their Mechanisms of Action

MT app Feature	Description [a]	BCT [b]	MoA (by strength of evidence/link) [c]
Children & Add Child	Add a photo and enter personalized information about your child or multiple children	None	N/A
Milestone Checklist/ Milestone Tracker	Track your child's developmental progress by looking for important milestones using an interactive, illustrated checklist	1.2. Problem solving 2.1. Monitoring of behavior by others without feedback	1.2. beliefs about capabilities [L], behavioral regulation [L], skill [I], environmental context and resources [I] 2.1. reinforcement [I], social influences [I]
When to Act Early	Know when it's time to "act early" and talk with your child's doctor about developmental concerns	4.1. Instruction on how to perform a behavior	4.1. knowledge [L], skill [L], beliefs about capabilities [L]
My Child's Summary/ Milestone Summary	Get a summary of your child's milestones to view, and share with or email to your child's doctor and other important care providers	4.1. Instruction on how to perform a behavior 15.1. Verbal persuasion about capability	4.1. knowledge [L], skill [L], beliefs about capabilities [L] 15.1. beliefs about capabilities [L]
Tips and Activities	Support your child's development at every age	4.1. Instruction on how to perform a behavior 15.1. Verbal persuasion about capability	4.1. knowledge [L], skill [L], beliefs about capabilities [L] 15.1. beliefs about capabilities [L]
Milestone Photos and Videos	Know what each milestone looks like so that you can better identify them in your own child	6.1. Demonstration of the behavior 6.2. Social comparison	6.1. beliefs about capabilities [L], social learning/imitation [L], skill [I] 6.2. social influences [L], norms [L], subjective norms [L], feedback process [L], social/professional role and identity [I]
Appointments	Keep track of your child's doctors' appointments and get reminders about recommended developmental screenings	7.1. Prompts/cues	7.1. memory, attention and decision processes [L], environmental context and resources [L], behavioral cueing [L], reinforcement [I]
Notifications	Notifications (e.g., reminders to complete checklists)	7.1. Prompts/cues	7.1. memory, attention and decision processes [L], environmental context and resources [L], behavioral cueing [L], reinforcement [I]

Note. BCTs = behavior change techniques; I = inconclusive link result; MoA = mechanisms of action; MT = *CDC's Milestone Tracker* app; L = link.

[a] Descriptions of app features from www.cdc.gov/MilestoneTracker. [b] BCTs are classified using Michie et al (2013) BCT Taxonomy (v1). [c] Theory and techniques tool for linking BCTs and their MoA available from https://theoryandtechniquetool.humanbehaviourchange.org/.

awareness and knowledge of developmental milestones, confidence in ability to discuss concerns with a clinician about a child’s development, and positive attitudes regarding developmental surveillance (Abercrombie et al., 2021; Campbell et al., 2019; Chödrön, Barger, et al., 2021; Chödrön, Pizur-Barnekow, et al., 2021; Daniel et al., 2009; Farmer et al., 2022; Gadomski et al., 2018; Gallagher et al., 2019; Graybill et al., 2016; Raspa et al., 2015). The only published evaluation of *CDC’s Milestone Tracker* app was a series of usability studies during the design phase of app development of an initial version called “ActEarly” conducted with 23 English and Spanish-speaking adults (Armenta et al., 2019). The evaluation identified pain points (i.e., potential problems users may experience) to develop and test a redesigned prototype of the app (Armenta et al., 2019). Accordingly, further evaluation of *CDC’s Milestone Tracker* app would be beneficial.

Motivation and Research Positionality Statement

I approached this work from the perspective of a behavioral scientist and acknowledge there are interdisciplinary differences in how engagement is conceptually defined and measured. I did not attempt to reconcile these differences and do not propose that a cross-discipline standardized framework for measuring engagement with behavior change apps is the solution; rather I sought to investigate a gap I first encountered a decade ago. While working at the CDC conducting research to develop and evaluate national HIV health communication campaigns, I observed a fundamental shortcoming in how DBCI are evaluated. At the CDC, I was responsible for managing evaluation activities including monthly data collection for thirteen campaigns. While I could access a mountain of data, what these data illustrated and how to translate the information into recommendations for improving our campaigns was puzzling. As a team, we struggled to determine what data to capture and how data could inform campaign implementation. Often, we could only report number of “page views” and “likes,” neither of which were sufficient indicators of actual engagement with our HIV prevention content. These metrics were of limited use for informing future communication efforts since they did not provide information about how or why people interact with content, both of which are important evaluation components. The scale of this problem reaches far beyond the projects I was involved in at the CDC. The development of novel methods to improve measurement and evaluation of DBCI are necessary to ensure intervention success and improve public health.

As a parent of a toddler who was recently identified with a developmental delay and as someone committed to conducting research that examines social determinants of health, I am particularly invested in this dissertation project, which focuses on a tracking tool to help families identify developmental delays and disabilities in young children. I know firsthand the importance of developmental surveillance and screening to identify delays early and understand the significance of resources like *CDC’s Milestone Tracker* app to learn about developmental milestones and empower parents and caregivers to discuss potential concerns and seek services if needed.

It is important to note that the focus of this dissertation research was to explore behavior change apps for groups experiencing disadvantages, as opposed to behavior change apps for people from certain racial, ethnic, or socioeconomic backgrounds. As others have established, race and ethnicity are imperfect measures that should be examined along with other sociodemographics (Flanagin et al., 2021). Therefore, this dissertation project aimed to study inequities in health, research, and behavior change app development, use, and evaluation using a study population that is more inclusive and does not conflate race and SES. Further, because the language and terminology we use have meaning and power, I use the term “groups experiencing disadvantages" as opposed to vulnerable, minority, marginalized, or underserved which are all vague and problematic terms (American Psychological Association, 2021; Centers for Disease Control and Prevention, 2021b; Flanagin et al., 2021; Walker & Fox, 2018; Wrigley & Dawson, 2016). Recent inclusive language guidance from the CDC and American Psychology Association recommend person-first language and being as specific as possible when describing subpopulations (American Psychological Association, 2021; Centers for Disease Control and Prevention, 2021b). See the methods sections of Chapters 2 and 3 for how I defined groups experiencing disadvantages for study inclusion in the systematic review and the sample to recruit for the surveys with parents and caregivers of young children.

Study Purpose

This dissertation explored how engagement with behavior change apps for groups experiencing disadvantages is conceptually defined and how measurement can be improved. Driven by gaps in the research discussed above (see Table 2 for a summary of main research gaps and related research questions), the goal of this dissertation was to conduct formative research to understand how to better measure engagement with behavior change apps for groups experiencing disadvantages. This was accomplished by reviewing the literature to synthesize existing conceptual definitions and evaluation methods and practices, followed by surveys to understand parent and caregiver engagement with *CDC's Milestone Tracker* app and interviews to explore pediatric clinician perceptions of engagement with *CDC's Milestone Tracker* app. The overarching question guiding this research was how to better measure engagement with behavior change apps for groups experiencing disadvantages? Related research questions include:

RQ1: How is engagement with behavior change apps for groups experiencing disadvantages conceptually defined?

RQ2: How is engagement with behavior change apps for groups experiencing disadvantages measured?

RQ3: What methods are used to measure engagement with behavior change apps for groups experiencing disadvantages?

RQ4: How do parents and caregivers of young children engage with the *CDC's Milestone Tracker* app?

RQ5: What features of the *CDC's Milestone Tracker* app are engaging for parents and caregivers of young children?

RQ6: What are FQHC pediatric clinician perceptions of the potential utility of the *CDC's Milestone Tracker* app?

RQ7: What methods and measures should be used to evaluate engagement with the *CDC's Milestone Tracker* app?

Thus, the specific aims of the proposed dissertation research seek to:

Aim 1: Synthesize existing conceptual definitions and current evaluation methods and practices of engagement with behavior change apps for groups experiencing disadvantages.

Aim 2: Explore engagement behaviors and identify factors that influence engagement with *CDC's Milestone Tracker* app among parents and caregivers with young children from groups experiencing disadvantages.

Aim 3: Explore perceptions regarding engagement with *CDC's Milestone Tracker* among pediatric clinicians at FQHCs.

Table 2

Research Gaps, Related Research Questions, and Potential Contributions of this Dissertation Project

	Research Gaps	Related Research Questions	Potential Contributions
1	High attrition for use of behavior change apps makes it difficult to evaluate effects, resulting in a need to study why and how people use apps and reasons for nonadherence to understand how to keep users engaged	RQ4, RQ5, and RQ6	Theoretical/conceptual
2	Limited engagement research with diverse populations (e.g., groups experiencing disadvantages)	RQ1, RQ2, RQ3, RQ4, RQ5, and RQ6	Empirical
3	Little consensus on how to conceptualize and measure engagement with behavior change apps	RQ1, RQ2, and RQ3	Theoretical/conceptual, Empirical
4	Few engagement scales and frameworks are specific to health and most do not include behavior change as a component of behavioral engagement	RQ4, RQ5, and RQ7	Methodological
5	Only one published evaluation of the *CDC's Milestone Tracker* app and more research is needed to evaluate engagement with the app [a]	RQ4, RQ5, and RQ6	Empirical

Note. Research gaps included in this table are the main gaps relevant to this dissertation project.

[a] Evaluation of *CDC's Milestone Tracker* app is a secondary goal of this dissertation project.

Dissertation Chapter Outline

Chapter 2 of this dissertation explores the variability in conceptual definitions of engagement through a systematic review of the literature to synthesize conceptual definitions, measures, and methods of engagement based on common features described in the literature specific to behavior change apps for groups experiencing disadvantages across disciplines (Aim 1). This review of the literature on behavior change apps for groups experiencing disadvantages (N = 21 articles included in the review) helped to identify existing conceptual definitions of engagement and current evaluation practices and provided insight into the current intervention landscape for this specific population of interest. Conclusions from the systematic review provided context and informed measures and questions for the second and third studies of this dissertation described in Chapter 3 and Chapter 4.

The third chapter of this dissertation explores parent and caregiver's perceptions of engagement with *CDC's Milestone Tracker* to provide insight into how engagement with behavior change apps for groups experiencing disadvantages can be measured (Aim 2). This study used surveys conducted at two time points (baseline and one-month follow-up) with parents and caregivers of young children (birth to age five) from groups experiencing disadvantages (N = 72) to explore how parents and caregivers engage with the *CDC's Milestone Track*er app and identify the multidimensional factors (behavior, cognition, and affect) that influence subjective engagement. This study also included a secondary analysis of *CDC's Milestone Tracker* app usage data as an objective measure of engagement. These data were integrated with the conclusions from the systematic review (Chapter 2) and interviews with pediatric clinicians to suggest measures to include in an engagement measurement framework to evaluate engagement with the *CDC's Milestone Tracker* app (described in Chapter 5).

Chapter 4 of this dissertation explored perceptions of engagement with *CDC's Milestone Tracker* app among pediatric clinicians at FQHCs (Aim 3). Twenty individual in-depth interviews were conducted using a semi-structured interview guide designed to capture perceptions of the three dimensions of engagement. This study was meant to complement the surveys conducted with parents and caregivers to understand perceptions about engagement from a secondary audience of *CDC's Milestone Tracker*.

Finally, Chapter 5 builds on the three studies described in Chapters 2, 3, and 4 to integrate findings and suggest methods and measures (i.e., metrics) to include in an engagement measurement framework to evaluate engagement with the *CDC's Milestone Tracker* app. This applied research culminates in a better understanding of engagement with behavior change apps for groups experiencing disadvantages and recommendations for how engagement with behavior change apps can be evaluated. Findings from this dissertation have potential to guide subsequent behavior change app evaluation and design activities to advance public health initiatives.

Significance of the Study

Research on engagement is of great interest to many in the health behavior change intervention community across industry, government, and academia. Recent calls to action encourage interdisciplinary teams of investigators to study "a science of engagement" (Pagoto & Waring, 2016) by conducting research to discover effective strategies to generate "meaningful patient engagement" (engagement leading to "healthy changes in knowledge, attitudes, and behaviors") that can be scaled for population-level implementation (Pagoto et al., 2019, p. 4). Additionally, this work aligns with two health equity-focused principles recommended by the ConNECT framework, a model to advance health equity in behavioral sciences (Alcaraz et al., 2017). This dissertation actualizes the second principle of ConNECT (fostering a norm of inclusion) by centering diverse participants in the research that meaningfully reflects the inequities of developmental delays and disabilities among young children. The fourth principle of ConNECT (harnessing communication technology) is addressed by conducting research on technology that can be leveraged to increase engagement through interactivity and increase awareness, presenting opportunities to achieve health equity (Alcaraz et al., 2017). This dissertation is a response to these calls and aims to explore key engagement research questions such as how people from groups experiencing disadvantages engage with behavior change apps and what factors are most effective at promoting engagement to facilitate behavior change (Perski et al., 2018). The answers to these questions not only improve evaluation practice, but have implications for how to design better and more equitable health interventions that are more likely to result in desired public health outcomes.

Chapter 2

Assessing Engagement with Digital Behavior Change Interventions for Groups Experiencing Disadvantages: A Systematic Review

Introduction

Digital strategies are increasingly utilized to reach audiences with important health messages and interventions (Atkin & Rice, 2013). DBCI are used to promote healthy behaviors, manage chronic conditions, and provide access to treatment in everyday life. DBCI can engage groups experiencing disadvantages in terms of social and economic attainment with tailored health content and have potential to improve health outcomes (Kreps & Neuhauser, 2010; Kreuter et al., 2013; Lustria et al., 2009). Research has identified strategies for the development, utility, and usability of DBCI based on user experience and design principles (Jaspers, 2009), but less is known about best practices for evaluating DBCI (Perski, Blandford, West, et al., 2017; Short et al., 2018). Recently, the fields of public health, medicine, psychology, and human-computer interaction (HCI) have focused on creating DBCI that are more interactive to bolster engagement and retain users' attention in a crowded media climate; yet little consensus exists on how to conceptually define and measure engagement with DBCI and what the data currently associated with engagement truly represent (Cole-Lewis et al., 2019). Additionally, methods to measure engagement—which include analysis of system usage data, qualitative user experience data, and self-report questionnaires—are applied inconsistently since there is little guidance for which methods are most appropriate to assess engagement (Short et al., 2018).

Variation in engagement conceptual definitions and measurement within the DBCI evaluation landscape is underscored by interdisciplinary differences in how engagement is conceptually defined. For example, behavioral scientists view engagement as synonymous with acceptability, satisfaction, or intervention adherence while computer scientists perceive engagement as increased immersion, attention, and enjoyment with an intervention (Doherty & Doherty, 2018; Klasnja et al., 2011; Short et al., 2018). However where these interdisciplinary differences converge is in the conceptual models of DBCI engagement that suggest that engagement requires an interaction between behavioral and psychosocial (e.g., cognitive and affective) processes (Cole-Lewis et al., 2019; Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017; Yardley, Spring, et al., 2016); however, few models have been validated (Kelders, Kip, et al., 2020; Perski, Blandford, et al., 2019; Perski, Lumsden, et al., 2019; Short et al., 2018). Cole-Lewis et al (2019) pose an engagement definition that reflects this interaction behavioral and psychosocial processes. The authors describe two forms of DBCI engagement: (1) how users interact with features of the DBCI designed to encourage use such as quizzes, games, and chats, and (2) how users interact with the behavior change components within the DBCI such as goal-setting, tracking tools for self-monitoring behaviors, and other behavior change techniques. It is argued that engagement with the health behavior is dependent on engagement with the technology since if users do not interact with DBCI features and enjoy their experience, their exposure to the behavior change components will be limited and less likely to result in desired behavior change. This suggests that both forms of engagement are important to measure with regards to evaluating engagement, but it is unclear if this occurs in DBCI evaluation practice.

Recently, numerous systematic reviews have synthesized the literature to develop conceptual definitions of engagement (Cole-Lewis et al., 2019; Doherty & Doherty, 2018; Flaherty et al., 2021; Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017; Szinay et al., 2020; Yardley, Spring, et al., 2016), identify engagement methods and measures (Alkhaldi et al., 2016; Doherty & Doherty, 2018; Kelders & Kip, 2019; Kelders, Kip, et al., 2020; O'Connor et al., 2016; Perski, Blandford, et al., 2019; Perski, Lumsden, et al., 2019; Short et al., 2018; Simblett et al., 2018), and examine engagement within the context of various health issues as it relates to health outcomes such as mental health (Baumel et al., 2019), weight loss and diet (Delaney et al., 2018; Patel et al., 2021; Sharpe et al., 2017), and cancer (Escriva Boulley et al., 2018). These reviews offer insight into how engagement is conceptually defined, measured, and what methods are used, yet the findings are not necessarily applicable to DBCI for groups experiencing disadvantages, broadly defined as people from minoritized racial and ethnic groups, people with disabilities, people with lower incomes, people who have lower educational attainment, or who have lower SES. These groups may be described as marginalized, disadvantaged, "hard-to-reach," minority, or vulnerable in the literature. Despite calls to explore engagement with diverse populations (Brager et al., 2019; Cole-Lewis et al., 2019; Flaherty et al., 2021), at the time of writing, there are no known reviews on DBCI engagement for groups experiencing disadvantages.

The purpose of this study (Aim 1) was to generate a comprehensive review of the literature on mobile app DBCI (hereafter referred to as behavior change apps) for groups experiencing disadvantages to identify key components of engagement and current evaluation methods and practices. More specifically, this review aimed to describe and synthesize how engagement is conceptually defined, how it is measured, and what methods are used to measure engagement. To narrow the scope, this review focused on one type of DBCI, mobile apps. There is less of a chance for potential confounding factors by studying a single platform. Additionally, given that 85% of adults in the United States own a smartphone, mobile apps have become the most utilized digital channel for health intervention dissemination (Pew Research Center, 2021b). Mobile apps are also increasingly important to reach people who depend on smartphones to access the Internet (who are most often persons with lower incomes and lower educational attainment). Specific research questions for this systematic review are as follows:

RQ1: How is engagement with behavior change apps for groups experiencing disadvantages conceptually defined?

RQ2: How is engagement with behavior change apps for groups experiencing disadvantages measured?

RQ3: What methods are used to measure engagement with behavior change apps for groups experiencing disadvantages?

The results from this systematic review can inform methods and measures to evaluate engagement with behavior change apps for groups experiencing disadvantages.

Methods

Search Strategy

This systematic review was conducted in accordance with Cochrane Collaboration guidelines for conducting systematic reviews (Higgins, Thomas, et al., 2019) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for reporting systematic reviews (Liberati et al., 2009; Moher et al., 2009). A protocol with additional details was pre-registered in the Northwestern University Institutional Repository Arch (Getachew-Smith et al., 2020). An iterative search strategy was used to identify literature on DBCI for groups experiencing disadvantages published through July 2020 (no start date). To reflect the interdisciplinary nature of this research, a comprehensive search was performed across various research databases and sources. Additionally, searches were conducted in the reference lists of included articles for other potentially relevant articles.

Search Terms

A comprehensive search strategy developed in consultation with an academic librarian was used to identify potentially relevant articles. Search terms were informed by prior systematic reviews and other reviews on DBCI and engagement (Alkhaldi et al., 2016; Delaney et al., 2018; Escriva Boulley et al., 2018; Grady et al., 2018; Mclaughlin et al., 2018; Perski, Blandford, West, et al., 2017; Sawesi et al., 2016; Sharpe et al., 2017; Yang et al., 2019). Search queries were conducted using Boolean operators and controlled vocabulary or index terms, amended as necessary for each database (where applicable). Preliminary searches were conducted to refine search terms. Searches included six a priori concepts: “mobile apps” *and* “health” *and* “intervention” *and* “engagement” *and* “diverse” or “disadvantaged” (see Appendix A for example search query).

Databases

The following eight databases were queried on the same day: PubMed, Embase, PsycINFO, Communication Source, Web of Science, Scopus, the Association for Computing Machinery (ACM) Digital Library and IEEE Xplore Digital Library. These databases were selected because they have been used in previous systematic reviews and other reviews on DBCI and engagement (Alkhaldi et al., 2016; Escriva Boulley et al., 2018; Grady et al., 2018; Perski, Blandford, West, et al., 2017; Sawesi et al., 2016; Sharpe et al., 2017) and are considered essential to conducting a systematic review (Higgins, Thomas, et al., 2019). The ACM and IEEE libraries were searched for HCI-related conference proceedings. Searches were repeated one day later for accuracy and exported as .csv, .txt, and .ris files and imported into the citation manager EndNote X9 (Clarivate Analytics). De-duplication was conducted using both the EndNote compare function and hand searching.

Inclusion and Exclusion Criteria

Published articles with no restriction on country of origin or publication date were included in this review. Articles were included that (1) described mobile app (i.e., disseminated via smartphone or tablet) DBCI, (2) with a specified health outcome, (3) for groups experiencing disadvantages (defined as priority population or population recruited included a significant [greater than 50%] proportion of individuals who were at least one of the following: people with lower incomes, people who have lower educational attainment or who are considered low SES [as defined by the authors of the article], people from minoritized racial and ethnic groups, or people with disabilities), (4) assessed user engagement or interactivity with the app, (5) published in English, and (6) peer-reviewed. Articles were excluded for the following reasons: (1) published in a language other than English, (2) not peer-reviewed (i.e., white and grey literature), (3) technical report, dissertation, thesis, book, or book chapter, (4) was a non-empirical study (i.e., did not generate original data), (5) the article did not describe an intervention, (6) the article described an intervention to improve participant recruitment, participation, and retention or was a data collection tool or study protocol, (7) the intervention did not specify a health outcome and/or was not designed to promote healthy behaviors, manage chronic conditions, provide access to treatment, or improve health outcomes, (8) the intervention was not disseminated via digital channels (e.g., was conducted in-person), (9) the intervention was not a smartphone or tablet mobile app (i.e., was a website, wearable such as Fitbits and Apple watch for Fitness tracking or other web-based platform), (10) was a social media-based intervention (e.g., disseminated via Facebook, Twitter, Instagram, Snapchat) or disseminated using messaging platforms (e.g., WhatsApp, GroupMe, Facebook Messenger or other SMS text intervention) or utilized virtual reality (VR), (11) the intervention was not explicitly for groups experiencing disadvantages (defined as priority population or population recruited included a significant [greater than 50%] proportion of individuals who were at least one of the following: people with lower incomes, people who have lower educational attainment or who are considered low SES [as defined by the authors of the article], people from minoritized racial and ethnic groups, or people with disabilities), or (12) the article did not assess engagement or interactivity with an app. While most these criteria were included in the protocol, some were determined post hoc and during the screening and full-text review processes.

Screening and Review

A multi-phased process was used to screen and review citations for inclusion. De-duplicated citations were uploaded to Rayyan (https://rayyan.qcri.org), a free web and mobile application that helps collaborators quickly sort citations during the initial title and abstract screening process (Ouzzani et al., 2016). Inclusion and exclusion criteria keywords were entered into Rayyan for use during sorting. Three researchers (HGS, NA, and AB) conducted an independent title and abstract screening in Rayyan to pilot the screening process using a decision tree we created based on the inclusion and exclusion criteria (Wilson, 2009). Following this calibration process, the decision tree was revised, and the three researchers conducted an independent title and abstract screening in Rayyan by sorting citations into 12 categories based on the exclusion criteria. At least two of the three researchers reviewed each citation included in the title and abstract screening process. The three researchers met to discuss discrepancies and resolved them using (1) a consensus-based approach (for citations that were screened by all three researchers) and (2) a third researcher to resolve conflicts (for citations that were screened by two researchers). Full-text was located and independently reviewed for eligibility by at least two researchers using a modified version of the decision tree used in title and abstract screening, noting the reason for exclusion in Rayyan. The one addition to the decision tree was to exclude articles that did not specify whether the intervention intended to change health behaviors (i.e., was not a behavior change intervention, did not describe behavior change theory or behavior change techniques such as goal-setting, prompts/cues, rewards based on the Behavior Change Techniques taxonomy; Michie et al., 2013). The three researchers met to discuss and compare full-text review and discrepancies were reviewed by the third researcher who did not initially screen the full-text, and discrepancies were resolved by consensus.

Data Extraction and Analysis

A data extraction form adapted from other DBCI engagement systematic reviews (Alkhaldi et al., 2016; Escriva Boulley et al., 2018; Grady et al., 2018; Perski, Blandford, West, et al., 2017; Short et al., 2018) was programmed in Qualtrics software (Version September 2020, Provo, UT). The data extraction form included: intervention overview (e.g., intervention or app name, country of study, funding agency or organization, health issue or health behavior, intervention duration, use of theory or framework, and use of behavior change techniques), details about the study (e.g., aims and objectives, study design, population characteristics, and total sample size), data collection methods and scales used, engagement measures, outcome measures, and the definition of engagement used in the article. The data extraction form was piloted and refined increasing the validity of study findings (Higgins, Thomas, et al., 2019). Two researchers (HGS and NA) divided the articles and independently examined the full-text to extract relevant information. To ensure consistency, both researchers coded two articles independently and then discussed what information was extracted and why for one of the articles. The two researchers also conducted a spot check of a random selection of 10% of the other researcher's set of articles. Any disagreements were resolved by discussion between the two researchers.

In addition to the citations identified through the database searches, a snowball approach was used to identify related citations from the reference lists of articles included in the full text review that did not appear in the searches but met inclusion criteria. Forty related citations were screened for title and abstract review using the same decision tree. Among these, five duplicates were identified. The remaining 35 were screened for full text review following the aforementioned process and included in extraction.

Extracted data were recoded for clarity and consistency. A narrative synthesis on study design, intervention characteristics, measures and data collection methods, and engagement definitions is provided and summarized statistically, where possible, in the following section.

Results

Search Results

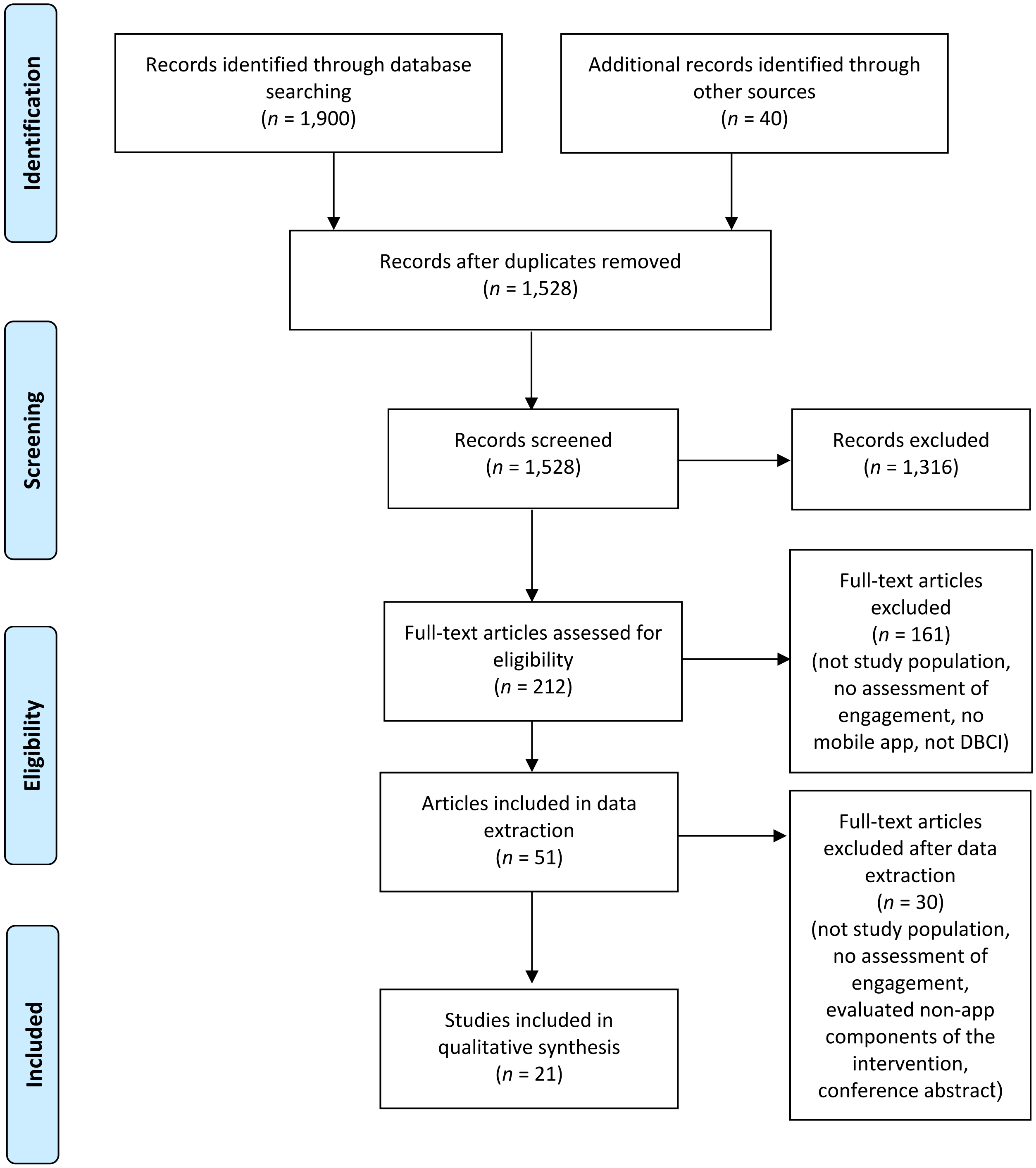
Database queries identified 1,900 citations and an additional 40 citations were identified from the reference lists of articles included in the full-text review. Once duplicate citations were removed (*n* = 412), titles and abstracts of 1,528 potentially relevant citations were screened and 1,316 were excluded based on eligibility criteria. A total of 212 full-text articles were skimmed for eligibility using the revised decision tree, 161 of which were excluded. Fifty-one articles were included in data extraction and an additional 30 articles were excluded because they did not meet eligibility criteria upon reading the entire article. Following the full-text examination, 21 articles published between 2012 and 2020 met inclusion criteria and were included in this review (Figure 3).

Study Characteristics

Table 3 reports article characteristics and Table 4 reports descriptive statistics about studies for included articles. Most articles (*n* = 19, 90.5%) were published in the United States (Aharonovich et al., 2017; Ben-Zeev et al., 2018; Bennett et al., 2018; Chaudhry et al., 2016; Dillingham et al., 2018; Godino et al., 2016; Hartzler et al., 2016; Hayes et al., 2014; Khan et al., 2012; Kim et al., 2015; Mauriello et al., 2016; Mendu et al., 2018; Pratap et al., 2018; Quante et al., 2019; Rung et al., 2020; Saksono et al., 2020; Schaefbauer et al., 2015; Thomson et al., 2017; Wang et al., 2018), in either computing conference proceedings (*n* = 7, 33.3%; Chaudhry et al., 2016; Hakobyan et al., 2016; Hayes et al., 2014; Khan et al., 2012; Mendu et al., 2018; Saksono et al., 2020; Schaefbauer et al., 2015), medical journals (*n* = 5, 23.8%; Aharonovich et al., 2017; Ben-Zeev et al., 2018; Dillingham et al., 2018; Godino et al., 2016; Quante et al., 2019), or medical informatics journals (*n* = 4, 19.0%; Pratap et al., 2018; Rung et al., 2020; Taki et al., 2017; Wang et al., 2018). Apps addressed several health topics including physical activity (*n* = 8, 38.1%; Bennett et al., 2018; Godino et al., 2016; Hartzler et al., 2016; Khan et al., 2012; Kim et al., 2015; Saksono et al., 2020; Schaefbauer et al., 2015; Taki et al., 2017), mental health (*n* = 3, 14.3%; Ben-Zeev et al., 2018; Pratap et al., 2018; Rung et al., 2020), pregnancy (*n* = 2, 9.5%; Hayes et al., 2014; Mauriello et al., 2016), among others (*n* = 8, 38.1%; Aharonovich et al., 2017; Chaudhry et al., 2016; Dillingham et al., 2018; Hakobyan et al., 2016; Mendu et al., 2018; Quante et al., 2019; Thomson et al., 2017; Wang et al., 2018).

Figure 3

Behavior Change Apps for Groups Experiencing Disadvantages Systematic Review Flow Diagram



Median intervention duration was 3.5 months ranging from 0.5 months (or 2 weeks) to 24 months. Participants were predominantly Asian, Black, Latinx, or other race or ethnicity, had lower educational attainment, and reported lower levels of income. Only 5 (23.8%) articles specified disability status of participants (Ben-Zeev et al., 2018; Hakobyan et al., 2016; Hartzler et al., 2016; Pratap et al., 2018; Wang et al., 2018). Sample size ranged from 4 to 404 with a median of 36.

Nearly two-thirds of articles (*n* = 13) described apps that were informed by theoretical frameworks or models such as Bandura's (1986) social cognitive theory (*n* = 5, 23.8%; Godino et al., 2016; Khan et al., 2012; Mendu et al., 2018; Schaefbauer et al., 2015; Thomson et al., 2017), Petty and Cacioppo's (1986) elaboration likelihood model (*n* = 2, 9.5%; Khan et al., 2012; Mendu et al., 2018), Green and Brock's (2000) transportation theory (*n* = 2, 9.5%; Khan et al., 2012; Schaefbauer et al., 2015) and other behavior change theories including self-determination theory (Ryan & Deci, 2000) and the transtheoretical model of behavior change (Prochaska & DiClemente, 1983). The majority of articles (*n* = 16, 76.2%) described specific BCTs or constructs used in the intervention such as goal-setting, self-monitoring of behavior, prompts or cues, social support, instruction on how to perform a behavior, and personalized feedback (Aharonovich et al., 2017; Bennett et al., 2018; Chaudhry et al., 2016; Godino et al., 2016; Hartzler et al., 2016; Hayes et al., 2014; Khan et al., 2012; Kim et al., 2015; Mauriello et al., 2016; Mendu et al., 2018; Pratap et al., 2018; Quante et al., 2019; Saksono et al., 2020; Schaefbauer et al., 2015; Thomson et al., 2017; Wang et al., 2018).

Conceptually Defining Engagement

This review provided limited insight to answer the first research question: how is engagement with behavior change apps for groups experiencing disadvantages conceptually defined? Among the 21 articles in this review, few (*n* = 6, 28.6%) explicitly conceptually defined engagement. Conceptual definitions ranged from more comprehensive to study and health-issue specific. For example, Taki et al (2017) use a modified version of Lalmas et al (2014) definition of engagement: “the quality of the user experience, the positive aspects of their interaction, and their desire to use the app over longer periods of time or repeatedly” (p. 2). Others conceptually defined it as behavior indicating interaction with the mobile app (Kim et al., 2015) or completion of surveys and assessments (Pratap et al., 2018). Finally, three articles simply conceptualized engagement as app usage (e.g., amount of use; Ben-Zeev et al., 2018; Chaudhry et al., 2016; Godino et al., 2016).

Measuring Engagement

The second research question—how is engagement with behavior change apps for groups experiencing disadvantages measured?—was answered by examining how articles reported on engagement measures and outcome measures, as well as the scales used to measure engagement. Like engagement conceptual definitions, there was heterogeneity in engagement measures and was context and health-issue-specific (Table 5). Nearly half (*n* = 10, 47.6%) measured engagement as the number and type of entries or app pages accessed (Bennett et al., 2018; Chaudhry et al., 2016; Dillingham et al., 2018; Godino et al., 2016; Hakobyan et al., 2016; Hayes et al., 2014; Kim et al., 2015; Schaefbauer et al., 2015; Thomson et al., 2017; Wang et al., 2018), while 42.9% (*n* = 9) measured perceived ease of app use and usefulness (Chaudhry et al., 2016; Hakobyan et al., 2016; Hartzler et al., 2016; Hayes et al., 2014; Kim et al., 2015; Mendu et al., 2018; Quante et al., 2019; Saksono et al., 2020; Taki et al., 2017), and 38.1% (*n* = 8) measured interactions with various features of the app

Table 3

Characteristics of Included Articles (N = 21)

	Study			Participants					
Author, Year	Location	Health topic	Intervention duration	*n*	*M* (*SD*) age (years)	Age range (years)	% Race, ethnicity	% Low income, low education, or low SES	% People w/ disabilities
Aharonovich, 2017	USA	Drug and alcohol use	NR	47	51 (7)	34-63	79% B, 17% L, 4% O	93% unemployed/disabled/ public assistance; 81% HS	NR
Ben-Zeev, 2018	USA	Mental health	3 months	163	49 (NR)	NR	65% B, 7% O/M, 27% W	61% <HS	100% serious mental illness [e]
Bennett, 2018	USA	PA, obesity, and nutrition	12 months	351	51 (9)	NR	52% B, 13% L, 5% O/NR, 29% W	67% income <$35k; 30% income <FPL; 15% <HS, 36% HS/GED	NR
Chaudhry, 2016	USA	Chronic kidney disease	6 weeks	9	49 (8)	NR	100% B	100% low income neighborhood; 100% Medicaid insured	NR
Dillingham, 2018	USA	HIV	12 months	77	36 (12)	NR	1% A, 49% B, 8% L, 6% M, 1% NR, 34% W	58% income <50% FPL; 19% <HS, 35% HS, 10% GED; 26% unstable housing	NR
Godino, 2016	USA	PA, obesity, and nutrition	24 months	404	23 (4)	NR	2% AIAN, 24% A, 4% B, 31% L, 9% M, 20% O, 42% W [d]	NR	NR
Hakobyan, 2016	UK	Macular degeneration	6 weeks	9	77 (NR)	65-89	NR	NR	100% age-related macular degeneration
Hartzler, 2016	USA	PA, obesity, and nutrition	Once	15	67 (10)	36-77	33% A, 7% B, 60% W	25% income <$50k	100% type 2 diabetes an

Table 3

Characteristics of Included Articles (N = 21)

	Study			Participants					
Author, Year	Location	Health topic	Intervention duration	*n*	*M* (*SD*) age (years)	Age range (years)	% Race, ethnicity	% Low income, low education, or low SES	% People w/ disabilities
									major depression
Hayes, 2014	USA	Pregnancy	4 months	14	29 (6)	20-41	7% A, 7% B, 64% L, 21% W	93% income <$30k	NR
Khan, 2012	USA	PA, obesity, and nutrition	Once	26	1: 36 (9) [a] 2: 15 (2) [a]	NR	77% B, 23% L	100% public housing	NR
Kim, 2015	USA	PA, obesity, and nutrition	6 months	34	18 (NR)	13-24	15% A, 21% B, 47% L, 6% O/M, 12% NR	62% patients at public hospital in low income neighborhood	NR
Mauriello, 2016	USA	Pregnancy	24 weeks	335	27 (6)	18-45	3% A, 9% B, 65% L, 1% M, 1% O, 21% W	51% unemployed; 28% <HS, 41% HS/GED	NR
Mendu, 2018	USA	Cancer	NR	66	39-47 (NR) [b]	NR	100% L	12% income <$20k, 88% income <$30k	NR
Pratap, 2018	USA	Mental health	12 weeks	345	35 (11)	NR	1% AIAN, 7% A, 7% B, 31% L, 1% O, 53% W	30% income <$20k, 26% income $20k-$40k; 23% cannot make ends meet	100% depression
Quante, 2019	USA	Sleep	2 weeks	27	16 (1)	14-18	67% A, 11% B, 22% L, 15% O, 7% W [d]	100% low/middle income neighborhood	NR
Rung, 2020	USA	Mental health	6 weeks	236	46 (10)	NR	40% B/L/O/M, 60% W	50% income <$50k; 34% unemployed; 61% <HS	NR

Table 3

Characteristics of Included Articles (N = 21)

	Study			Participants					
Author, Year	Location	Health topic	Intervention duration	*n*	*M* (*SD*) age (years)	Age range (years)	% Race, ethnicity	% Low income, low education, or low SES	% People w/ disabilities
Saksono, 2020	USA	PA, obesity, and nutrition	3 months	36	1: 34 (7) [a] 2: 6 (2) [a]	NR	78% B, 22% O	100% income <150% FPL; 100% low income neighborhood or public housing; 33% HS	NR
Schaefbauer, 2015	USA	PA, obesity, and nutrition	12 weeks	20	1: 40-43 (NR) [a,c] 2: 14-14 (NR) [a,c]	1: 31-58 [a] 2: 12-18 [a]	10% A, 10% B, 50% L, 30% W	100% public housing	NR
Taki, 2017	AU	PA, obesity, and nutrition	9 months	225	30 (NR)	NR	NR	76% unemployed; 100% low SES region	NR
Thomson, 2017	USA	Asthma	NR	4	NR	NR	100% B	100% Medicaid insured	NR
Wang, 2018	USA	Type 2 diabetes	6 months	26	56 (NR)	NR	4% AIAN, 4% A, 65% B, 23% L, 27% W [d]	92% income <$20k, 8% income $20k-$30k	31% disabled

Note. AIAN = American Indian/Alaskan Native; A = Asian; B = Black or African American; FPL = U.S. federal poverty level; HS = High school; L = Latinx; low SES = low socioeconomic status; M = Multiracial; NR = Not reported; O = Other race or ethnicity; W = White.

[a] Mean age or age range for two study populations (primary and secondary) reported separately. [b] Range for mean age provided because mean age reported separately for each of the three studies included in the article. [c] Range for mean age provided because mean age reported separately for intervention and control groups. [d] Race and ethnicity reported separately and therefore percentages sum to more than 100%. [e] Serious mental illness includes schizophrenia or schizoaffective disorder, bipolar disorder, and major depressive disorder.

Table 4

Descriptive Statistics about Studies for Included Articles (N = 21)

	n	%
App health topic		
Physical activity, obesity, and nutrition	8	38.1
Mental health	3	14.3
Pregnancy	2	9.5
Other [a]	8	38.1
Study location		
United States	19	90.5
United Kingdom	1	4.8
Australia	1	4.8
Publication type		
Computing conference proceedings	7	33.3
Medical journal	5	23.8
Medical informatics journal	4	19.0
Public health journal	2	9.5
Engineering conference proceedings	2	9.5
Computing journal	1	4.8
Study type [b]		
Feasibility study	7	33.3
Randomized control trial	7	33.3
Usability testing	5	23.8
Pilot or beta testing	5	23.8
Acceptability testing	4	19.0
Other [c]	10	47.6
App informed by theory or framework (Yes)	13	61.9
Behavior change techniques (BCTs) or constructs reported		
Yes	16	76.2
Yes, but not specified	3	14.3
No	2	9.5
Conceptually define engagement (Yes)	6	28.6
Methods used to assess engagement [d]		
System usage data	17	80.9
User experience data (self-report questionnaires)	12	57.1
User experience data (qualitative)	12	57.1
Observation	1	4.8
Outcome measures reported (Yes)	15	71.4

Note. [a] Other health issues include asthma, cancer, chronic kidney disease, drugs and alcohol use, HIV, macular degeneration, sleep, and type 2 diabetes. [b] Study type categories are not mutually exclusive and therefore sum to more than 100%. [c] Other study types include prototype testing, participatory design, efficacy testing, and naturalistic evaluation. [d] Methods used to assess engagement are not mutually exclusive and therefore sum to more than 100%.

(Ben-Zeev et al., 2018; Godino et al., 2016; Hartzler et al., 2016; Pratap et al., 2018; Saksono et al., 2020; Taki et al., 2017; Thomson et al., 2017; Wang et al., 2018).

One-third of articles (*n* = 7, 33.3%) measured engagement as number or proportion of days the app was used (Aharonovich et al., 2017; Chaudhry et al., 2016; Mauriello et al., 2016; Quante et al., 2019; Rung et al., 2020; Schaefbauer et al., 2015; Thomson et al., 2017), while seven articles (33.3%) measured engagement as level of satisfaction and enjoyment (Chaudhry et al., 2016; Hayes et al., 2014; Quante et al., 2019; Saksono et al., 2020; Taki et al., 2017; Thomson et al., 2017; Wang et al., 2018). Five articles (23.8%) measured number of assessments (e.g., surveys and quizzes; Ben-Zeev et al., 2018; Bennett et al., 2018; Dillingham et al., 2018; Mendu et al., 2018; Pratap et al., 2018), and four (19.0%) focused on facilitators and barriers to app use (Chaudhry et al., 2016; Kim et al., 2015; Quante et al., 2019; Saksono et al., 2020). Two articles (9.5%) measured each of the following: time spent in the app (Mendu et al., 2018; Taki et al., 2017), number of app launches or logins (Dillingham et al., 2018; Rung et al., 2020), number of clicks in the app (Mendu et al., 2018; Thomson et al., 2017), and general phone use (e.g., call time, duration, and text length; Pratap et al., 2018; Quante et al., 2019).

Nearly 60% of articles (*n* = 12) measured engagement using measures that were categorized as other measures, since they differed from the groups of measures described above. These included the user's overall experience using the app (Kim et al., 2015), participant's or user's attitudes toward the intervention (Quante et al., 2019; Wang et al., 2018), preferences for app features (Khan et al., 2012), acceptability of the intervention (Wang et al., 2018), perceived importance of the health behavior (Hayes et al., 2014), day of the week and time of day the app was used (Rung et al., 2020), number of push notifications opened (Taki et al., 2017), mobility

data (Pratap et al., 2018), and how families engage with one another as facilitated by the app (Schaefbauer et al., 2015).

Table 5

Engagement Measures of Included Articles (N = 21)

	n	%
Other measures [a]	12	57.1
Number and type of entries or app pages accessed	10	47.6
Perceived ease of app use and usefulness	9	42.9
Interaction with app features	8	38.1
Number of days app used	7	33.3
Level of satisfaction and enjoyment	7	33.3
Number of assessments completed	5	23.8
Facilitators and barriers to use	4	19.0
Time spent in the app	2	9.5
Number of app launches or logins	2	9.5
Number of clicks	2	9.5
General phone use (e.g., call time, duration, text length)	2	9.5

Note. Engagement measures are not mutually exclusive and therefore sum to more than 100%.

[a] Other measures include overall experience using the app, attitudes toward intervention, preferences for app features, acceptability of the intervention, perceived importance of health behavior, time of day app used, number of push notifications opened, mobility data, and how families engage with one another and facilitated by the app.

Nearly 60% of articles (*n* = 12) measured engagement using measures that were categorized as other measures, since they differed from the groups of measures described above. These included the user's overall experience using the app (Kim et al., 2015), participant's or user's attitudes toward the intervention (Quante et al., 2019; Wang et al., 2018), preferences for app features (Khan et al., 2012), acceptability of the intervention (Wang et al., 2018), perceived importance of the health behavior (Hayes et al., 2014), day of the week and time of day the app was used (Rung et al., 2020), number of push notifications opened (Taki et al., 2017), mobility data (Pratap et al., 2018), and how families engage with one another as facilitated by the app (Schaefbauer et al., 2015).

Fifteen articles (71.4%) specified primary and secondary outcome measures, reporting behavioral antecedents, health behavior outcomes, and clinical and health outcomes that were the result of the intervention or app deployment. Among these 15 articles, five reported behavioral antecedents such as changes in knowledge (Chaudhry et al., 2016; Hayes et al., 2014; Mendu et al., 2018), awareness (Hakobyan et al., 2016; Hayes et al., 2014; Schaefbauer et al., 2015), self-efficacy (Chaudhry et al., 2016; Mendu et al., 2018), empowerment (Hayes et al., 2014), motivation (Hakobyan et al., 2016; Hayes et al., 2014), and behavioral intentions (Hakobyan et al., 2016; Mendu et al., 2018). Eight articles described health behavior change outcomes such as healthier diet and eating habits (Chaudhry et al., 2016; Hakobyan et al., 2016; Mauriello et al., 2016; Rung et al., 2020; Schaefbauer et al., 2015), increased physical activity (Rung et al., 2020), reduced drug and alcohol use (Aharonovich et al., 2017), tracking infant's health (Hayes et al., 2014), adoption of stress reduction and mindfulness techniques (Mauriello et al., 2016; Rung et al., 2020), and improved sleep quality (Quante et al., 2019; Rung et al., 2020). Eight articles (38.1%) reported outcome measures based on clinical and health outcomes, including weight change and body mass index (measured by glucose, lipids, and HbA1c levels; Bennett et al., 2018; Godino et al., 2016; Rung et al., 2020; Wang et al., 2018), blood pressure and heart rate (Godino et al., 2016), mental health severity (primarily depression; Ben-Zeev et al., 2018; Pratap et al., 2018; Rung et al., 2020), urine screens for drug use and breathalyzers for alcohol use (Aharonovich et al., 2017), CD4 counts and viral load (Dillingham et al., 2018), and disability status (Pratap et al., 2018).

Four articles (19.0%) reported measures related to app use as primary and secondary outcome measures. These included app use (Schaefbauer et al., 2015), app feasibility, acceptability, and characteristics of app utilization (Rung et al., 2020), level of engagement (i.e., amount of use) with intervention components (Godino et al., 2016), and barriers to app use and intervention fidelity (Hayes et al., 2014).

Finally, few articles (*n* = 3, 14.3%) reported measuring engagement using usability or engagement scales. Hartzler et al (2016) used the system usability scale (Brooke, 1996) and Saksono et al (2020) used the technology acceptance survey (Davis, 1989). Taki et al (2017) was the only article to report an engagement-specific scale which was a scale they developed called the engagement index adapted from the Web Analytics Demystified visitor engagement index (Peterson & Carrabis, 2008).

Methods to Measure Engagement

The final research question—what methods are used to measure engagement with behavior change apps for groups experiencing disadvantages?—was answered by examining study design, study type, and reported methods. Articles included in this review presented randomized control trials (*n* = 7, 33.3%; Aharonovich et al., 2017; Ben-Zeev et al., 2018; Bennett et al., 2018; Godino et al., 2016; Mauriello et al., 2016; Pratap et al., 2018; Wang et al., 2018), feasibility studies (*n* = 7, 33.3%; Aharonovich et al., 2017; Chaudhry et al., 2016; Mendu et al., 2018; Pratap et al., 2018; Rung et al., 2020; Taki et al., 2017; Wang et al., 2018), pilot or beta testing (*n* = 5, 23.8%; Aharonovich et al., 2017; Dillingham et al., 2018; Hartzler et al., 2016; Schaefbauer et al., 2015; Thomson et al., 2017), usability testing (*n* = 5, 23.8%; Chaudhry et al., 2016; Hakobyan et al., 2016; Hartzler et al., 2016; Mendu et al., 2018; Rung et al., 2020), and acceptability testing (*n* = 4, 19.0%; Table 4; Hakobyan et al., 2016; Mendu et al., 2018; Quante et al., 2019; Rung et al., 2020). Ten articles (47.6%) presented other types of studies including prototype testing (Khan et al., 2012), participatory design (Kim et al., 2015), efficacy testing (Godino et al., 2016), and naturalistic evaluation (Saksono et al., 2020), among others.

All 21 articles reported methods to measure engagement; however these methods were heterogenous across articles (Table 4). Most (*n* = 17, 80.9%) reported system usage data (e.g., tracking time spent in the app and number of clicks) to measure engagement (Aharonovich et al., 2017; Ben-Zeev et al., 2018; Bennett et al., 2018; Chaudhry et al., 2016; Dillingham et al., 2018; Godino et al., 2016; Hakobyan et al., 2016; Hartzler et al., 2016; Hayes et al., 2014; Kim et al., 2015; Mendu et al., 2018; Pratap et al., 2018; Rung et al., 2020; Schaefbauer et al., 2015; Taki et al., 2017; Thomson et al., 2017; Wang et al., 2018). More than half of articles (*n* = 12, 57.1%) reported user experience data collected via self-report questionnaires and surveys implemented at different time points ranging from multiple times per day to once or twice during the study (Bennett et al., 2018; Dillingham et al., 2018; Hayes et al., 2014; Kim et al., 2015; Mauriello et al., 2016; Mendu et al., 2018; Pratap et al., 2018; Rung et al., 2020; Saksono et al., 2020; Schaefbauer et al., 2015; Taki et al., 2017; Wang et al., 2018). Twelve articles (57.1%) reported qualitative user experience data collected via interviews and focus groups to assess app satisfaction, enjoyment, facilitators, and barriers to use (Chaudhry et al., 2016; Hakobyan et al., 2016; Hartzler et al., 2016; Hayes et al., 2014; Khan et al., 2012; Kim et al., 2015; Mendu et al., 2018; Quante et al., 2019; Saksono et al., 2020; Schaefbauer et al., 2015; Thomson et al., 2017; Wang et al., 2018). Only one article (4.8%) reported using direct observation (i.e., watching the user) interact with features of the app to measure engagement (Mendu et al., 2018).

Discussion

This systematic review summarizes current literature on evaluating engagement with behavior change apps for groups experiencing disadvantages. Consistent with other reviews (Cole-Lewis et al., 2019; Yardley, Spring, et al., 2016), among the 21 articles included, few articles reported engagement conceptual definitions. Further, only one conceptual definition included psychosocial constructs (both cognitive and affective) of engagement and all other conceptual definitions focused on engagement behaviors. Other reviews have identified engagement as a multidimensional process (Flaherty et al., 2021; Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017). Therefore, only including the single dimension of behavioral engagement means many studies in this review may not have captured engagement in its entirety. Given the variation in conceptualizations of engagement, Cole-Lewis et al (2019) and others (e.g., O'Brien & Toms, 2008) recommend ensuring engagement is defined by interactions with DBCI features as well as the desired health behavior, which would enhance application to other interventions. All articles described how engagement was measured and which methods were used to measure engagement, revealing some indication of useful methods and measures. The sizable heterogeneity in measures, methods, and outcomes prevents researchers from identifying effective engagement patterns among behavior change apps for groups experiencing disadvantages. This finding is consistent with other reviews on DBCI (Alkhaldi et al., 2016; Escriva Boulley et al., 2018). Most engagement measures were unvalidated, illuminating the need to validate engagement measures and scales.

Interestingly, despite rarely including psychosocial constructs in the conceptual definitions, articles described a fair number of psychosocial measures. These psychosocial measures included: perceived ease of app use and usefulness, level of satisfaction and enjoyment, facilitators and barriers to app use, attitudes toward the intervention, preferences for app features, acceptability of the intervention, and perceived importance of the health behavior. To truly understand engagement, it may be necessary to go beyond engagement behavior to include measures that assess cognitive (e.g., attention, questions, concerns, and knowledge) and affective processes (e.g., satisfaction and enjoyment).

Notably, the results of this review show most studies used system usage data, which is unsurprising given that it is an objective measure of engagement that is readily available. However, while system usage data may be a beneficial form of passive data collection in that it reduces participant burden (Short et al., 2018), it also can drain device battery and use a great deal of cellular data, which can be detrimental for those with lower income since cellular data is expensive and stable and consistent Internet connection may be limited (Lee et al., 2022). These barriers, known as "connection maintenance costs" or the hidden costs (e.g., financial, time, and energy resources) that are associated with using digital health tools, can widen existing inequalities (Lee & Viswanath, 2020). It will be important to identify the most effective combination of engagement measures while maintaining reduced user burden and accounting for differences in data access. Notably, a systematic review on factors that influence uptake and engagement with smartphone apps for health and well-being found substantial heterogeneity in factors related to engagement (Szinay et al., 2020). Given that only 19% of articles (*n* = 4) reported app use or engagement as a primary or secondary outcome, there is a need for more research in this area. Finally, what qualifies as "meaningful engagement" may be dependent on the priority audience, health issue, setting, and outcome(s) of interest.

Future Directions

This review identified gaps in evaluating engagement with behavior change apps for groups experiencing disadvantages. While it is important to conduct controlled trials to develop and evaluate interventions, there is a need to conduct more practical evaluations to understand how users from groups experiencing disadvantages engage with behavior change apps in an ecological way (i.e., outside of controlled settings with researcher oversight and where they are incentivized to use an app). Extending the Mohr et al (2015) trials of intervention principles framework, which tests theoretical constructs of DBCI, may help to increase generalizability of findings across studies and provide more accurate assessments of engagement. Similar to Li et al (2022), it is necessary to explore dose-response effects of engagement on outcomes of interest (particularly behavioral antecedents, health behaviors, and health and clinical outcomes). It is crucial to know how much engagement is necessary to produce desired changes in health behavior and health outcomes and for which groups of people or types of users. Studying the relationship between engagement and desired intervention outcomes will increase the likelihood that DBCI are beneficial to improving the health and well-being of those disproportionately affected by health inequities.

Implications for Dissertation Research

Conclusions from this review provided context and informed survey measures and interview questions for the second study described in Chapter 3 (Aim 2) and third study described in Chapter 4 (Aim 3). Connections to each chapter are described below in more detail.

Chapter 3: Exploring Parent and Caregiver Engagement with CDC's Milestone Tracker App

Methods and measures that were most frequently reported in the articles included in the systematic review were selected for the second study of this dissertation, exploring engagement with *CDC's Milestone Tracker* app. I considered methods that were frequently used, described in enough detail to replicate, and feasible to implement given logistical constraints. Using this approach, two findings emerged: (1) the proposed mixed methods procedure was likely unnecessary and (2) the second study should generate user experience data to assess the psychosocial aspects of engagement. Originally, the proposed methods for this study included two surveys, an individual in-depth interview, eye tracking assessment, and retroactive think-aloud session. While some of these planned methods were restricted by the COVID-19 pandemic, the in-person eye-tracking assessment and retrospective think-aloud session were deemed unnecessary since only one article included in the systematic review used direct observation to assess engagement. Instead, the study was designed to collect user experience data via self-report surveys conducted at two time points (baseline and one-month follow-up) similar to Rung et al (2020), Taki et al (2017), Thomson et al (2017), and nine others articles included in the systematic review. The eye tracking assessment was replaced with heat maps in the Qualtrics survey and the some of the interview and think aloud questions were added into the surveys as open-ended response items.

To determine the measures used in the second study, I used the final data extraction spreadsheet to sort the various engagement measures reported from articles in the systematic review and identify constructs that should be measured (e.g., perceived ease of use and usefulness and level of satisfaction and enjoyment). Then I revisited the full-text (or supplemental materials to the full-text) to generate a list of potential survey items that articles in the review used to measure the constructs deemed most salient to my study. Thus, many of the measures included on the two surveys were derived from studies included in this review and modified for *CDC's Milestone Tracker* app. For example "This app is easy to use" (Hakobyan et al., 2016; Hartzler et al., 2016; Saksono et al., 2020; Taki et al., 2017) and "I was concerned about the Internet data usage on my phone when using the app" (Quante et al., 2019; Taki et al., 2017) were modified versions of items from the review that I added to the follow-up survey to assess perceived ease of use and barriers to use.

Similar to the majority (80.9%) of articles in this review, the second study also examined system usage data from the *CDC's Milestone Tracker* app to assess engagement. I mapped the list of available metrics from the LTSAE team to the different ways engagement was measured by articles in the review to determine which app usage data to analyze. For example, the number of milestone checklist ages viewed and completed served as a proxy for intervention dose and interactions with app features (Hartzler et al., 2016; Mauriello et al., 2016; Pratap et al., 2018). The number of app launches represents frequency of app use similar to the total and average number of days used (Rung et al., 2020). Data included in Taki et al (2017) engagement index were particularly helpful to inform data examined in the second study. Unfortunately, it was not possible to apply the full index to *CDC's Milestone Tracker* due to the limited data LTSAE currently has access to and the way the data are aggregated. However, individual indices such as interaction index (number of push notifications opened) were explored.

Chapter 4: Pediatric Clinician Perceptions Regarding Engagement with CDC's Milestone Tracker App

Building on the twelve articles that employed qualitative methods to measure engagement, the third study of this dissertation was designed to collect qualitative user experience data via individual in-depth interviews to assess awareness, attention, perceived use, and facilitators, and barriers to use. Using a similar process described above to identify salient constructs, measures that were most frequently reported, detailed, and feasible were incorporated into the interview guide. Some of the interview questions (e.g., "Would you recommend this app to parents and caregivers? Why or why not?"; Taki et al., 2017) measure satisfaction, an indication of affective processing, which was a prevalent construct among articles included in this review. Similarly, the coding scheme used to analyze transcripts reflected the behavioral, cognitive, and affective constructs that arose from pediatric clinician's perspectives and experiences with *CDC's Milestone Tracker*.

Study Limitations

Only peer-reviewed, English language published articles were included in this review resulting in potential publication bias. Despite efforts to apply an inclusive search strategy, articles may have been excluded from if they were published as white papers or in the gray literature, as searching the gray literature for this topic with the specified inclusion and exclusion criteria was not recommended by the librarian consulted for this study. Future reviews should include relevant unpublished work. It was not possible to determine the most effective measures and methods to evaluate engagement because of the variation in measurement and methods. Several articles did not report desired information and therefore data extraction was limited. In some instances, the researchers were left to infer study population using the definition developed for this study when it was not explicitly stated as being designed with or for groups experiencing disadvantages. There is likely recency bias as conceptualizations of engagement that predate the release of the iPhone and other smartphones in 2007 may be different from how we currently think about engagement and how users engage with technology. Finally, it was not possible to assess risk of bias because not enough information was provided for most studies, however article quality (i.e., the elements attributable to reporting and study design) was accounted for in this review (Higgins, Savović, et al., 2019).

Conclusions

This chapter discussed a systematic review of the literature to synthesize existing conceptual definitions of engagement and current evaluation practices (Aim 1). While few articles reported engagement definitions, all described how engagement was measured and the methods used to measure engagement. Engagement measurement may depend on the health issue, context of the intervention, as well as related outcomes. In addition to examining engagement behavior, it may be necessary to assess the cognitive and affective processes that comprise engagement. This would allow for a better understanding of app effectiveness and inform future interventions. Conclusions from this review are revisited in Chapter 5 of this dissertation.

Chapter 3

Exploring Parent and Caregiver Engagement with *CDC's Milestone Tracker* App

Introduction

Developmental delays (e.g., cognitive, speech and language, gross and fine motor, and social delays) and disabilities (e.g., ASD, ADHD, blindness, hearing loss, learning and intellectual disability) are common in young children (Zablotsky et al., 2019) but if identified early, delays can be improved through EI services (Adams & Tapia, 2013; Bruder, 2010). Early identification and timely referral to EI can result in improved outcomes and reduce long-term costs associated with developmental delays and disabilities (Adams & Tapia, 2013; American Academy of Pediatrics, 2001b; Bradshaw et al., 2015). However, racial, ethnic, and socioeconomic disparities persist in knowledge of child development, developmental surveillance and screening, timely referral to early identification, and participation in EI (Barger et al., 2022; Bilaver et al., 2020; Bradley & Corwyn, 2002; Daniels & Mandell, 2013; Donohue et al., 2017; Durkin et al., 2017; Shaw et al., 2021; Wittke & Spaulding, 2018; Zuckerman et al., 2014). Children from disadvantaged racial, ethnic, and socioeconomic backgrounds are more likely to be identified later and therefore receive fewer EI services than their counterparts (Daniels & Mandell, 2013; Feinberg et al., 2011; Magnusson & Mistry, 2017; Ratto et al., 2015).

Since children develop at different rates, mild delays are often difficult to detect, resulting in delayed identification. The AAP guidelines recommend developmental surveillance (also known as developmental monitoring) and developmental screening for early detection of developmental delays and disabilities, which when implemented together, improve early identification (Barger et al., 2018). While no single tool is appropriate for all children (American Academy of Pediatrics, 2006), many tools exist to monitor developmental milestones, and identify those at risk for developmental delays and disabilities (American Academy of Pediatrics, 2001a). One such tool is *CDC's Milestone Tracker* (www.cdc.gov/MilestoneTracker) from the CDC's LTSAE program.

CDC's Milestone Tracker is a behavior change app with seven BCTs to help families (particularly those from groups experiencing disadvantages) learn about child development, track developmental milestones, and act early if they have concerns about their child's development. *CDC's Milestone Tracker* is free to download and available in English and Spanish. Following its release in 2017, the app has undergone several updates to improve functionality and content. As of January 2022, the app includes 10 interactive, age-appropriate developmental milestone checklists (for ages 2, 4, 6, 9, 12, 18 months, 2, 3, 4, and 5 years) with photos and videos for each milestone that can be observed in natural settings as well as space to enter notes. The app generates a milestone summary that parents and caregivers can share with clinicians and other early childhood providers. Also built into the app are appointment reminders and tips and activities to support early development.

CDC recently announced that the *CDC's Milestone Tracker* app surpassed one million downloads (as of January 2022). Evaluations of the paper versions of the LTSAE milestone checklists demonstrate positive outcomes (e.g., increased knowledge and awareness of developmental milestones, confidence to discuss concerns, and positive attitudes toward developmental surveillance; Abercrombie et al., 2021; Daniel et al., 2009; Farmer et al., 2022;

Gadomski et al., 2018; Graybill et al., 2016), however the only published evaluation of the app was a series of usability studies during the design phase of app development of an initial version called “ActEarly” (Armenta et al., 2019). User testing was conducted with 23 English and Spanish-speaking adults—less than half of which were parents or caregivers of young children—and identified pain points (i.e., potential problems users may experience) to develop and test a redesigned prototype of the app (Armenta et al., 2019). Thus, more research is needed to evaluate current versions of the app to understand how parents and caregivers use and interact with *CDC’s Milestone Tracker* app content and features.

As discussed in Chapters 1 and 2, engagement, comprised of cognitive (e.g., attention, interest, and immersion in the behavior change app), affective (e.g., enjoyment and satisfaction), and behavioral (e.g., behavior change app usage) processes that characterize the user experience (Cole-Lewis et al., 2019; Flaherty et al., 2021; Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017; Yardley, Spring, et al., 2016), is an important component of intervention evaluation. This multidimensional concept functions as the interplay between behavioral, cognitive, and affective factors and is believed to be a determinant of health behaviors (Kelders, van Zyl, et al., 2020; Perski, Blandford, et al., 2019; Short et al., 2018). How users interact with the behavior change components of the app is dependent on how users interact with app features, and it is important to measure both aspects (Cole-Lewis et al., 2019). With the increase in resources invested in developing and disseminating behavior change apps, it is important to identify a clear conceptualization of engagement and corresponding measures. For interventions delivered using traditional channels (e.g., in-person health education classes for smoking cessation), a higher level of engagement is an indicator of effectiveness, however this does not necessarily translate to digital spaces where users must actively participate in the exchange of information, thoughts, and ideas. Thus, measures to understand the complexities of user engagement with behavior change apps are needed.

As demonstrated in Chapter 2, engagement with behavior change apps is rarely explored among groups experiencing disadvantages. Recent recommendations support assessing engagement with diverse populations (Brager et al., 2019; Cole-Lewis et al., 2019; Flaherty et al., 2021; Yardley, Spring, et al., 2016). Within the realm of early child development, McClure et al (2018) suggest more research is needed to assess how families from groups experiencing disadvantages—whose children experience the burden of disparities related to developmental delays and disabilities—access and use behavior change apps. Further, the authors highlight the need for a better understanding of how specific behavior change app features or elements increase engagement and promote behavior change to support early development. Therefore, the primary purpose of this study (Aim 2) was to explore engagement behaviors and identify factors that influence engagement with *CDC's Milestone Tracker* app among parents and caregivers with young children from groups experiencing disadvantages (defined in the following section). A secondary purpose was to evaluate the *CDC's Milestone Tracker* app. This study was guided by the following research questions:

RQ1: How do parents and caregivers of young children engage with the *CDC's Milestone Tracker* app?

RQ2: What features of the *CDC's Milestone Tracker* app are engaging for parents and caregivers of young children?

This research builds on the models of engagement proposed by Kelders et al (2020) and Cole-Lewis et al (2019) and much of the emerging engagement literature that positions engagement as multidimensional consisting of behavioral (interaction with the technology and health behavior), cognitive, and affective processes. This chapter reports on an evaluation of engagement with *CDC's Milestone Tracker* and serves as a complement to findings presented in Chapter 4 to provide insight into how engagement with behavior change apps for groups experiencing disadvantages can be measured.

Methods

Study Design

This study employed both a within-subjects and a between-subjects design. Data were collected through two self-administered online (accessible via desktop or mobile) surveys (baseline and one-month follow-up) conducted November 5, 2021 through January 27, 2022 using Qualtrics software (Version October 2021, Provo, UT). Survey research is used to obtain information from a sample of individuals (Fowler Jr, 2013) and is considered an accessible way to assess subjective measures of engagement (Short et al., 2018). There are many benefits to using survey research methods including that surveys enable systematic data collection from individual responses to questions that can be compared within or across samples and aggregated to make generalizations about the sample or population (Fowler Jr, 2013). Further, because surveys can be self-administered and respondents can complete them on their own time, they tend to be a low-cost/low-investment option for data collection, helping to minimize the burden on study participants.

This study was designed to evaluate engagement in real-world settings. As others (e.g., Roberts et al., 2019) have noted, research that involves participants' experiences using an app likely provides increased ecological validity and a better understanding of perceptions and preferences while reducing trail bias (Baumel et al., 2019). Further, as a somewhat natural experiment, there was no requirement for participants to download and use the *CDC's Milestone*

Tracker at any point during the study period. This mimics real-world application, where individuals are exposed to behavior change apps, but exposure may (or may not) lead to downloading the app and intervention uptake. Although it is challenging to study engagement in real-world settings (Perski, Blandford, et al., 2019), this study was designed to mitigate participation bias (where more highly engaged users participate in engagement research) by not recruiting a sample based on prior use of the app, however since potential respondents might know about the app from other sources, familiarity with the app was captured at baseline. Respondents were introduced to the app via a heat map task embedded within the baseline survey instrument and received pre-use engagement measures followed by post-use items on the one-month follow-up survey, reducing the potential for participation bias.

Additionally, aggregated *CDC's Milestone Tracker* app usage data from January 1, 2019 through November 30, 2021 were provided by the LTSAE program to examine general engagement patterns (e.g., number of app downloads, number of app launches, and number of milestone checklists are completed; see Measures section below for complete list). The timeframe for these data is inclusive of a major app update that occurred in March 2021 where new features and analytics were added to improve the app. Data were de-identified and not directly connected to any participants in this study.

Sample

A non-probability based purposive sampling strategy combined with snowball sampling was used to recruit potential participants. The intended sample size was 70 to 80 respondents to complete both the baseline and follow-up survey. This sample size was tailored to certain logistical limitations including time and funding available for this study. As this was an exploratory pilot study that included a repeated measures design, 70 to 80 respondents across two time points would provide sufficient data about engagement measurement (Hertzog, 2008). Similar sample sizes have been used in other studies to assess user engagement (Mendu et al., 2018; Pagoto et al., 2018).

To be eligible for this study, respondents had to be 18 years of age or older, a United States resident (including United States territories), a parent or caregiver of young children (birth to age five) who had not been identified with a developmental delay or disability (broadly defined as any speech and language, motor, hearing, or cognitive delay or diagnosed with ASD, ADHD, Cerebral Palsy, Down Syndrome, etc.), self-report their child had not received therapies or services for developmental delays and disabilities, comfortable speaking and reading English, and identify as someone from groups experiencing disadvantages. For the purpose of this study, groups experiencing disadvantages included people from minoritized racial and ethnic groups, persons with lower levels of income, and/or persons with disabilities (based on self-report disability status). This study population is one of the priority audiences for *CDC's Milestone Tracker*, and is consistent with calls for additional research with diverse populations to better understand engagement (Brager et al., 2019; Cole-Lewis et al., 2019; Yardley, Spring, et al., 2016).

Parents and caregivers of children with identified developmental delays and disabilities were ineligible for the study because they are not the intended audience for the *CDC's Milestone Tracker* app. Similar to a study that evaluated paper versions of the LTSAE developmental milestone checklists (Gallagher et al., 2019), respondents were excluded if any child in their family receive (or received in the past) any of the following therapies or services: (1) speech, occupational, physical, or other therapy provided by a county or state EI or special education program (such as Part C or EI services or therapy, (2) speech, occupational, physical, or other therapy provided by other agencies or organizations, or (3) other therapies or services for developmental delays or disabilities.

Procedures

Potential respondents were recruited via flyers and posts (Appendix B) to email listservs for families (e.g., members of the Northwestern University Early Intervention Research Group Parent Registry who met eligibility criteria, FamTech Founders Collaborative and their parent test market, and Illinois Medical Professionals Action Collaborative Team among others), posts shared via Instagram, and closed (i.e., private, members only) Facebook groups for parents and caregivers. After the baseline survey was launched and inadvertently posted to a Facebook group that was not closed, the survey received fraudulent, invalid responses (*n* = 49; duplicate responses, odd responses to open-ended items, implausible referrals, suspicious email addresses, did not pass Qualtrics quality check and flagged as bots). The survey was immediately closed (i.e., de-activated) to new responses per the recommendation of Qualtrics support and other researchers (Ballard et al., 2019; Pratt-Chapman et al., 2021; Storozuk et al., 2020). A second baseline survey (with a new URL and updated recruitment flyer) was launched and shared with my personal network of parents via email as well as posted to a listserv, however this survey was also attacked by additional bots and fraudulent responses and promptly closed. Among the 216 responses, only four (1.9%) were valid (see Appendix C for process used to detect fraudulent responses). For the third version of the baseline survey, a new strategy was used to recruit participants via email, text message, and direct messaging on Instagram to my personal network of parents of young children. Several Instagram followers then reposted the study flyer and direct URL to the baseline survey to share with their networks.

Potential participants who clicked the baseline survey were directed to the to the consent and if they agreed (by clicking the "I Agree" button), continued to the screening instrument which included 13 questions to screen for eligibility (see Appendix D for baseline survey instrument). The screening questions assessed age, smartphone ownership, language most comfortable speaking and reading, number of children and if any were under six years of age, if they had ever been told any of their children has a developmental delay or disability, whether any of their children were receiving or previously received therapies or services (see full list of excluded services in the Sample section above), state or territory of residence, self-reported race, self-reported ethnicity, level of difficulty paying monthly bills (Hanmer & Cherepanov, 2016), and self-reported disability status (Ipsen et al., 2021). Eligible participants were then automatically sent to complete the 20 to 25 minute baseline survey. Respondents who completed the baseline survey and provided a valid email address were sent an email, text message, or Instagram direct message with a personalized URL for the 15 to 20 minute follow-up survey approximately one month after they completed the baseline survey (see Appendix E for follow-up survey instrument). Non-responders received reminders via different channels and were contacted up to four times. Respondents received an incentive up to $50 for their time and participation ($25 gift card for completing the screener and baseline survey and an additional $25 gift card for completing the follow-up survey). This study was deemed exempt by the Institutional Review Board at Northwestern University (IRB STU00215896).

Due to the high number of bots and fraudulent responses on the first and second versions of the baseline survey, it was not possible to calculate the total number of valid responses. The third version of the baseline survey garnered 162 responses, 149 (92.0%) of which completed a screener. Among those, 85 (57.0%) were eligible and 71 (47.7%) completed the entire baseline survey. No other records were excluded. The four valid survey responses from the second version of the baseline survey were added to the responses from the third version of the baseline survey, resulting in 75 baseline surveys completed. All 75 respondents were sent the follow-up survey URL and among them, 72 (96.0%) completed the follow-up survey. The three respondents who did not complete the follow-up survey were excluded resulting in a final analytic sample of 72 parents and caregivers of young children.

Measures

The survey instruments were developed in collaboration with the LTSAE team, who requested certain items to pretest for inclusion in an in-app survey for a future version of the *CDC's Milestone Tracker*. Both surveys were pilot tested (baseline survey: $n = 4$; follow-up survey: $n = 4$) and revised, as necessary. Measures were informed by findings from the systematic review discussed in Chapter 2 as well as other studies evaluating engagement, app use, and early childhood development. The analyses reported in this chapter utilized several types of measures with task-based activities, closed-ended items, and open-ended items. The following sections organized by survey instrument describe each measure in detail.

Baseline Survey Measures

The baseline survey assessed child development, *CDC's Milestone Tracker* app perceived use, acceptability, behavioral intentions, media use, and demographics. Measures pertinent to the current analysis are described below.

Child Development. Respondents were asked to indicate their level of familiarity with (or awareness of) five early childhood development concepts, tools, and resources: (1) developmental milestones, (2) EI services (i.e., Part C of IDEA), (3) special education services for children three years and older (i.e., Part B 619 of IDEA), (4) CDC milestone checklists and other LTSAE resources, and (5) *CDC's Milestone Tracker* app using a 5-point scale ranging from 1 = *Not familiar at all* to 5 = *Extremely familiar*. A similar open-ended item about awareness of LTSAE was asked in the interviews with pediatric clinicians discussed in Chapter 4 (adapted from Gerndt & Mitchell, Forthcoming 2022). Respondents were also asked how concerned they are about their child's development on a scale ranging from 1 = *Not at all concerned* to 5 = *Extremely concerned*.

Heat Map Analysis. In lieu of the proposed eye-tracking assessment with a retrospective think-aloud, a heat map analysis was performed within the Qualtrics platform to serve as a proxy for eye-tracking. Heat map items in Qualtrics have been successfully used for message testing to understand which features capture respondent's attention (Hoek et al., 2016) and heat maps have been used in studies to demonstrate behavior change app engagement (Bell et al., 2020). First, respondents were presented with a new survey page with three different images of *CDC's Milestone Tracker* app and received the instructions "Click as fast as you can on the image of the app that catches your attention. Please click only one spot." The three images were presented together on the same page and selected because they collectively depict all features of the app. The images of the app were divided into 13 regions (similar to areas of interest used in eye-tracking) based on different app features (Appendix F). These regions were hidden to respondents and clicks or touches outside of those designated areas were captured as "other." Following the heat map task was an open-ended item asking respondents to explain why they clicked on that spot, an approach often used when conducting retrospective think-aloud studies (Short et al., 2018).

The heat map task and open-ended item was repeated using the same three images of the app for five additional sets of questions (with different instructions) assessing cognitive and affective processing: (1) most interesting part of the app ("Click on the most interesting image of the app."), (2) app feature respondents would go to first if using the app for the first time ("Click the image of the app that you would you go to first if you were using the app for the first time."), (3) likes ("Click the images of the app that you like."), (4) dislikes ("Click the images of the app that you do not like."), and (5) images that are unclear ("Click the parts of the app that are unclear."). Only one click or touch was recorded for most interesting part of the app and features respondents would go when first using the app whereas respondents could click or touch as many areas as they wanted for likes, dislikes, and confusing images (although only 10 clicks or touches were recorded). Respondents' click or touch behavior for each heat map task were aggregated to create visual depictions of popular and unpopular areas of the app (i.e., heat maps). While it was not possible for the Qualtrics survey to capture proper eye-tracking measures such as fixation-based gaze metrics and point-of-gaze metrics most often used to operationalize attention and depth of cognitive processing (King et al., 2019), timing items were added to each of the six heat map task pages to capture time to first click (as a proxy for reaction time) and total time spent on the page (as a proxy for total time looking at the whole screen; Cho et al., 2019; Holmqvist et al., 2011).

***CDC's Milestone Tracker App*: Engagement and Related Constructs.** Following the heat map task, respondents were asked a series of questions about engagement and related constructs including acceptability and satisfaction with *CDC's Milestone Tracker* app. Respondents were asked to indicate their level of agreement with 10 statements assessing engagement dimensions and related constructs including behavioral antecedents (e.g., "I would use this app to help me talk about my child's development with a doctor or healthcare provider"), trustworthiness of the app (e.g., "I trust this app to help me identify concerns about my child's development”), and relatability of the app (e.g., “This app was designed for someone like me”) which measures identity. Assessing behavioral antecedents (e.g., self-efficacy and behavioral intentions) is beneficial because they demonstrate engagement in the behavior change process (i.e., engagement with the health behavior; Short et al., 2018). These items were selected because they represent multidimensional constructs from the existing engagement scales such as TWEETS (see Appendix G for comparison of items; Kelders & Kip, 2019; Kelders, Kip, et al., 2020) and uMARS (Stoyanov et al., 2016) and were of interest to the LTSAE team. Response options ranged from 1 = *Strongly disagree* to 5 = *Strongly agree*.

Perceptions of acceptability (referred to as “pre-use acceptability” or “prospective acceptability,” see Perski and Short 2021 for review) was assessed using the single star rating item (“How would you rate this app? 1-5 stars with 5 being the highest”) paired with an open-ended item asking respondents to explain their response as suggested by Perski and Short (2021). This measure may be useful for capturing respondents’ reactions to the app and identifying influencing factors (Perski & Short, 2021). Satisfaction was assessed by asking “Would you recommend this app to friends and family?” (with response options *No*, *Yes*, and *Maybe* recoded into a binary variable for analysis where 0 = *No* and 1 = *Yes and Maybe*) followed by an open-ended item asking respondents to explain why or why not. This item—also asked of pediatric clinicians during the interviews described in Chapter 4—was adapted from the feedback index of the engagement index (Taki et al., 2017; an article from the systematic review in Chapter 1).

Media Use. Survey items to assess media use included app use, frequency of use, and if respondents have ever used an app for their health or their child’s health (adapted from Perski, Blandford, Ubhi, et al., 2017).

Sociodemographics. Sociodemographic items included self-reported gender, children's age (in months or years), relationship status, educational attainment, household income, employment status, health insurance status and type of coverage. Contact information (first and last name, email address, and phone number) was also collected to provide the incentive, check for fraudulent responses, and recontact respondents to complete the follow-up survey.

Follow-up Survey Measures

Many of the items on the baseline survey were repeated on the follow-up survey to measure changes from baseline to one-month follow-up. The follow-up survey assessed child development, *CDC's Milestone Tracker* app engagement and related constructs and behavioral outcomes. Measures specific to this analysis are described below.

Child Development. Respondents received the same items about child development that were asked on the baseline survey to measure knowledge acquisition.

***CDC's Milestone Tracker App*: Engagement and Related Constructs.** The 10 items assessing behavioral antecedents, trustworthiness, and relatability/identity were also asked again, however six new items were added to the follow-up survey and assessed using the same 5-point scale. These items evaluated satisfaction (e.g., "I like using this app to track my child's development"), perceived ease of use (e.g., "The app is easy to use"), and barriers to use (e.g., "I couldn't find all of the answers I needed in the app" and "I was concerned about the Internet data usage on my phone when using the app"). Perceived ease of use emerged as an important engagement measure from the articles included in the systematic review in Chapter 2 (Hakobyan et al., 2016; Hartzler et al., 2016; Hayes et al., 2014; Saksono et al., 2020; Taki et al., 2017) and is included in several DBCI evaluation models such as the TWEETS, uMARS, and technology acceptance models (Davis, 1989; Holden & Karsh, 2010). Likewise, barriers to use were measured by several articles included in the systematic review (e.g., Quante et al., 2019; Taki et al., 2017) and have been used to capture user burden (Suh et al., 2016).

Two items believed to be “best bets” for short measures of engagement were added to the follow-up survey (Perski, Blandford, et al., 2019; Perski, Lumsden, et al., 2019). Respondents were asked to indicate their level of agreement with the statement “I like using this app to track my child’s development.” using a 5-point scale ranging from 1 = *Strongly disagree* to 5 = *Strongly agree* (see Table G1 for comparison of items). They were also asked “How engaging is the *CDC’s Milestone Tracker* app?” using a 5-point scale ranging from 1 = *Not at all engaging* to 5 = *Extremely engaging* followed by two open-ended items to explain their answer and what would make the app more engaging. The same questions about acceptability (app rating scale) and satisfaction (if respondents would recommend the app) were asked on the follow-up survey.

***CDC’s Milestone Tracker App*: Subjective Use and Behavioral Outcomes.** The survey asked respondents if they downloaded *CDC’s Milestone Tracker* app and those that responded affirmatively received a series of questions to measure subjective usage of the app. Given that *CDC’s Milestone Tracker* app data is not collected at the individual user level, it was not possible to gather objective usage data (a similar challenge encountered by Kelders, Kip, et al., 2020). Therefore, subjective usage was assessed using items from the *DBCI Engagement Scale* about how they used the app, including how much time they spent in the app (Perski, Blandford, et al., 2019; Perski, Blandford, Ubhi, et al., 2017), which features of the app they used (Perski, Blandford, et al., 2019), and what prompted them to use the app. Respondents who downloaded the app were also asked two behavioral outcome items: (1) who they shared their child’s milestone summary with and (2) if the app helped identify any possible concerns about their child’s development.

CDC's Milestone Tracker App Usage Data

A secondary analysis of *CDC's Milestone Tracker* app data was conducted to explore engagement behaviors. De-identified data with aggregated totals by month were not directly linked to any of the participants in this study. Data in this analysis included: app downloads, app launches (i.e., each time a user opens the app from their device), average minutes spent per app launch, visitors to the app (unique visitors, returning unique visitors, and average daily users), visitors to specific features (e.g., tips and activities, milestone overview, show doctor, email doctor, appointments, notifications), milestone checklists started and completed by age, milestone checklist videos played, and child profiles added (i.e., ages and gender of children added to the app).

Data Analysis

This study was designed with the intention of collecting data to identify factors that predict engagement with the app features (e.g., frequency of use, type of use), engagement with the behavior change components of the app (e.g., behaviors such as completion of milestone checklists, identifying any potential concerns, acting early by talking to a clinician or early childhood care provider about concerns), cognitive engagement, affective engagement, and facilitators and barriers to use. The plan was to assess changes within-subjects from baseline to follow-up for behavioral antecedents (e.g., awareness, self-efficacy, etc.) and skip logic was built into the survey so that respondents who downloaded and used the app received the questions necessary to perform this exploratory analysis. Unfortunately, only about half of respondents in this sample downloaded the app, resulting in limited data about engagement with app features and most of the behavioral items to assess engagement related to behavior change.

Consequently, an alternate analytic plan was developed to create a new behavior change app engagement scale to measure subjective engagement with apps for groups experiencing disadvantages. Given that engagement is multidimensional, comprised of behavior, cognition, and affect (Kelders, van Zyl, et al., 2020), a single item to measure engagement may not be reliable and a composite measure is necessary to capture the different dimensions and the interaction between them. As such, the new goal for analysis was to use the items that theoretically encompass the multidimensional properties of engagement—behavior, cognition, and affect, which for the purposes of this study were categorized according to similar items on the TWEETS (Kelders & Kip, 2019; Kelders, Kip, et al., 2020)—to create a new behavior change app engagement scale (see Table G2 for comparison of survey items to items on the TWEETS). This new scale was used to explore engagement with the *CDC's Milestone Tracker* app among parents and caregivers from groups that experiencing disadvantages.

Data from the baseline survey and one-month follow-up survey were cleaned and then merged into a final dataset (Van den Broeck et al., 2005). Aggregated, de-identified data were analyzed using SPSS 28 statistical software (IBM Corp., Armonk, New York). Since the only engagement scale identified in the systematic review discussed in Chapter 2 required the use of both objective and subjective user data (Taki et al., 2017), the new behavior change app engagement scale drew on constructs from some of the articles about behavior change apps for groups experiencing disadvantages included in the review but was primarily derived from the TWEETS. Given that the TWEETS is a scale comprised of all three engagement dimensions and has demonstrated good validity and reliability, it seemed reasonable to employ a similar scale. Items that were asked on both the baseline and follow-up surveys were mapped to the TWEETS to determine which items should be included on the initial version of a behavior change app engagement scale. The new scale was comprised of behavioral antecedents (behavior engagement), trustworthiness of the app (cognitive engagement), and relatability/identity (affective engagement). Internal consistency was calculated using Cronbach alpha, where a coefficient greater than .70 indicated good item covariance and high internal consistency (Tavakol & Dennick, 2011). Test-retest reliability (i.e., degree to which the scales are related) was calculated using Pearson correlations between the behavior change app engagement scale at baseline and follow-up, where the effect size greater than .5 indicated strong reliability (Ferguson, 2016).

Descriptive statistics were used to assess engagement scores, app ratings, and recommendations among the total sample. Other baseline sample characteristics (e.g., sociodemographics and level of concern about child's development) were examined descriptively. McNemar's tests were used to compare changes in awareness of early childhood development, tools and resources from baseline to follow-up among respondents that downloaded the app. Statistical significance was determined as $p < .05$. As a preliminary step, open-ended survey response options were examined to provide examples to support quantitative findings, however more in-depth qualitative analyses are warranted. Finally, aggregated usage data from *CDC's Milestone Tracker* were analyzed to report descriptive statistics for the app and provide context for how current users engage with the app.

Results

Sample Characteristics

Among the 72 respondents, most were women (n = 55, 76.4%), non-Latinx Black (n = 40, 55.6%) with reported annual income greater than $90,000 ($n$ = 54, 75.0%), although little more than one third of the sample reported difficulty meeting monthly bills (n = 26, 36.1%). On average, respondents were 34.6 years (*SD* = 5.1, range 28-66) with 1.6 children (*SD* = 0.8, range 1-5) who were mean age of 43.2 months (*SD* = 60.1, range 0-300). Table 6 provides full sociodemographic information about the sample.

More than half of respondents (*n* = 39, 54.2%) had at least some concerns about their child's development at baseline, which mostly remained consistent on the one-month follow-up (*n* = 35, 48.6%). Concerns about child development did not differ by those who downloaded *CDC's Milestone Tracker* compared to those who did not (*p* = .984). While all respondents reported that they use apps they currently had on their phones daily, fewer reported using apps for their health (*n* = 61, 84.7%) or their child's health (*n* = 36, 50.0%).

Behavior Change App Engagement Scale Internal Consistency and Reliability

Descriptive statistics for the behavior change app engagement scale items are reported in Table 7. The Cronbach alpha of the behavior change app engagement scale at baseline and follow-up was .92 and .90, respectively, indicating high internal consistency. Pearson correlation of the behavior change app engagement scale at baseline and follow-up was significant (*r* = .69, *p* <.001), showing strong test-retest reliability.

Engagement with *CDC's Milestone Tracker* App

An exploratory analysis was used to answer the first research question: How do parents and caregivers of young children engage with the *CDC's Milestone Tracker* app? Mean scores on the behavior change app engagement scale were higher at baseline (*M* = 27.3, *SD* = 5.4) than at one-month follow-up (*M* = 26.0, *SD* = 5.1). Constructs related to engagement were similar at baseline and follow-up, with the median app rating of 4.0 for both baseline (IQR 3.0-4.0) and follow-up (IQR 3.0-5.0). Nearly the entire sample said they would recommend the app to friends and family at both baseline (*n* = 69, 95.8%) and follow-up (*n* = 70, 97.2%; Table 8).

Table 6

Sample Characteristics (N = 72)

Sociodemographics	*N*	%
Age (*M*, *SD*) [a]	34.6	5.1
Gender		
Woman	55	76.4
Man	15	20.8
Non-binary	2	2.8
Race and Ethnicity		
Non-Latinx Asian	4	5.6
Non-Latinx Black or African American	40	55.6
Latinx	12	16.7
Non-Latinx Multiracial	5	6.9
Non-Latinx other	3	4.2
Non-Latinx White	8	11.1
Relationship Status		
Married/domestic partner/civil union	62	86.1
Living with partner	6	8.3
Single	4	5.6
Education		
High school graduate or GED	3	4.2
Some college or certificate	3	4.2
Associate's degree	1	1.4
Bachelor's degree (BA, BS)	22	30.6
Post graduate degree (MA, JD, MD, PhD)	43	59.7
Income		
< $89,999	11	15.3
$90,000-$199,999	33	45.8
> $200,000	21	29.2
Prefer not to answer	7	9.7
Difficult to meet monthly bills [b]	26	36.1
People with disabilities [b]	7	9.7
Health insurance type		
Private through workplace	62	86.1
Private through the ACA	4	5.6
Medicaid	2	2.8
Medicare	1	1.4
Prefer not to answer	3	4.2
Number of children (*M, SD*) [c]	1.6	0.8
Child age in months (*M, SD*) [d]	43.2	60.1

Note. MT = *CDC's Milestone Tracker* app; ACA = Affordable Care Act Health Insurance Marketplace.

[a] Sample age range 28-66.[b] Reflects the number and percentage of respondents answering "yes" to this question. [c] Sample number of children range 1-5. [d] Sample child age in months range 0-300.

App Engagement Among Respondents who Downloaded *CDC's Milestone Tracker*

An evaluation of app engagement with *CDC's Milestone Tracker* among those who downloaded the app was also used to answer the first research question. On the baseline survey, 62.5% (*n* = 45) of respondents indicated they plan to use *CDC's Milestone Tracker*. Among the 35 respondents (48.6% of total sample) who downloaded the *CDC's Milestone Tracker*, 74.3% (*n* = 26) completed a milestone checklist (Table 9). While one third of respondents (*n* = 12) shared their child's milestone summary with family or friends, only three (8.6%) shared the summary with their child's doctor or health care provider. Another one third did not share their child's summary with anyone, citing that "there was nothing to share" because their child met milestones and they did not have any concerns. Average ratings for the two statements measuring barriers to use—"I couldn't find all of the answer I needed in the app" (*M* = 2.8, *SD* = 1.2) and "I was concern about Internet data usage when using the app" (*M* = 1.7, *SD* = 1.1)—were both low. Less agreement with the statements suggests respondents in this study encountered few barriers that might result in less engagement. Additional descriptive statistics are reported in Table 9.

Respondents also received the two short measures of engagement. The mean rating of the single item engagement measure "How engaging is the app?" was 3.8 (*SD* = 0.9 out of 5) while the mean rating for the item asking respondents how much they like using the app was lower (*M* = 3.6, *SD* = 0.9 out of 5), indicating a medium-to-high level of engagement with the app. Open-ended responses for why respondents found the app engaging primarily focused on the videos and ease of use (e.g., "straight-forward and informative," "easy to navigate," and "Checklist is helpful and the videos are engaging"). Others pointed out that the app does not necessarily need to be more engaging since it is a health tracking tool and not a social media platform explaining, "While I enjoy using the app, I don't find a reason to utilize it for too long" and "It feels like if I'm not at a milestone, there's no push to use it." Respondents noted push notifications with recommended tips and activities between milestone checklists could make the app more engaging.

There were statistically significant differences in changes of awareness from baseline to one-month follow-up among parents and caregivers who downloaded the app for awareness of LTSAE, $\chi^2(1) = 7.692$, exact $p = .003$, and awareness of *CDC's Milestone Tracker* app, $\chi^2(1) = 14.087$, exact $p = < .001$ (Figure 4). There were no changes in awareness of developmental milestones ($p = .500$), EI ($p = .683$), or Part B 619 (special education; $p = .219$).

Heat Map Analysis: Engaging App Features

A heat map analysis was used to answer the second research question: What features of the *CDC's Milestone Tracker* app are engaging for parents and caregivers of young children? The heat map task captured the features of *CDC's Milestone Tracker* app respondents found engaging along two engagement dimensions: cognition and affect. The frequency and percentage of respondents who clicked or touched an area with exemplary responses from the open-ended items that accompanied the heat map tasks are summarized below. Time to first click and total time spent on page, which serve as proxies for fixation-based gaze metrics typically captured in eye-tracking assessments are included in Table 10.

Table 7

Descriptive Statistics for Behavior Change App Engagement Scale Items at Baseline (N = 72)

	Range	Mean (*SD*)	Variance	Skewness	Kurtosis
Behavioral Engagement					
1. "I plan to use this app to track my child's development in the future."	1-5	3.8 (1.0)	1.02	-0.59	0.01
2. "I would use this app to learn more about tips and activities that can help my child learn and grow."	2-5	4.2 (0.8)	0.65	-0.82	0.37
3. "I would use this app to share concerns about my child's development with my doctor or healthcare provider."	2-5	3.9 (1.0)	0.92	-0.57	-0.61
Cognitive Engagement					
4. "I trust this app to help me track my child's development."	1-5	3.9 (0.9)	0.86	-0.77	0.90
5. "I trust this app to help me identify concerns about my child's development."	1-5	3.9 (0.8)	0.69	-0.69	1.04
6. "I trust this app to provide accurate information about what milestones my child should be reaching for their age."	1-5	4.0 (0.9)	0.84	-1.06	1.63
Affective Engagement					
7. "This app was designed for someone like me."	1-5	3.7 (1.1)	1.13	-0.90	0.60

Note. Respondents asked to indicate their level of agreement on a scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Table 8

Behavior Change App Engagement Scale and Related Constructs (N = 72)

Behavior Change App Engagement Scale [a]	*M* (*SD*)
Baseline [b]	27.3 (5.4)
Follow-up [c]	26.0 (5.1)
“How would you rate this app?” [d]	*Mdn* (IQR)
Baseline	4.0 (3.0-4.0)
Follow-up	4.0 (3.0-5.0)
“Would you recommend this app to friends and family?” [e]	*n* (%)
Baseline (Yes)	69 (95.8%)
Follow-up (Yes)	70 (97.2%)

Note. IQR = Interquartile range; MT = *CDC’s Milestone Tracker* app. Follow-up survey administered at one-month post baseline survey.

[a] Behavior change app engagement scale includes 7-items. [b] Behavior change app engagement scale at baseline range 10-35. [c] Behavior change app engagement scale at follow-up range 12-35. [d] App rating is on a scale from 1 to 5 stars, with 5 being the highest. [e] Reflects the number and percentage of respondents answering “yes” or “maybe” to this question.

Table 9

Descriptive Statistics for App Engagement Among Respondents who Downloaded MT (n = 35)

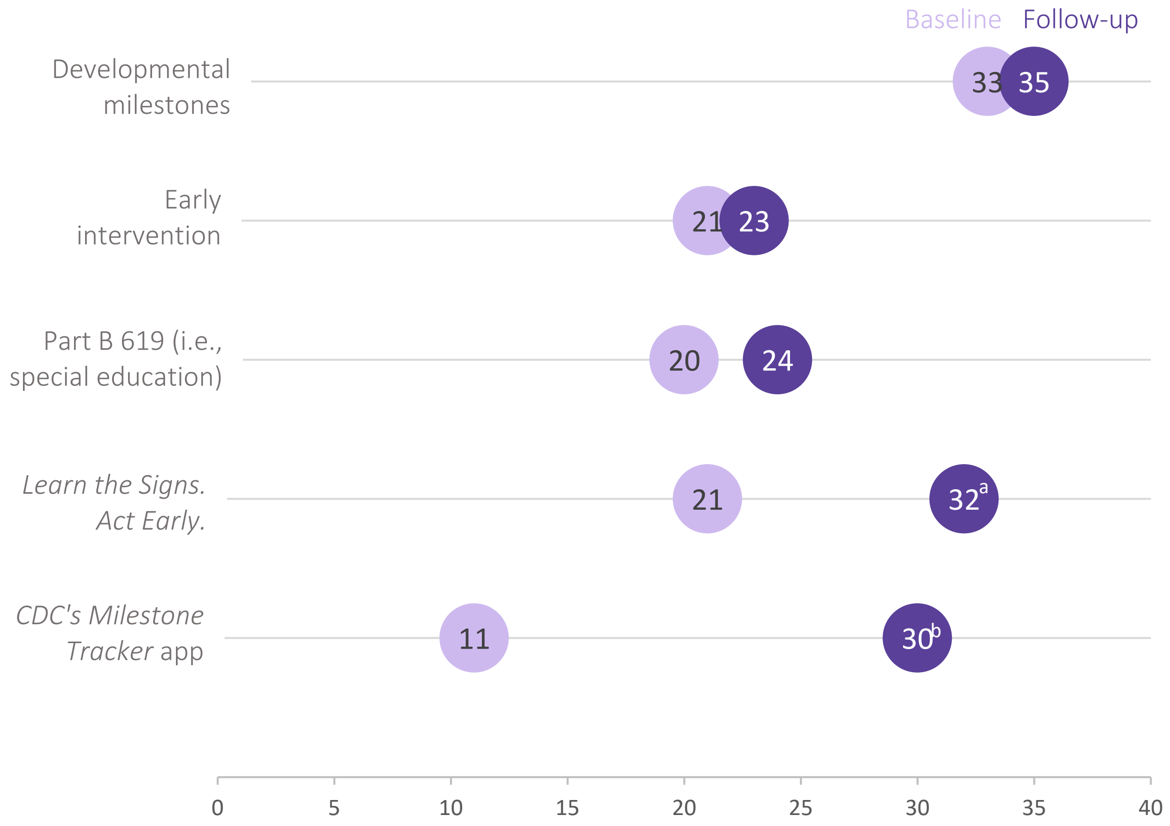
	n	%	
Use MT monthly	11	31.4	
Completed milestone checklist (Yes)	26	74.3	
"What typically prompts you to use the MT app?" [a]			
When I have a possible concern	16	45.7	
Preparing for doctor's appointment (e.g., well-child visit)	13	37.1	
Other [b]	8	22.9	
Notifications from the app	5	14.3	
"Who did you share your child's milestone summary with?" [c]			
Family or friends	12	34.3	
Doctor or healthcare provider	3	8.6	
Did not share summary	12	34.3	
"Did the MT app help you identify any possible concerns related to your child's development?" (Yes)	5	14.3	
	M	*SD*	Range
How much time spent (in minutes) [d]	21.8	25.7	0-120
"I couldn't find all of the answers I needed in the app." [e]	2.8	1.2	1-5
"I was concerned about the Internet data usage on my phone when using the app." [e]	1.7	1.1	1-4
"The app is easy to use." [e]	4.2	0.7	3-5
"I like using this app to track my child's development." [e]	3.6	1.2	1-5
"How engaging is the MT app?" [f]	3.8	0.9	1-5

Note. MT = *CDC's Milestone Tracker* app.

[a] Respondents could select all prompts that apply therefore percentages do not add up to 100%. [b] Other prompts to use MT included "Going through phone at night," "When I think of it," and "When I'm trying to plan our days." [c] Respondents could select all that apply for who they shared their child's milestone summary with and therefore percentages do not add up to 100%. [d] How much time spent over past two weeks in minutes. [e] Respondents asked to indicate their level of agreement on a scale ranging from 1 (strongly disagree) to 5 (strongly agree). [f] How engaging is the MT app is on a scale ranging from 1 (not engaging at all) to 5 (extremely engaging).

Figure 4

Changes in Awareness of Early Childhood Development, Tools, and Resources from Baseline to One-Month Follow-up Among Respondents who Downloaded MT (n = 35)



Note. MT = *CDC's Milestone Tracker* app. Counts represent the frequency of respondents who reported at least some familiarity with each item. Respondents were asked "Please indicate how familiar you are with each of the following" on a scale ranging from 1 (not at all familiar) to 5 (extremely familiar). Response options were collapsed into not familiar (1) and familiar (2-5).

[a] McNemar's test deteremined statistically significant differences between baseline and one-month follow-up, $\chi^2(1) = 7.692$, exact $p = .003$. [b] McNemar's test deteremined statistically significant differences between baseline and one-month follow-up, $\chi^2(1) = 14.087$, exact $p = < .001$.

Table 10

Descriptive Statistics for Heat Map Time to First Click and Total Time Spent on Page (N = 72)

Construct	Directions	Time to First Click *M (SD)*	Total Time Spent on Page *M (SD)*
Attention [a]	"Click as fast as you can on the image of the app that catches your attention."	8.9 (5.7)	43.2 (29.3)
Most Interesting [a]	"Click on the most interesting image of the app."	11.0 (16.7)	48.3 (70.2)
First Use [a]	"Click on the image of the app that you would go to first if you were using the app for the first time."	9.5 (13.8)	38.8 (37.7)
Confusion [a]	"Click on the images of the app that are unclear."	10.7 (13.5)	44.8 (73.1)
Satisfaction (Like) [b]	"Click on the images of the app that you like."	11.6 (21.8)	50.2 (36.4)
Dissatisfaction (Dislike) [b]	"Click on the images of the app that you do not like."	8.4 (12.3)	40.0 (33.9)

Note. Time in seconds.
[a] Assesses cognitive engagement. [b] Assesses affective engagement.

Cognitive Engagement

Four heat map tasks evaluated cognitive engagement: attention, most interesting, first use, and confusion. Figure 5 depicts respondent's selections of features of the *CDC's Milestone Tracker* that caught their attention. The bright areas in the heat map indicate the app features that received clicks or touches, where the brighter the spot, the greater the number of clicks or touches. As the figure shows, the video in the milestone checklist was the region most clicked on (*n* = 30, 41.7%), followed by the milestone checklist feature (*n* = 16, 22.2%) and child name and age (*n* = 12, 16.7%). A similar trend was observed for the part of the app respondents found most interesting, where more than half of the sample (*n* = 39, 54.2%) selected the video in the milestone checklists and another 15.3% (*n* = 11) selected the milestone checklist (see Figures H1 through H5 for remaining heat maps). They cited visual appeal of the app (e.g., "simple, clean look" and "bright colors") and images and pictures of a "real-life cute baby" as being inviting and relatable for reasons why they found these features most interesting.

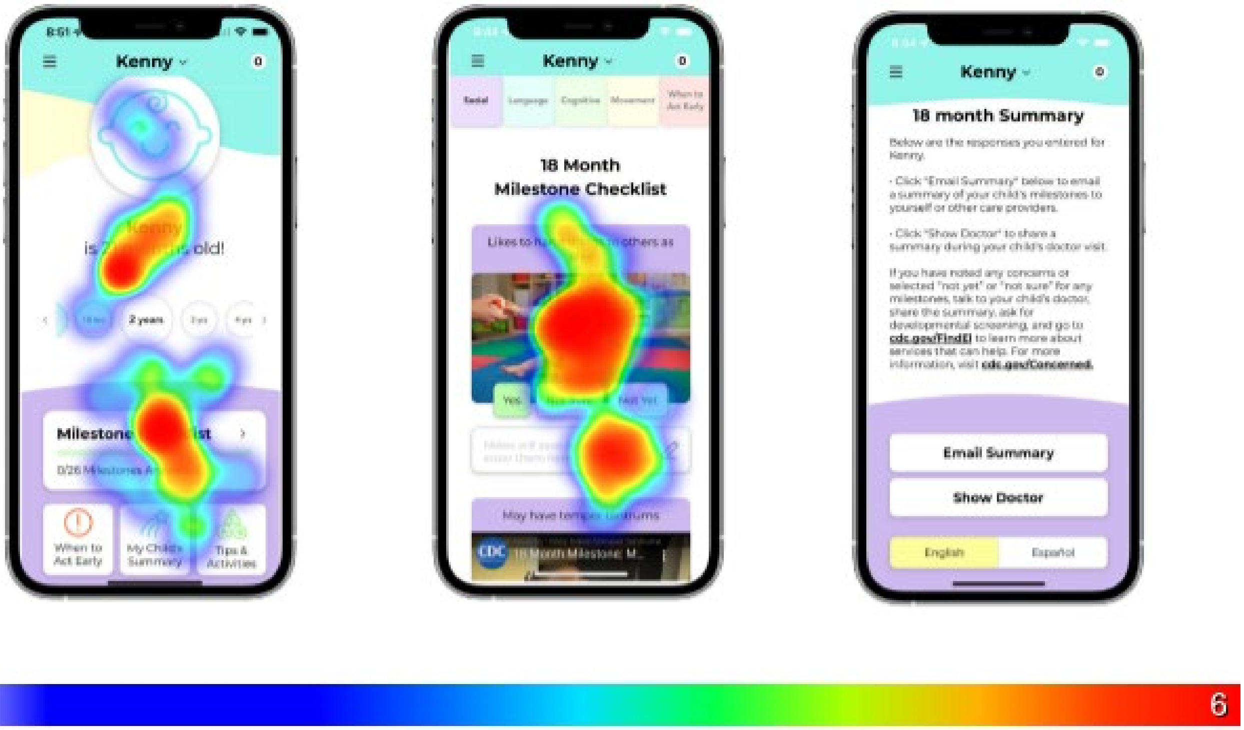
A third of parents and caregivers in this sample selected the milestone checklist as the first feature they would use if using the app for the first time (*n* = 25, 34.7%). Respondents explained this was because it looked like the starting place for the app and they would want to see what information the app provides, noting that milestones are the main purpose of the app. The final cognition heat map task sought to identify areas of the app that were confusing. Respondents indicated that the summary (*n* = 18, 25.0%), show doctor (*n* = 6, 8.3%), and child name and age (*n* = 6, 8.3%) features were most unclear; however it should be noted that some respondents skipped this item or only provided one click or touch even though they could select as many spots as they wanted (up to 10). Respondents found these features to be unclear because they were not sure what the summary was referring to, what 'act early' meant, and why they would need to show the summary to a doctor.

Affective Engagement

Two heat map tasks—where respondents could click or touch up to 10 areas—evaluated affective engagement: likes and dislikes. Unsurprisingly the respondents preferred videos and images over features that relied mostly on text (e.g., child's summary). Most liked features included the video in the milestone checklists (*n* = 41, 56.9%), milestone checklist (*n* = 38, 52.8%), and child name and age (*n* = 30, 41.7%). Respondents liked that the app looked easy to use and navigate and that the videos and images made understanding milestones simple, without requiring the user to read a lot of text. Accordingly, the summary feature (*n* = 39, 54.2%) was the most frequently disliked app feature because it was text heavy (e.g., "comes off as intimidating" and "too busy and too wordy").

Figure 5

Heat Map of Clicks or Touches on MT App Assessing Attention



18 month Summary

Below are the responses you entered for Kenny.

• Click "Email Summary" below to email a summary of your child's milestones to yourself or other care providers.

• Click "Show Doctor" to share a summary during your child's doctor visit.

If you have noted any concerns or selected "not yet" or "not sure" for any milestones, talk to your child's doctor, share the summary, ask for developmental screening, and go to cdc.gov/FindEI to learn more about services that can help. For more information, visit cdc.gov/Concerned.

Note. Respondents were asked "Click as fast as you can on the image of the app that catches your attention. Please click only one spot." The bright spots of the heat map represent the relative intensity of total clicks or touches on each part of the MT app image. Red indicates the part of the app that received the highest number of clicks or touches.

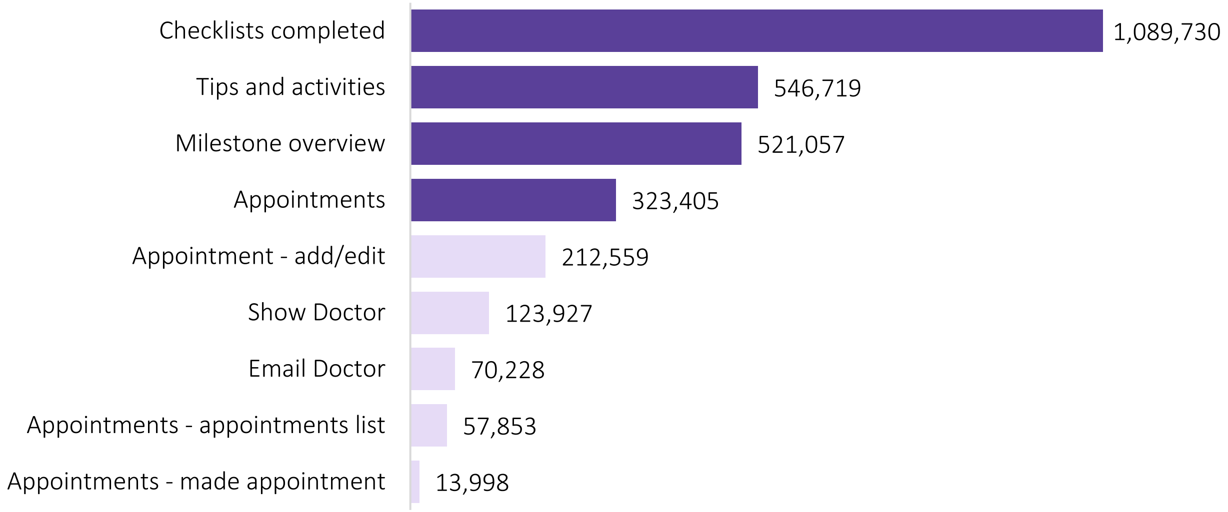
CDC's Milestone Tracker App Aggregated Usage Data

Usage data from the *CDC's Milestone Tracker* app was used to answer both research questions (how parents and caregivers engage with the app and which features they find engaging). Between January 2019 and November 2021, the *CDC's Milestone Tracker* app was downloaded nearly 800,000 times (975,063 total downloads since launch in 2017). On average, there were 86,675.7 unique users per month (56,268.5 returning users) and 2,904.2 active users per day. The app was launched more than 3.3 million times and users spent an average of 3.4 minutes on the app per visit. More than one million children profiles were created (*n* =

1,087,093, *M* = 31,059.8 per month, *SD* = 18,255.9), 54.4% of which were for male children. Expectedly, the app features used most were the milestone checklists. Among the 1,281,338 milestone checklists started, 1,089,730 were completed (85.0% completion rate), with an average of 31,135.1 (*SD* = 16,011.8) checklists completed per month. The 6 month, 4 month, and 2 month milestone checklists have been completed the most, accounting for 45.3% of all completed checklists since 2019. The videos embedded within the milestone checklists were played more than 5.3 million times (*M* = 154,256.3 plays per month, *SD* = 219,640.2). After milestone checklists, tips and activities (546,719 unique users), milestone overview (521,057 unique users), and appointments (323,405 unique users) were the most frequently visited features of the app (Figure 6). Users opened a total of 858,824 push notifications (*M* = 24537.8 notifications per month, *SD* = 19271.0), however data on the total number of notifications sent were unavailable prior to June 2021 and therefore data were not included in this analysis.

Figure 6

Most Frequently Visited Features of MT (January 2019-November 2021)



Note. MT = *CDC's Milestone Tracker* app. Data from 5/24/21 through 6/23/21 is incomplete and missing some data due to an error in the app configuration.

Discussion

This study explored engagement with the *CDC's Milestone Tracker* app among parents and caregivers with young children from groups experiencing disadvantages. Very little is known about engagement with behavior change apps among diverse communities, and this study provides insight to how parents and caregivers who primarily self-identify as non-White and/or people with disabilities use behavior change apps and which features they find engaging. Findings underscore the importance of using a measurement scale that positions engagement as both multidimensional and temporal to understand parent and caregiver engagement with the app. Understanding how users engage with behavior change apps and identifying factors that facilitate engagement is particularly important within the context of the COVID-19 pandemic. At a time when there have been significant decreases in developmental screening, referrals, and service provision for young children due to the COVID-19 pandemic and COVID-19 mitigation strategies (Wong et al., 2020; Yoshikawa et al., 2020), tools like *CDC's Milestone Tracker* app that allow for asynchronous intervention and education have considerable potential to change behaviors and improve health outcomes, particularly among groups experiencing disadvantages.

Two online surveys (baseline and follow-up) administered one month apart and a secondary analysis of system usage data produced evidence to answer the two research questions: (1) How do parents and caregivers of young children engage with the *CDC's Milestone Tracker* app? and (2) What features of the *CDC's Milestone Tracker* app are engaging for parents and caregivers of young children? Overall, pre-use ratings of *CDC's Milestone Tracker* app were relatively high which demonstrates that upon being introduced to the app, parents and caregivers had positive reactions toward the app. This aligns with other evaluations of LTSAE print materials and further substantiates the LTSAE suite of materials as valuable tools to improve early identification of developmental delays and disabilities (Abercrombie et al., 2021; Bright et al., 2019; Chödrön, Pizur-Barnekow, et al., 2021; Daniel et al., 2009; Gallagher et al., 2019; Graybill et al., 2016). From heat map analysis and system usage data it is clear that cognitive and affective engagement are driven (at least in part) by parents' and caregivers' satisfaction with the videos included in the milestone checklists and that the visual appeal of the app, which make it inviting and relatable. The incredibly high checklist completion rate (85.0%)—reflective of behavioral engagement—is a good indication that users find the checklists easy to use and complete in one sitting.

Among the 45 respondents who said they planned to use the app at baseline, 77.8% (*n* = 35) downloaded the app, which is considerably higher than the median conversion rate of 18.5% for health and fitness apps (Knotko, 2019). Respondents felt the app was easy to use, which was also reflected in the pre-use heat map tasks and accompanying open-ended responses. Notably, respondents in this sample did not report encountering barriers known to reduce app use and engagement (Brager et al., 2019), as evidenced by being able to find the answer they needed in the app and not being concerned about Internet data usage. A concern voiced by pediatric clinicians discussed in Chapter 4 was that parents and caregivers are too busy to use the *CDC's Milestone Tracker* app, however this was not a barrier mentioned by the parents and caregivers in this study. This is likely a function of this sample being both digitally savvy and relatively affluent, reducing financial burdens that other less affluent groups may experience when using apps (Suh et al., 2016). There is a future opportunity to explore barriers to use with groups who may struggle with technology and/or are less affluent users.

This study sought to identify factors that influence engagement among groups experiencing disadvantages using a conceptualization of engagement as multidimensional as suggested by Kelders et al (2020) and others. Given some of the limitations of the data, the analysis plan changed post hoc to develop a new subjective engagement scale for behavior change apps. The new behavior change app engagement scale created for this study drew on constructs identified in the systematic review of behavior change apps for groups experiencing disadvantages (discussed in Chapter 2) and the TWEETS. This new behavior change app engagement scale—which includes the behavioral, cognitive, and affective processes that characterize engagement—demonstrated high internal consistency, an indication that the scale items are measuring the same construct (i.e., engagement with the app). Further there was a positive correlation between the scale at baseline and follow-up. These findings add support to the argument that engagement is indeed multidimensional and as such, measurement should go beyond system usage data to examine all three dimensions of behavior, cognition, and affect (Kelders & Kip, 2019; Kelders, Kip, et al., 2020; Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017; Yardley, Spring, et al., 2016). While the initial scale demonstrated good internal consistency and decent test-retest reliability, the sample was too small to perform the necessary tests to evaluate the validity of the scale, which would require at least 200-300 respondents to develop a stable and generalizable scale (Boateng et al., 2018). As such, the scale should be re-examined with a larger sample of parents and caregivers from groups experiencing disadvantages to perform an exploratory factor analysis and evaluate the internal consistency, reliability, and validity (convergent, divergent, and predictive) of the behavior change app engagement scale.

Participants in this study scored relatively highly on the behavior change app engagement scale, an indication that they found the *CDC's Milestone Tracker* app engaging, largely based on the apps' clean aesthetic and the videos included in the milestone checklists. While average app rating were relatively high and most of the sample said they would recommend the app to friends and family, these measure of satisfaction should be explored in future studies to understand how they relate to engagement and app use.

As the two "best bet" items assessing respondents' perceptions of engagement (how engaging the app is and if they like using the app) were only asked on the follow-up survey, they were not included in the behavior change app engagement scale. While Perski et al (2019) found that the item asking users how engaging an app is was not a significant predictor of app logins, they did find that asking users how much they liked the app predicted future behavioral engagement. There is value in exploring the relationship between these engagement items and other objective measures of app use (e.g., time spent in the app or number of tasks completed).

Findings from this study highlight the temporal nature of engagement. Interestingly, some parents and caregivers noted that they do not think an app used for periodic health tracking (e.g., *CDC's Milestone Tracker*) necessarily needs to be engaging. Compared to some behavior change apps for tracking diet, exercise, and sleep that require active user participation on a more frequent basis to be effective (McVay et al., 2019), the *CDC's Milestone Tracker* app was designed to correspond with the schedule for well-child visits for children birth to five years (Zubler et al., 2022). Similar to other studies of "just in time" resources (Taki et al., 2017), engagement with the app seemingly fluctuates with the timing of the milestone checklists. This finding supports the idea that engagement encompasses different stages or phases (O'Brien & Toms, 2008) that are dependent on the needs of the user, the purpose of the app, and the behavior the app is designed to change. Users may move from initial engagement (e.g., creating a child profile and completing their first milestone checklist) to periods of disengagement (e.g., when there are no new checklists to complete) to re-engagement (e.g., when a well-child visit is approaching and there are checklists available for their child's age), which suggests that the quality of the behavior may be more important for engagement than quantity (or frequency) of the behavior (Kelders & Kip, 2019). While developmental surveillance is a continuous process, the number of logins per day or number of minutes used per week are not relevant or necessary for this type of behavior change app and checklist completions may be a better indicator of engagement.

The question then becomes how to encourage initial engagement and re-engagement with the app so that parents and caregivers complete the milestone checklists and act early when concerns are identified. Respondents' suggestions for timely push notifications as prompts to remind users to revisit incomplete checklists may help to address this issue. For example, incorporating GPS-based notifications to remind users to share any potential concerns while they are at a well-child visit might increase the number of families with concerns who raise those concerns to their clinician. Notably, prompts or cues are a BCT with strong evidence of links to three MoAs: (1) memory, attention, and decision processes, (2) environmental context and resources, and (3) behavioral cueing (Carey et al., 2018). Given that parents and caregivers who downloaded the app showed no changed in awareness of developmental milestones from baseline to follow-up, more information about child development within the app would be beneficial. One option might be to incorporate features like digital versions of LTSAE's children's books to support early development. Additional content embedded within the app might remind parents and caregivers that developmental surveillance is an ongoing process and bolster re-engagement.

A strength of this study was that it was conducted with the intended users of the app—parents and caregivers of young children who experience disadvantages—rather than university students or the general population. This provides a more authentic depiction of engagement among this audience. Another strength of this study was that it was designed to explore engagement as it might naturally occur, whereby participants were introduced to *CDC's Milestone Tracker* through the activities within the baseline survey. A link to download the app was included in the closing screens of both the baseline and follow-up surveys but participants were not instructed to download the app. Much app evaluation research has limited external validity (Baumel et al., 2019), since studies often employ designs that incentivize participants to download and use an app (or in some cases, equip participants with devices and data to use throughout the duration of the intervention). This essentially eliminates some of the major barriers to engagement, including high attrition of apps (Eysenbach, 2005). While these research methods are useful to inform app development and design, observing and assessing engagement in an ecologically valid way likely provides more accurate translation to everyday settings because it provides deeper insight into how people actually use an app and the reasons for non-use (Baumel et al., 2019; Roberts et al., 2019). The combination of pre-use questions to understand subjective perceptions of engagement and questions at one-month follow-up to explore actual engagement with the app enhanced the findings of the current study. This might be a good approach for apps that have already been developed and deployed and have at least some system usage data to inform updates to see how they could be improved and what might make them more engaging.

Future Directions

This study assessed engagement with *CDC's Milestone Tracker* app among parents and caregivers of young children from groups experiencing disadvantages. Utilizing a conceptualization of engagement as a multidimensional process involving behavioral, cognitive, and affective factors was useful to understand parent and caregiver engagement with *CDC's Milestone Tracker* app. Given the promising findings using the initial version of the behavior change app engagement scale developed for this study, more research is needed to further refine the scale and use it to explore engagement with this and other health apps. The features of the app that parents and caregivers find the most engaging are the milestone checklists, particularly the videos embedded in the checklists, but a closer examination of which checklists are preferred and why would be helpful to increase the already high checklist completion rate. Many open-ended items included in the two surveys were not fully analyzed due to time constraints, and more in-depth qualitative analyses might help more deeply understand parent and caregiver engagement with the app. Finally, some of the recommendations parents and caregivers provided on the follow-up survey are important to revisit to inform future updates to the app.

Implications for Dissertation Research

Findings from this study add to results from the first study presented in Chapter 2 (Aim 1) and the third study discussed in Chapter 4 (Aim 3). Connections to each chapter are described below and explore in more depth in Chapter 5.

Chapter 2: Assessing Engagement with Digital Behavior Change Interventions for Groups Experiencing Disadvantages: A Systematic Review

The findings from the systematic review helped to inform the design and direction of the study discussed in this chapter. Many of the measures included on the two surveys conducted with parents and caregivers were derived from studies included in the systematic review and modified for *CDC's Milestone Tracker* app. For example, constructs identified in the review were used to inform items to include in the initial behavior change app engagement scale. Additionally, the number of milestone checklist ages viewed and completed served as a proxy for intervention dosage and interactions with app features (Hartzler et al., 2016; Mauriello et al., 2016; Pratap et al., 2018). Given that most articles included in the systematic review examined system usage data, it was important to include *CDC's Milestone Tracker* app as a secondary data source. While individual usage data would have been preferable, the aggregated data was still informative to answer the study research questions.

Chapter 4: Pediatric Clinician Perceptions Regarding Engagement with CDC's Milestone Tracker App

Many of the findings from this chapter converge with the findings reported in Chapter 4. As previously mentioned, awareness of LTSAE, initial reactions to *CDC's Milestone Tracker* app, and if participants would recommend the app was asked of both parents and caregivers and pediatric clinicians. Across the two studies, participants' responses related to affective engagement can be attributed to the videos, images, and colors that create a visually appealing and inviting app. Further, within the context of cognitive engagement, some of the concerns raised by the pediatric clinicians who participated in the interviews described in Chapter 4 were negated by the parents and caregivers in this sample, highlighting the importance of including the perspectives of both families and those who provide care to families in research about behavior change apps.

Study Limitations

This study had several limitations. First, while most participants self-identified as non-White and/or people with disabilities, I intended for a more economically diverse sample by recruiting via a wider range of channels but because the survey was attacked by bots (twice) the recruitment strategy relied on my personal network which the sample is a reflection of in terms of sociodemographics (e.g., highly educated, more affluent). This convenience sample limits generalizability of the findings. Second, there were flaws in the study design where there was likely too short of a time frame between the baseline and follow-up survey to truly assess engagement with *CDC's Milestone Tracker* and anticipated outcomes of using the app. Milestone checklists are only available for certain ages and as children get older, well-child visits are less frequent meaning there was less time for parents and caregivers to (1) complete milestone checklists, (2) identify concerns, and (3) take early action by raising those concerns with a clinician or other early childhood care provider. With the known delays in initiating EI services (Scherr et al., 2020), one month would not have been enough time to initiate services based on concerns identified within the study period. While beyond the scope of this study, re-evaluating behavioral outcomes six months after parents and caregivers start using the app might provide better insight into engagement behaviors and preferences and would be a better indicator if respondents continue to use and engage with the app.

Finally, the *CDC's Milestone Tracker* app data was only shared at the aggregate level which made it impossible to compare the subjective measures of engagement assessed in this study with objective measures of engagement. System usage data linked directly to participants combined with self-report engagement and perceptions of engagement would have been preferable and would have strengthened the findings from this exploratory study. Future research should collect data exploring behavioral, cognitive, and affective processes via surveys, questionnaires, or interviews in conjunction with system usage data at the individual level to examine engagement preferences and behaviors.

Conclusions

This chapter explored parent and caregiver engagement behaviors and attempted to identify factors that influence engagement with the *CDC's Milestone Tracker* app (Aim 2).

Parents and caregivers rated the app highly and found it easy to use and engaging. Considering engagement as multidimensional (inclusive of behaviors, cognition, and affect) is a helpful framework to understand engagement within the context of behavior change apps. Results from this study can assist with informing best practices in evaluating the *CDC's Milestone Tracker* app and may help to inform the development of an engagement measurement framework. Findings from this study are revisited in Chapter 5 of this dissertation.

Chapter 4

Pediatric Clinician Perceptions Regarding Engagement with *CDC's Milestone Tracker* App

Introduction

One in six children has a developmental delay or disability (Zablotsky et al., 2019), yet most are not identified as early as possible (Maenner et al., 2021; Rosenberg et al., 2008; van 't Hof et al., 2020). Early identification of developmental delays and disabilities is necessary to initiate EI that can improve language, physical, cognitive, social-emotional, and educational outcomes for young children and families (Adams & Tapia, 2013; Bruder, 2010). Children with identified delays who receive interventions at younger ages are more likely to benefit from EI services (Bradshaw et al., 2015). Yet children from lower socioeconomic backgrounds experience disparities in terms of screening and identification of delays, and participation in EI (Barger et al., 2022; Shaw et al., 2021; Zuckerman et al., 2014). Current AAP guidelines recommend periodic developmental and autism screening using validated screening instruments (e.g., Ages and Stages Questionnaire [ASQ], Parents' Evaluation of Developmental Status [PEDS], and Modified Checklist for Autism in Toddlers [M-CHAT]) at specific intervals during well-child visits or when a parent or caregiver raises a concern about their child's development (Lipkin, Macias, Council On Children With Disabilities, et al., 2020). Further, AAP guidelines recommend developmental surveillance (also known as developmental monitoring) which includes routinely eliciting parent or caregiver concerns and observing the child at every well-child visit (Lipkin, Macias, Council On Children With Disabilities, et al., 2020). As such, developmental surveillance requires participation from parents and caregivers as well as the early childhood providers and clinicians that care for and interact with young children. While the use of developmental screening tools has tripled since 2002 (Lipkin, Macias, Baer Chen, et al., 2020), more than 20 years after the AAP's recommendations, only 37% of parents and caregivers of children under three years report receiving developmental surveillance (Hirai et al., 2018). Children who receive a combination of developmental surveillance and screening are more likely to receive EI services (Barger et al., 2018), yet only 1 in 5 children under age three received both surveillance and screening (Hirai et al., 2018). Additionally, clinicians may overestimate how often they discuss screening with families and make referrals for EI services, indicating that strategies to improve these practices are needed (Bright et al., 2019).

The *CDC's Milestone Tracker* is a behavior change app with seven BCTs that was designed to help families (particularly those from groups experiencing disadvantages) learn about child development, track developmental milestones, and take early action if they have a concern about their child's development. Part of a suite of materials from the LTSAE program (www.cdc.gov/ActEarly), the interactive app allows users to complete age-appropriate milestone checklists, which include photos and videos of developmental milestones, written descriptions of developmental milestones, and space to enter notes. The app also includes appointment reminders and tips and activities to support early childhood development. While the app was designed to meet the needs of parents and caregivers, secondary audiences are early childhood providers and clinicians to support developmental surveillance. Notably, the app encourages parents and caregivers to share the summary of their child's development with their doctor or other clinician. While previous research demonstrates positive outcomes associated with LTSAE materials and resources (e.g., increased knowledge and awareness of developmental milestones, confidence to discuss concerns, and positive attitudes toward developmental surveillance; Abercrombie et al., 2021; Campbell et al., 2019; Chödrön, Barger, et al., 2021; Daniel et al., 2009; Gadomski et al., 2018; Graybill et al., 2016), there has only been one published evaluation of *CDC's Milestone Tracker* (Armenta et al., 2019), which assessed the initial version of the app (released in 2017). Further, formative research to develop the app was primarily conducted with parents and caregivers, with limited input from clinicians (Muñoz & Arriaga, 2015). As such, it is unknown how clinicians view the app and what role the app might play in how developmental surveillance and screening is conducted.

Engagement is comprised of cognitive (e.g., attention, interest, and immersion in the behavior change app), affective (e.g., enjoyment and satisfaction), and behavioral (e.g., behavior change app usage) processes that characterize the user experience with behavior change apps (Cole-Lewis et al., 2019; Flaherty et al., 2021; Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017; Yardley, Spring, et al., 2016). This multidimensional concept functions as the interplay between behavioral, cognitive, and affective factors and is believed to be a facilitator of health behaviors (Kelders, van Zyl, et al., 2020; Perski, Blandford, et al., 2019; Short et al., 2018). To date, engagement with *CDC's Milestone Tracker* app has not been evaluated but might help to provide insight to how the app can improve early identification of developmental delays and disabilities.

Historically, clinicians have been considered trusted sources of health information (Pearson & Raeke, 2000), although the current sociopolitical climate has seemingly eroded some of that trust (Arora et al., 2020). Nevertheless, clinicians play a key role in disseminating information (particularly about child development). As such, clinician's recommendations may have implications for behavior change app adoption and effectiveness (Womack et al., 2018). For example, a systematic review of engagement and recruitment to digital health interventions found patients more likely to sign up for DBCI if supported by clinicians and trusted healthcare organizations, whereas lack of clinical endorsement was a barrier to DBCI uptake (O'Connor et al., 2016). Similarly, Taki et al (2017) found that mothers referred by their health practitioner to an infant feeding app had higher engagement scores compared to participants recruited via other channels (e.g., the web). The authors hypothesize that this is likely a result of mother's perception of health practitioners as a trustworthy source of information and suggest that they are important "referral pathways" to evidence-based apps (Taki et al., 2017).

Accordingly, it can be reasoned that clinicians might influence adoption, use of, and engagement with *CDC's Milestone Tracker*. The purpose of this study was to explore perceptions regarding engagement with *CDC's Milestone Tracker* among pediatric clinicians at FQHCs. FQHCs serve nearly 29 million people, primarily from groups experiencing disadvantages, including 1 in 3 people experiencing poverty and 1 in 5 people who are uninsured (Health Resources and Services Administration, 2021). This study aimed to answer the following research question: What are FQHC pediatric clinician perceptions of the potential utility of the *CDC's Milestone Tracker* app? Given the weight of clinician referrals and that clinicians are a secondary audience for the app, this chapter reports preliminary results on engagement with *CDC's Milestone Tracker* from clinician's perspective as a complement to findings presented in Chapter 3.

Methods

Study Design and Sample

This study was a sub-study of a research project to understand how pediatric clinicians assess social-emotional development during well-child visits with toddlers (ages 2 to 3 years). A non-purposive sampling strategy was used to recruit participants from Alliance Community Health Centers (AllianceChicago), a network of more than 50 FQHCs in 20 states that serve more than 3.5 million people from diverse communities including from urban and rural communities, people who are low income and uninsured, people experiencing homelessness, people from minoritized racial and ethnic groups, LGBTQ community, and refugee populations (AllianceChicago, 2022). Potential participants were recruited via email between June and November 2021 by a co-investigator who works at AllianceChicago. To be eligible, participants needed to be a United States-based practicing primary care clinician (e.g., pediatrician, family medicine physician, advanced nurse practitioner) who provides care for toddlers (ages 2 to 3 years) at an FQHC and can communicate in English. If a potential participant agreed to learn more about the study, a member of the research team contacted the participant to schedule an interview and share the informed consent document. Clinicians received $100 in compensation for their time. This research received approval from the Institutional Review Board at Northwestern University (IRB STU00214600).

Two researchers (HGS and CLS) trained in qualitative methods conducted individual in-depth interviews using a semi-structured interview guide that was piloted with two pediatric clinicians (Appendix I). During the interview, the interviewer shared their screen to present a visual demonstration of *CDC's Milestone Tracker*. The demonstration consisted of a 45-second screen capture video of the app to show different features including the ability for a parent or caregiver to email or show a summary of their child's progress to a clinician for any milestones a child is missing or where they have concerns. After the video, participants saw the *CDC's Milestone Tracker* promotional flyer and two screenshots of the app that displayed different features of the app in two languages (one in English and one in Spanish; see Figure 7). During the demonstration, the interviewer gave a verbal description of app features. Interviews were conducted via Zoom and lasted between 33 and 55 minutes. Saturation, or the point when little to no new information relative to the study was identified (Guest et al., 2006), was reached after the first eight interviews, however additional interviews were conducted with clinicians at different clinics. After providing verbal consent, all interviews were recorded, professionally transcribed, and transcripts were reviewed for accuracy.

Measures

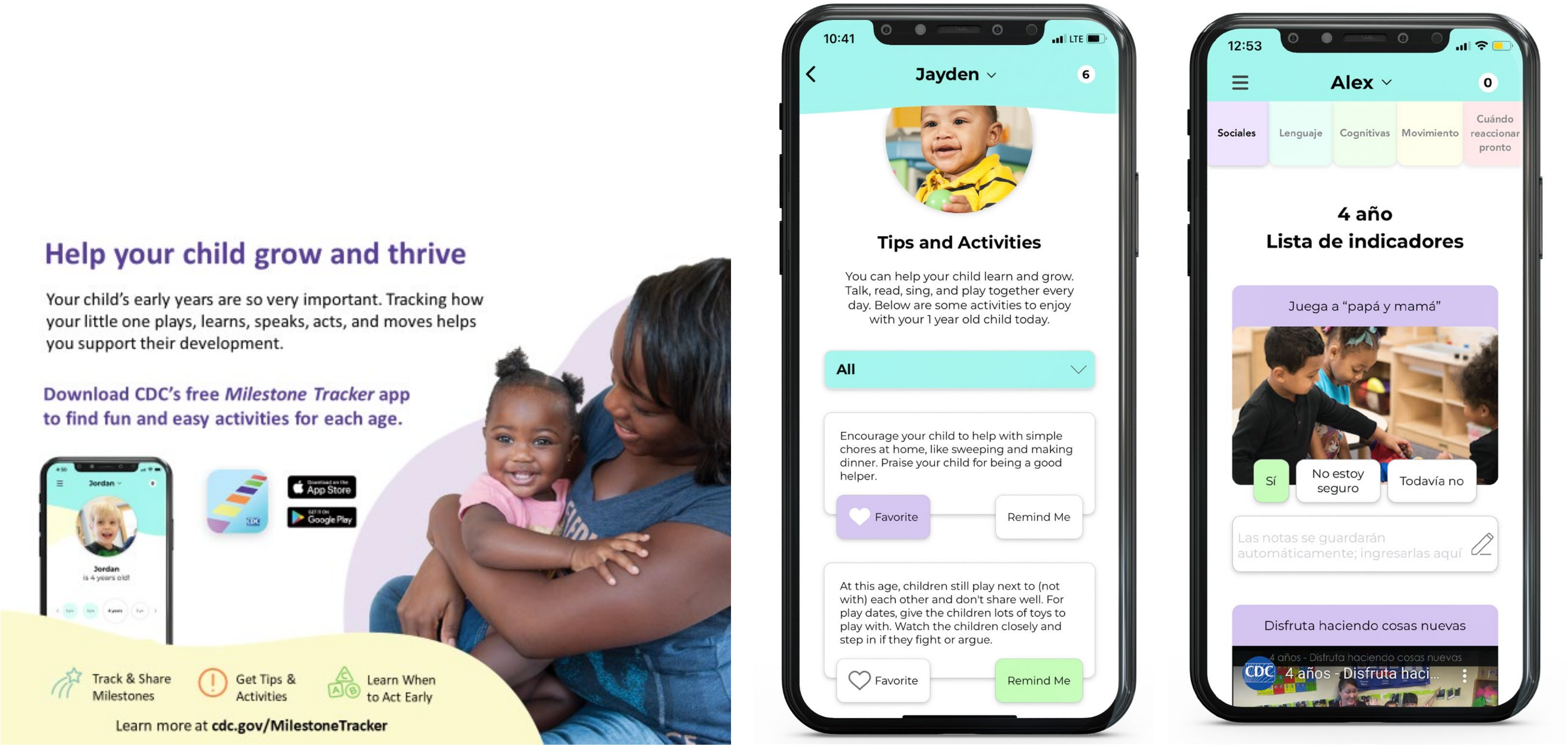
The interview aimed to understand if and how clinicians assess toddler development, including social-emotional well-being, how they discuss developmental and behavioral concerns with families, what resources they use to assess development, and their perceptions of the utility of *CDC's Milestone Tracker*. The sections of the interview guide that asked questions about general descriptions of well-child visits, assessing child development, resources, and sociodemographics were pertinent to this sub-study.

Descriptions of Well-child Visits and Assessing Child Development

The interview started with participant's describing their process for a well-child office visit with a toddler. Questions to understand processes for assessing child development were adapted from an AAP and CDC study to understand developmental surveillance in pediatric practice conducted by Gerndt and Mitchell (Forthcoming 2022). Participants were asked "Do you typically perform developmental surveillance (i.e., monitoring) or screening during routine

Figure 7

Visual Aids used for Demonstration of CDC's Milestone Tracker App



Help your child grow and thrive

Your child's early years are so very important. Tracking how your little one plays, learns, speaks, acts, and moves helps you support their development.

Download CDC's free *Milestone Tracker* app to find fun and easy activities for each age.

Learn more at **cdc.gov/MilestoneTracker**

Tips and Activities

You can help your child learn and grow. Talk, read, sing, and play together every day. Below are some activities to enjoy with your 1 year old child today.

Encourage your child to help with simple chores at home, like sweeping and making dinner. Praise your child for being a good helper.

At this age, children still play next to (not with) each other and don't share well. For play dates, give the children lots of toys to play with. Watch the children closely and step in if they fight or argue.

4 año

Lista de indicadores

Juega a "papá y mamá"

Las notas se guardarán automáticamente; ingresarlas aquí

Disfruta haciendo cosas nuevas

well-child visits?" and if they responded yes, were probed on how they practice developmental surveillance or screening and what tools and resources they use.

Resources

Constructs identified in the systematic review (described in Chapter 2) such as awareness, attention, perceived use, and facilitators, and barriers to use were asked in the resources section of the interview guide. For awareness, participants were asked questions adapted from Gerndt and Mitchell (Forthcoming 2022): "Are you familiar with the Centers for Disease Control and Prevention's (or CDC) *Learn the Signs. Act Early.* materials/resources?" and "If yes, do you use any of these materials/resources during well-child visits? Which ones?" After the demonstration of *CDC's Milestone Tracker*, participants were asked a question to gauge attention modified from Short et al (2018): "I'd like you to please share your initial reactions to this feature and the app in general. Do you see any components that might catch your attention? If so, which components? What would you do if a parent or caregiver shared this information with you?" Participants were also asked questions about expectations ("What impact do you think this app might have on how you conduct developmental surveillance during well-child visits?) and perceived use ("How do you think this app would help you discuss a parent or caregiver's concerns about their child's development?). Finally, participants were asked, "Would you recommend this app to parents and caregivers? Why or why not?," an item modified from the satisfaction questionnaire in the engagement index by Taki et al (2017) and the app subjective quality (section E) of the Mobile Application Rating Scale: user version (uMARS) by Stoyanov et al (2016). A similar question was also asked of parents and caregivers on the baseline and follow-up surveys described in Chapter 3.

Sociodemographics

Sociodemographics included self-reported gender, race and ethnicity, number of years practicing medicine, practice type, insurance type for patients served, and geographic area of practice.

Data Analysis

Qualitative data analysis of the individual in-depth interviews consisted of two processes. The first process used applied thematic analysis to identify and recognize patterns within the data, which in turn, became common themes within the different experiences of participants (Guest et al., 2012). Applied thematic analysis, which draws on features of grounded theory and phenomenology, was appropriate for this study because it offers a systematic, transparent, and efficient way to identify key themes within the text that were transformed into codes. Further it is useful to find solutions to real-world problems (Guest et al., 2012). To become familiar with the data, the interview transcripts were reviewed for accuracy. A codebook was developed using open coding through an inductive approach based on 10% ($n = 2$) transcripts. Two researchers (HGS and CLS) independently applied the codebook to 15% ($n = 3$) of transcripts and made revisions to the codebook after meeting to discuss each coded transcript.

The second process to analyze the data utilized a content analysis using a directed approach. Directed content analysis can offer supporting evidence to prior research, theories, or frameworks (Hsieh & Shannon, 2005). Using a deductive approach, a codebook was developed with predetermined categories that consisted of primary (or parent) codes based on the behavioral and psychosocial dimensions believed to constitute engagement (behavior, cognition, and affect; introduced in Chapter 1 and discussed further in the systematic review in Chapter 2; Kelders, van Zyl, et al., 2020). Sub codes that arose from pediatric clinicians' perceptions of and experiences with *CDC's Milestone Tracker* were added to each of the three primary codes. Two additional primary codes were added to capture the reasons participants would or would not recommend *CDC's Milestone Tracker* app and their suggestions for how to improve the app (see Table 11 for coding scheme). This coding scheme was piloted and refined along with the applied thematic coding described above.

Once a functional codebook was established, two researchers (HGS and SM) systematically review transcripts to assign codes using qualitative data management software MAXQDA Analytics Pro 2020 (VERBI GmbH Software, Berlin, Germany; Guest et al., 2012). The researchers coded 15% (*n* = 3) transcript together, 10% (*n* = 2) of transcripts independently, and then met to resolve discrepancies. Intercoder agreement was assessed on three (15%) additional transcripts coded independently using ReCal software (Freelon, 2010), achieving an acceptable level of agreement for a Krippendorff's alpha ($\alpha = .77$; Krippendorff, 2004). Additionally, to increase transparency, an audit trail was used to keep track of and document the entire data analysis process (e.g., analytic activities, data included and excluded in analysis and rationale for decisions, methods used to find themes and apply codes, changes to the codebook, and any coding checks; Guest et al., 2012). Sociodemographic data were analyzed using descriptive statistics showing distributions in SPSS 28 statistical software (IBM Corp., Armonk, New York).

Table 11

Qualitative Data Analysis Coding Scheme

Theme	Subtheme	Operational Definition
Behavioral engagement	Use (clinicians)	Descriptions of how participants (clinicians) would use the *CDC's Milestone Tracker* app and/or what impact they think it might have on developmental surveillance.
	Use (parents and caregivers)	Participants' perceptions of how parents and caregivers might use the *CDC's Milestone Tracker* app. This code includes both how clinicians view parent and caregiver use or non-use or their perceptions based on experiences (for those who currently recommend the app).
	Perceived audience	Perceptions of who would use the app (i.e., what type of parent or caregiver) or who the intended audience is of the app. Can also include which audiences they think would not use the app. Includes personality traits and demographics.
Cognitive engagement	Questions about the app	Any questions participants ask about the app, including questions about app features, intended audiences, app development, app utility, and app integration into processes and workflows.
	Concerns about the app	Any concerns raised by participants about the app, app features, app development, and app utility. This includes potential barriers to use and challenges related to recommending the app, using the app, and integrating the app into workflow.
Affective engagement	Satisfaction with the app	Participant comments about what they like about the app, including positive reactions to app features, look and feel, intended audiences, app development, app utility, and app integration into processes and workflows.
	Perceived usability of the app	Participant comments about how easy or difficult it would be to use the app, either from their perspective or the perspective of parents and caregivers.
Reasons would recommend or not recommend MT	---	Any reasons provided for why providers would or would not recommend the app to parents and caregivers. If they already recommend the app, capture reasons for why they currently recommend the app.
Areas for improvement	---	Provider's specific recommendations for how the app could be improved, including recommendations for app features, look and feel, integration into process and workflows, and promotion.

Note. MT = *CDC's Milestone Tracker* app.

Results

Sample Characteristics

All 20 pediatric clinicians who were contacted about this study completed an interview. The majority self-identified as non-Latinx White (*n* = 14, 70.0%) and as women (*n* = 15, 75.0%). The sample consisted of doctors of medicine/doctors of osteopathic medicine (*n* = 14, 70.0%), advanced nurse practitioners/nurse practitioners (*n* = 3, 25.0%), and physician assistants (*n* = 1, 5.0%) who have practiced medicine for 12.6 years on average (*SD* = 8.0). All 20 were affiliated with one of seven Chicago-area FQHCs, all of which primarily serve patients enrolled in Medicaid, however some were also affiliated with a university or medical school (*n* = 3, 15.0%), hospital (*n* = 3, 15.0%), and an independent practice (*n* = 1, 5.0%; Table 12).

Assessing Child Development

All participants (*n* = 20, 100.0%) reported performing developmental surveillance or screening during routine well-child visits with toddlers. Most (*n* = 17, 85.0%) reported regularly performing social-emotional surveillance or screening during routine well-child visits with toddlers (findings on social-emotional development discussed elsewhere; Scherr et al., In preparation). Most use the ASQ (*n* = 17, 85.0%;) to conduct developmental screening and the M-CHAT (*n* = 18, 90.0%) to conduct screening for autism. Participants also reported using other tools to conduct developmental surveillance and screening. Other tools to perform developmental surveillance and screening included: Ages & Stages Questionnaires: Social-Emotional (ASQ:SE), Screening Tool for Autism in Toddlers and Young Children (STAT), Survey of Well-being of Young Children (SWYC), CDC's LTSAE, books from Reach Out and Read (reachoutandread.org), blocks and toys from Prescription for Play by the LEGO Foundation

Table 12

Sample Characteristics (N = 20)

	n	%
Gender		
Woman	15	75.0
Man	5	25.0
Race and Ethnicity		
Non-Latinx Asian	3	15.0
Latinx	2	10.0
Non-Latinx Multiracial	1	5.0
Non-Latinx White	14	70.0
Clinician Type		
MD/DO	14	70.0
APN/NP/FNP/PNP	5	25.0
PA	1	5.0
Number of Years Practiced Medicine (*M, SD*) [a]	12.8	8.0
Practice Type [b]		
Federally Qualified Health Center (FQHC)	20	100.0
Affiliated with University or Medical School	3	15.0
Hospital-affiliated Practice	3	15.0
Independent Practice	1	5.0
Insurance Type of Patients Served		
% Enrolled in Medicaid (*M, SD*)	86.3	9.9
% Enrolled in Medicare (*M, SD*)	0.4	1.1
% Private Insurance (*M, SD*)	7.1	5.4
% Self-Pay or Uninsured (*M, SD*)	6.5	8.6
Perform Developmental Surveillance or Screening (Yes)	20	100.0
Perform Social-Emotional Surveillance or Screening (Yes)	17	85.0
Use Ages & Stages Questionnaires (ASQ; Yes)	17	85.0
Use Modified Checklist for Autism in Toddlers (M-CHAT; Yes)	18	90.0
Use Other Tools to Perform Surveillance or Screening (Yes) [c]	17	85.0

Note. APN = Advanced nurse practitioner; DO = Doctor of osteopathic medicine; FNP = Family nurse practitioner; MD = Doctor of medicine; NP = Nurse practitioner; PA = Physician Assistant; PNP = Pediatric nurse practitioner.

[a] Number of years practiced medicine at time of interview range 2-30. [b] Participants could select all practice types that apply therefore percentages do not add up to 100%. [c] Other tools to perform developmental and social-emotional surveillance and screening include: Ages & Stages Questionnaires: Social-Emotional (ASQ:SE), Screening Tool for Autism in Toddlers and Young Children (STAT), Survey of Well-being of Young Children (SWYC), CDC's *Learn the Signs. Act Early.*, books from Reach Out and Read, blocks and toys from Prescription for Play by the LEGO Foundation, and developmental questions built into electronic medical record (often based on Bright Futures).

(www.rx4play.org), and developmental questions built into the electronic medical record, often based on AAP's Bright Futures (brightfutures.aap.org; Table 12).

Awareness of Resources and Tools to Support Developmental Surveillance

A quarter of participants (*n* = 5, 25.0%) were aware of LTSAE materials and resources and 15.0% (*n* = 3) currently use *CDC's Milestone Tracker* as an additional developmental surveillance resource for families. Those who currently use *CDC's Milestone Tracker* explained they occasionally recommend the app to parents and caregivers, mostly informing them that it exists or in situations where they have identified some concerns and think the parent or caregiver might be reluctant to self-identifying developmental differences. As one participant explained:

> I just keep it on my phone, and then I'll just show them. I'll be like "Oh, if you have questions or concerns, it's a good thing." I include it, usually, in the patient handouts, and then just I don't... I keep them [milestone checklists] up-to-date on my own children. So I'll just show them like, "Oh, this can be really helpful for just going through and seeing what they're doing." (P10, nurse practitioner)

Participants noted that parents and caregivers who received information about the app have not mentioned it in subsequent appointments and therefore the clinicians were not sure how much they used the app. One participant reported that they have shared the app with the fellows they work with to increase awareness of developmental milestones among trainees. When asked how they learned about the app, participants explained that they came across it while looking for resources on the CDC's website or that they were introduced to it through a state-level early identification task force. One additional participant was aware of *CDC's Milestone Tracker* but explained they had not yet figured out how to integrate the app into their workflow.

Given that engagement is theorized to be comprised of behavioral, cognitive, and affective processes (Cole-Lewis et al., 2019; Kelders, Kip, et al., 2020; Kelders, van Zyl, et al., 2020; Perski, Blandford, West, et al., 2017; Yardley, Spring, et al., 2016), the results in the following sections are organized by each of the three dimensions and include illustrative quotations by clinician type.

Perceptions about Behavioral Engagement

Behavioral engagement focuses on how individuals use an app. A handful of participants used the terms "engage" and "engagement" when describing perceptions about anticipated behaviors. Some discussed how parents and caregivers might engage in the intended behavior (i.e., engage in developmental surveillance), stating that the app can "engage the parent or caregiver" in their child's health and development. However there was a sentiment that these developmental surveillance behaviors were mediated by the clinician, as explained by a nurse practitioner, "there has to be some engagement with the provider I think in order for it to work" (P16). Others described engagement with features of the app. One participant remarked, "I guess really anything that can sort of engage the parent has probably really high value" (P20, physician), attributing these benefits to the fact that "apps is where a lot of their [parent and caregiver] engagement and learning happens" (P12, physician).

Participants in this sample also described their perceptions about behaviors related to engagement for clinicians' use, parents and caregivers use, the perceived audience for *CDC's Milestone Tracker*, and intentions to recommend the app to parents and caregivers. Each is described in more detail in the following sections.

Utility for Clinicians

Nearly all participants (n = 19, 95.0%) described how they would use *CDC's Milestone Tracker*. Participants described hypothetical scenarios about how the app could help provide additional information about the patient, serving as a useful talking point to elicit more from the parent or caregiver. Several explained that they could potentially integrate data from the app with information gained from developmental assessments, observations, and a patient's medical history. They emphasized that if families were to use the app it would save time during the visit:

> It would be great if parents would already come with…already going through this on their own before coming in. And it might save me some time too with providing education because there's great deal of the well-check is providing education to parents. So they if already have this tool, they have the resources kind of at their hands. (P14, physician assistant)

Additionally, participants discussed the impact the app might have on developmental surveillance and screening, sharing ideas of how it could be introduced to patients and their families:

> I think again, bringing anything to the family is always helpful, and if it's something that they could do prior to the visit, with an email from our office that says, "Hey, make sure you fill this out before your visit," that definitely eliminates some barriers of filling out more paperwork. And if it's more interactive like that, I think it's always a hook for a family. (P17, physician)

Participants expanded on how parents and caregivers might use *CDC's Milestone Tracker* in the following section.

Utility for Parents and Caregivers

All 20 participants discussed their perceptions about how parents and caregivers would use *CDC's Milestone Tracker*. Participants believed parents and caregivers could use the app as a resource to learn about developmental milestones at their convenience, track their child's development, get concrete ideas for activities to help with development, and pay closer attention to their child's development. They explained how the app might help to facilitate parent and caregiver-engaged developmental surveillance which would, in turn, prepare families to fill out developmental screening forms in clinic, resulting in a more effective and productive visit. Some felt the app could empower parents and caregivers by giving them the language they need to raise concerns. As a nurse practitioner explained:

> I think it will validate them, right? So, if a [parent] is concerned and then they do this, and then this also shows that there's a concern, I think that just gives weight to the parents' thoughts. "Oh, I thought something was wrong here. Now I really am going to speak up about it." And I think that can be a valuable tool, right? Because I think if a provider isn't prodding about development or really asking some targeted questions about development, then a parent might not necessarily bring it up. So, while I can say, "Oh, what are your concerns today?" That doesn't always mean that a parent is going to bring it up proactively unless I say, "Is your child talking? Do you have concerns about your child's speech or about how your child is behaving or processing emotion?" If I don't ask this specifically and a parent does have concerns, they don't always bring it up. So, this is another way to potentially help them bring it forward to their provider, and that's where I think it would be most applied. (P01)

Several participants indicated that the *CDC's Milestone Tracker* would be useful for families given that phones are how many people get information and are connected to the world as evidenced by statements such as “most people are already on their phone all the time anyways” (P19, physician). This suggests there are certain audiences who they deem more likely to use *CDC’s Milestone Tracker*, as described below.

Perceived Audience

Without prompting, most participants (*n* = 17, 85.0%) explained who they believe is the intended audience for *CDC’s Milestone Tracker*, often explicitly describing the characteristics of parents and caregivers (e.g., personality traits, demographics, and life circumstances) who would or would not use the app (see Table 13 for examples). In general, participants in this sample believed there would be low engagement with *CDC’s Milestone Tracker* among their patient population. They thought the app is for digitally savvy parents and caregivers who are interested in their child’s development or already have concerns. One participant explained the perceived audience as:

> Parents that have the time for it…And I’m not inferring that the parents don’t necessarily care. I just think it’s a matter of there are some parents that have the bandwidth that this is something that they are concerned about. And I think other parents have bigger struggles, and this is not on their radar…Like social life struggles, daycare, employment, jobs, financial struggles, transportation. (P20, physician)

While participants had many positive reactions the app, they were skeptical if the app would be appropriate for the families they work with, citing concerns about health literacy, competing priorities, and overall less interest in and familiarity with early childhood development and developmental milestones.

Table 13

Examples of Clinician's Perceptions of the Intended Audience for CDC's Milestone Tracker

MT app is for parents and caregivers who:	MT app is not for parents and caregivers who:
• are motivated, engaged, and interested in child development • are digitally savvy • have more time on their hands • have fewer competing priorities (i.e., more financially stable) • are visual learners • have concerns about their child's development (i.e., are worried) • are unfamiliar with developmental milestones	• are unmotivated or unengaged • do not have a smartphone or tablet • are struggling financially • have limited or unreliable Internet access and cellular data plans • do not speak English or Spanish • have low-literacy skills • do not trust the federal government

Note. MT = *CDC's Milestone Tracker* app.

Intentions to Recommend CDC's Milestone Tracker

Nearly all participants (*n* = 17, 85.0%) said they would recommend *CDC's Milestone Tracker* to the families seen in clinic, particularly those who raise concerns about their child's development. Some participants were hesitant to say they would recommend it outright and would need to test the app first themselves or would want to review data about the validity of the app before recommending it to parents and caregivers. As noted:

> I'm going to have to download it, review it because I tend to be very watchful of what I recommend to parents. But because it's from the CDC almost certainly the answer would be yes, but I still want to play with it a little bit before. (P13, physician)

A few seemed to understand the power they hold in influencing patient's behaviors, imagining that if someone shared completed checklists they would congratulate and thank the parent or caregiver for taking initiative in hopes of encouraging them to continue to use the app. Those who said they would not recommend *CDC's Milestone Tracker* explained they were concerned that no one would use it and would only recommend it if it were required by their clinic.

Cognitive Engagement

Cognitive engagement is presumed to include attention, interest, and immersion in the app. Since participants did not have an opportunity to use the app, perception about cognition was conceptualized as how clinicians process information and their perceptions about how parents and caregivers process information. Participants in this sample described the features that caught their attention, concerns about the app, and questions about the app, all of which are discussed below.

Attention and Initial Reactions to CDC's Milestone Tracker

Participants were asked to share their initial reactions to *CDC's Milestone Tracker* and identify specific features that caught their attention. Responses varied, but comments were primarily positive about the app and its capabilities. They liked the colors and thought the app was visually appealing. Initial reactions most often mentioned the videos of developmental milestones embedded in the milestone checklists and the benefits of watching and learning about age-appropriate milestones, as one physician explained:

> I think it would prompt parents a little bit more to be like, "oh, he's not doing this. Or, oh, he is doing this." I think sometimes they're like, "Yeah, I think they do that." And I think it's harder, but if they're actively tracking it at home, then they're clearly paying a little bit more attention to it. (P06)

There were mixed reactions to the app feature where a user could email or share a child's summary with their clinician. While some said it would be "great," others felt it would be burdensome for clinicians since email does not sync with a patient's electronic medical record (EMR), and some worried about privacy and security related to emails. Even those with secure email said they do not want patients constantly emailing them and would prefer direct integration with the EMR. Additional concerns about the app are described in the following section.

Concerns about CDC's Milestone Tracker

Although participants were not asked directly about potential concerns, 85.0% ($n = 17$) of participants expressed some concerns about *CDC's Milestone Tracker*. Concerns largely focused on logistics, especially related to documentation and integration with current workflows and processes, including outdated EMR and software. Limited time to explain the app to patients was often mentioned as a potential barrier to use in clinic, with some expressing concerns about how self-explanatory the app is and the amount of time it would take to introduce the app, help families download, and learn how to use it. Interestingly, two participants described concerns related to attrition with health apps, where users often download and forget about an app, resulting in non-use. As one participant explained:

> I think some families would use this and follow through and some people would download it and then never use it. I think if they're motivated, they'll use it, but also, just like with any app, if they're not really concerned or not really interested, then they won't. (P18, physician)

Participants recommended push notifications to remind users to complete milestone checklists—a feature that currently exists in the app but was not described in the demonstration during the interview—as a strategy to increase engagement with the app.

Some voiced concerns related to their patient population, explaining there are potential access issues since some families do not have data, high-speed Internet, or a compatible device. They also worried about parents and caregivers who struggle with literacy and those who have limited time and competing priorities and felt they probably would not use the app. One participant felt that the families they work with would not use the app because of mistrust of the federal government, stating “They’re going to believe what their grandma says on Facebook more than they’re going to believe what the CDC says” (P05, nurse practitioner). More than one clinician noted concerns about the lack of diversity in the images of children, parents, and caregivers in the app:

> Are there pictures of diverse kids in there?...think it is really important. Especially if you click on the Spanish option, not all kids who speak Spanish look the same. But I think having images of the kids that I serve I think is always really helpful. (P04, physician)

Finally, while participants repeatedly mentioned that they believe “smart people at CDC created the app,” participants expressed concerns about the validity of *CDC’s Milestone Tracker* and wanted to review the research used to inform app content and research on the effectiveness of the app.

Questions about CDC’s Milestone Tracker

Fewer participants (*n* = 9, 45.0%) asked questions about *CDC’s Milestone Tracker*. Questions were primarily about what languages the app is available in, if it could replace developmental screening tools such as the ASQ, compatibility with the EMR, and if the app would be able to notify the clinician automatically about any red flags of where a patient is missing developmental milestones. Some speculated if and how the app could be integrated into their workflow and hinted they might ask their medical directors if there are opportunities to bring the app to their clinic.

Affective Engagement

Affective engagement includes user satisfaction and enjoyment with the app. Participants in this sample described what they liked or did not like about *CDC's Milestone Tracker* and perceptions of the usability of the app.

Satisfaction and Enjoyment with CDC's Milestone Tracker

Overall, satisfaction and enjoyment with *CDC's Milestone Tracker* was high. Participants reported that they enjoyed the overall look and feel of *CDC's Milestone Tracker*, appreciating that the app is available in two languages and that there are options for both visual and auditory learners. Many stressed the advantages of an app over paper screening tools and handouts, using terms such as "interactive" and "engaging" to describe *CDC's Milestone Tracker*. As one clinician explained:

> I do like the videos. Because a lot of times I think, some of the questions we ask and some of the skills can be a little challenging. So if they could see a video of like, "Oh yeah, my kid does that." Some people aren't audio learners, maybe [for] visual learners seeing a video of a child doing something might jog their memory more so than just asking a question about it. (P19, physician)

They also liked that the app included more updated pictures and examples compared to other screening tools that reference "toys from the 1980s" (P06, physician).

Perceptions of the Usability of CDC's Milestone Tracker

Clinicians' perceptions about usability focused mostly on ease of using the app, considering themselves or patients who receive care at other clinics (in different neighborhoods of Chicago) who might be more digitally savvy as the end users. Participants noted the app looked user-friendly, readily accessible, and does not require much effort on behalf of the user. As one clinician described:

> I think just because it seems to be a more...How do I say it? A more friendly way to bring these things up. I do think it's more user-friendly. A lot of things in medicine are very not patient-friendly or user-friendly. This seems to be more like the color scheme is very user-friendly, so I can see how this is less scary than a doctor telling you. (P08, physician)

While some clinicians in this study worried about their patient population not having enough time to use the app, others speculated that it would be quick for users to complete the checklists within the app, especially since so many parents and caregivers spend a lot of time on their phones.

Suggestions to Improve *CDC's Milestone Tracker*

A little more than one third of participants (*n* = 7) suggested how *CDC's Milestone Tracker* could be improved. Suggestions primarily centered on incorporating the app into their workflow and the EMR to notify the clinician of any areas of concern, however others made more specific suggestions. One such suggestion was to use more inclusive language since other types of clinicians provide care to young children:

> It says click show doctor to share summary during your child's doctor visit. A lot of pediatric providers now are nurse practitioners or physician assistants, so I think it would be really nice if we could change that to show provider or something like that to recognize that some people's primaries aren't always doctors. (P01, nurse practitioner)

Addressing clinicians concerns and suggestions for improvement may help to increase uptake of *CDC's Milestone Tracker* among pediatric clinicians and the families they serve.

Discussion

This study is the first to evaluate *CDC's Milestone Tracker* with clinicians at FQHCs, who primarily serve people from groups experiencing disadvantages. The ways clinicians speculated they and the families they work with would use *CDC's Milestone Tracker* highlights the potential utility and benefits of the app. For example, several participants noted that the app would provide visual examples of the written questions asked on developmental screening tools, which can be difficult for parents and caregivers to answer in clinic. They also believed the app would give parents and caregivers necessary language to talk about concerns and empower them to raise those concerns with their clinician, a deep-rooted barrier to early identification of developmental delays and disabilities (Scherr et al., 2020).

This is also the only study known to explore perceptions about app engagement from the perspective of clinicians, a secondary audience of the app. By eliciting feedback on *CDC's Milestone Tracker*, it is evident that the way clinicians conceptualize app engagement aligns with the definitions posited by Cole-Lewis et al (2019) and others. Many participants comments were related to behavioral engagement with the technology when describing perceptions about behaviors, concerns, and satisfaction with app features. Others described the potential for engagement with the behavior change components of the app when describing the potential for the app to encourage parent and caregiver-engaged developmental surveillance to assess child development. A few clinicians in this sample also keyed into the reason why there is increased interest in studying DBCI engagement when they voiced concerns about high attrition of health apps, where users download and forget about an app thus resulting in non-use (and limited effects and behavior change). Clinicians in this study recognized the same challenges that behavioral scientists and researchers from the computing community grapple with, which further confirms that engagement is an important concept that is of interest to a variety of audiences. This critical issue could be addressed by determining which factors help to facilitate engagement. One such idea was raised by a clinician in this study who suggested reminders and push notifications as a way to increase engagement with the app. While this is already a feature of *CDC's Milestone Tracker*, push notifications are an example of the BCT 'prompts and cues' and have been demonstrated to be effective at increasing medical adherence and intervention uptake among parents (Sardi et al., 2020) and should be revisited for *CDC's Milestone Tracker*.

All 20 pediatric clinicians in this sample reported performing developmental surveillance and screening during routine well-child visits with toddlers, which is considerably higher compared with the national average (63%; Lipkin, Macias, Baer Chen, et al., 2020). As previously mentioned, other studies have found that clinicians overestimate how often they discuss screening with families (Bright et al., 2019). Consistent with previous literature (Bright et al., 2019; Gerndt & Mitchell, Forthcoming 2022), findings from this study indicate that additional tools to facilitate developmental surveillance (and support more efficient developmental screening) are both needed and desired. The *CDC's Milestone Tracker* app could help to fill the gap, but more research is needed to understand optimal integration into existing workflows.

Currently, healthcare providers represent about 2.0% (*n* = 5,983/292,518) of all *CDC's Milestone Tracker* users, however it should be noted that the CDC only started collecting self-report user profile information in the updated version of the app released March 2021. As evidenced by the relatively small number of clinicians familiar with LTSAE and/or using *CDC's Milestone Tracker*, there is a need to increase awareness. Outreach efforts to clinicians should focus on the benefits of the app to enhance developmental surveillance and how it complements—not replaces—formal developmental screening. Importantly, when communicating with clinicians it is necessary to share information on app development and evidence related to outcomes to generate clinician buy-in. For example, emphasizing that CDC applied plain language principles to develop the milestones (Zubler et al., 2022) may help to ease clinician's concerns. Further, given that other studies have identified a relationship between clinicians as the referral source and app uptake and engagement (O'Connor et al., 2016; Taki et al., 2017), it is important to examine this potential causal link between clinician recommendations and engagement with DBCI. While some clinicians in this sample recognized the influence they have on intervention uptake, an explicit reminder about the influence of their endorsement on adoption and use might be helpful.

A benefit of DBCI is that they mimic the experience of interpersonal communication for the user, however, participants in this study echoed what others have noted as a potential limitation of DBCI—that they may be most useful for younger users with higher levels of educational attainment, and those who are already in the "action stage" of behavior change (Carroll et al., 2017; Kontos et al., 2014; Krebs & Duncan, 2015). The pediatric clinicians in this study thought *CDC's Milestone Tracker* was a good idea, but not necessarily for the families they work with, even going as far as pathologizing parents and caregivers with lower incomes and lower literacy levels who receive care at FQHCs. Some of the descriptions about perceived users of the app demonstrated biases about who "cares" about child development. Clinicians seem to be making assumptions about parent and caregiver's interest and motivation to attend to their children's development. Counter to clinician's biases and beliefs that parents and caregivers with lower incomes are not interested in child development, results from a study conducted to integrate LTSAE developmental monitoring checklists (i.e., the paper versions of the milestone checklists in *CDC's Milestone Tracker*) in WIC)—which serve a similar population as FQHCs—demonstrated that overwhelmingly, families were willing to complete checklists (Farmer et al., 2022). Only 20% of WIC staff reported lack of interest from families as a barrier to integration. Further, 95% of WIC staff reported that the checklists were easy for families to understand (Farmer et al., 2022), and the content in the app was written at or below a 7th grade reading level (Zubler et al., 2022), refuting another concern voiced by the pediatric clinicians in the current study. Nonetheless, this raises an important design question about how the app (and other behavior change apps like the *CDC's Milestone Tracker*) can be redesigned for users with lower incomes and lower literacy levels.

The emphasis on time as a potential benefit of the app but also as a barrier to use was particularly noticeable in clinician's discussion about perceptions regarding engagement with *CDC's Milestone Tracker*. Participants in this sample mentioned that the app could save time during well-child visits but also expressed concerns about it potentially requiring additional time (that they do not have) to introduce and explain it to parents and caregivers. This finding echoes other research that has identified lack of time as a major barrier to use. For example, Thies et al (2017) found that asking patients to download an app while in clinic was challenging and burdensome for clinic staff and patients alike, and did not integrate into existing workflows. Some also stated that due to competing priorities, they believed the families they work with would not have time to complete the milestone checklists within *CDC's Milestone Tracker*. Other studies have demonstrated that parents and caregivers from lower income communities do not complete DBCI when they struggle to balance their caregiver responsibilities (Brager et al., 2019). Relevantly, recent research demonstrates that within WIC settings, the paper-based milestone checklists are easy to use, easy to integrate into workflow, and took less than five minutes to complete (Farmer et al., 2022), which is promising for FQHCs. To integrate *CDC's Milestone Tracker* in FQHCs, it might be useful to consider principles from Implementation Science—a field dedicated to identifying methods to translate research into practice and promote successful uptake of evidence-based and empirically-supported interventions (Bauer et al., 2015; Koh et al., 2018). Proctor et al (2011) suggest a focus on eight specific outcomes including acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration, and sustainability. Examining each of these within the context of integrating *CDC's Milestone Tracker* in FQHCs might help to improve developmental surveillance and screening among young children from families experiencing social and economic disadvantages.

Future Directions

This study demonstrated that pediatric clinicians at FQHCs expect low engagement with *CDC's Milestone Tracker* among families but hypothesize that it would likely improve developmental surveillance and screening in clinic, and therefore it has utility and could be helpful. Yet much more research is needed to understand the underlying mechanisms influencing engagement with behavior change apps, particularly among secondary audiences such as clinicians, who often serve as gatekeepers and trusted sources of information for patients. It is important to explore the relationship between behaviors (e.g., perceptions about use and intended audiences), psychosocial processes (e.g., cognition and affect) and referrals or recommendations for behavior change apps. Additionally, research to assess the outcomes and impact of clinician referrals on user engagement with behavior change apps are necessary.

Implications for Dissertation Research

Findings from this study build on conclusions from the first study described in Chapter 2 (Aim 1) and the second study described in Chapter 3 (Aim 2). Connections to each article are described below in more detail.

Chapter 2: Assessing Engagement with Digital Behavior Change Interventions for Groups Experiencing Disadvantages: A Systematic Review

The study design and concepts investigated in the interview guide for the current study were informed by findings from the systematic review presented in Chapter 2. Additionally, the coding scheme used to analyze the interviews with pediatric clinicians was based on gaps identified in the existing literature on behavior change app engagement for groups experiencing disadvantages, which failed to produce multiple articles that conceptualized, operationalized, and assessed engagement as multidimensional. While the interaction between behavioral and psychosocial (e.g., cognitive and affective) processes was not the focus of this research, a strength is that all three constructs were embedded and assessed throughout the design and analysis of the study.

Chapter 3: Exploring Parent and Caregiver Engagement with CDC's Milestone Tracker App

Several concepts explored in this chapter were also examined via the parent and caregiver surveys reported in Chapter 3. Items to assess awareness of LTSAE were asked of pediatric clinicians as described in this chapter and of parents and caregivers on the surveys discussed in Chapter 3. Across the two studies, 51.1% (n = 47/92) were aware of LTSAE materials and resources. Initial reactions to the *CDC's Milestone Tracker* app were evaluated via the surveys with parents and caregivers and the interviews with pediatric clinicians. Videos, images, and colors were most frequently noted as features that caught the attention of participants across the two studies. The question about whether or not participants would recommend *CDC's Milestone Tracker* to parents and caregivers was asked on surveys with parents and caregivers and of pediatric clinicians during the interviews described in this chapter. Of the 92 participants across the two studies, 94.6% (*n* = 87/92) indicated they would recommend *CDC's Milestone Tracker* to other parents and caregivers, suggesting high overall satisfaction with the app among participants.

Some of the concerns expressed by clinicians this study differed from the findings from the surveys with parents and caregivers. For example, while clinicians thought parents and caregivers might be too busy to use *CDC's Milestone Tracker*, respondents in the study reported in Chapter 3 did not mention time as a perceived barrier to use. This divergence may be the result of the differences in populations, where the sample of parents and caregivers was more affluent and had high educational attainment as compared to the patients who receive care at FQHCs (a point revisited in Chapter 5), further underscoring that research about behavior change apps needs to be more inclusive of diverse populations.

Study Limitations

This study had some limitations. Given safety measures resulting from the COVID-19 pandemic, all interviews were conducted remotely. Participants did not have an opportunity to use *CDC's Milestone Tracker*, and therefore their responses are limited to what they were exposed to in the demonstration during the interview (as selected by the researcher) rather than self-exploration of the app. With the exception of the few clinicians who already use *CDC's Milestone Tracker* with patients, the results focused on clinicians perceptions as opposed to their experiences with the app. This study explored engagement from clinician's perspectives and is missing the perspectives of families who receive care at FQHCs. Future research should examine engagement with *CDC's Milestone Tracker* with a sample of parents and caregivers who receive care at FQHCs to compare findings. There is likely participation bias given that all 20 clinicians who expressed interest in the study completed an interview. Participants in this sample might be more likely to adhere to AAP guidelines for developmental surveillance and screening, and consequently were motivated to share their experiences, indicated by their enthusiasm to participate in an interview about how they perform developmental surveillance and screening. This sample was also mostly non-Latinx White and women. It is unclear what other themes would have emerged from a more diverse sample of clinicians. Consequently, more research with clinicians who do not self-identify as non-Latinx White (specifically Asian, Black, Latinx, and Native American or Alaskan Native clinicians) is warranted.

Conclusions

This chapter examined perceptions regarding engagement with *CDC's Milestone Tracker* from the perspective of pediatric clinicians at FQHCs (Aim 3). Findings indicate that while pediatric clinicians hold positive perceptions about the app, they believe adoption and engagement would be low among families who receive care at FQHCs. Despite expectations about low engagement and use, most clinicians in this study indicated they would recommend the app to parents and caregivers, and described potential utility and benefits of the app. More work is needed to address identified concerns and increase awareness about *CDC's Milestone Tracker* as a tool to support early identification of developmental delays and disabilities, particularly among families with lower incomes and those who care for them. Findings from this study are revisited and summarized in the context of this broader dissertation project in Chapter 5 of this dissertation.

Chapter 5

Discussion, Future Directions, and Conclusion

Dissertation Research Summary

In recent years, there has been a proliferation of health interventions disseminated using digital technologies. DBCI—with wide reach and tailored content specific to audiences—have potential to improve health outcomes. Responding to calls for more inclusive engagement research, I set out to answer the following guiding research question: How to better measure engagement with behavior change apps for groups experiencing disadvantages? Using a three article model approach and drawing on models of engagement that center engagement as a multidimensional phenomenon inclusive of behavioral, cognitive, and affective processes, this dissertation consisted of three studies: (1) a systematic review of the literature ($N = 21$ articles included) to identify existing conceptual definitions, measures, and methods to measure engagement with behavior change apps for groups experiencing disadvantages (Aim 1), (2) baseline and one-month follow-up surveys and app usage data to understand parent and caregiver ($N = 72$) engagement with the *CDC's Milestone Tracker* app (Aim 2), and (3) interviews to explore perceptions of engagement with *CDC's Milestone Tracker* app from the perspective of pediatric clinicians at FQHCs ($N = 20$; Aim 3).

In answering the first six research questions posed in Chapter 1, the three studies described in the aforementioned chapters attempted to fill a gap in our understanding of engagement specific to behavior change apps for groups experiencing disadvantages and inform the potential methods and measures that could be used to measures engagement with behavior change apps. The primary contributions of this work are theoretical (or conceptual). Theoretical contributions include an improved conceptual definition of engagement as (1) multidimensional, where the dimension of behavioral engagement encompasses both engagement with the technology and the health behavior, and (2) temporal, which ebbs and flows depending on the needs of the user and the health behaviors the app is intended to change. Further, theoretical contributions of this research illuminate how and why parents and caregivers from groups experiencing disadvantages engage with and use a behavior change app (the *CDC's Milestone Tracker*), which can help to inform app design to improve the quality of app engagement.

Secondary contributions of this work are empirical and methodological. Empirical contributions include a synthesis of measures and methods used to measure engagement with behavior change apps for groups experiencing disadvantages that emerged from the systematic review, which help us understand the different contexts to use these engagement methods and measures. Importantly, another empirical contribution is the fact that each of these three studies on engagement centered people from groups experiencing disadvantages. Additionally, the studies described in Chapters 3 and 4 add to the few evaluations of the *CDC's Milestone Tracker* app as the first research to evaluate engagement with the app. Finally, a methodological contribution of this research is the development of a new behavior change app engagement scale, which while not validated, is a step toward improved measures to evaluate engagement with behavior change apps.

Given the importance and pervasiveness of DBCI and behavior change apps, it is essential to evaluate their effectiveness to understand if they create change and improve health and well-being. The alarming disparities that exist and overall poor rates of early identification of developmental delays and disabilities in young children affirm the necessity of tools like *CDC's Milestone Tracker* app that support developmental surveillance using BCTs with known MoAs. Yet limited evaluations of *CDC's Milestone Tracker* have been conducted to date, and it is unknown how and why families from groups experiencing disadvantages engage with the app. As such, a secondary goal of this study was to evaluate the *CDC's Milestone Tracker* app. Parents and caregivers and pediatric clinicians all responded favorably to *CDC's Milestone Tracker* app, with 95% reporting that they would recommend the app to others.

The final research question posed in Chapter 1 was what methods or measures should be used to evaluate engagement with the *CDC's Milestone Tracker* app? To answer this question, I used a working definition of engagement as multidimensional (discussed further in detail below) to create a data integration matrix (organized with the dimensions of engagement along the y-axis and each of the three dissertation studies along the x-axis) that was populated with results from this dissertation. For quantitative findings, a threshold of at least 25% of articles, respondents, or participants and survey items with Likert scale response options that had an average rating of 3.5 or higher were used for inclusion in the matrix. Key qualitative findings were included as they were inherently the most prevalent themes identified in the data. Based on the data integration matrix, I selected measures that could be used to evaluate engagement with *CDC's Milestone Tracker* (Table 14). Using a combination of system usage data (objective measurement) and an in-app survey (subjective measurement), behavioral engagement with *CDC's Milestone Tracker* can be measured by assessing (1) how users use the app (engagement with the technology) and (2) what the outcomes of using the app are (engagement with the health behavior). Cognitive engagement with the app can be measured by assessing attention

Table 14

Proposed Measures and Methods to Evaluate Engagement with CDC's Milestone Tracker App

Engagement Dimension	Measures (Example Survey Items)	Method
Behavioral engagement (interaction with the technology and the health behavior)	• Child profiles created • Type and number of age-appropriate milestone checklists completed • Milestone checklists return rate (return to complete subsequent milestone checklists calculated as the proportion of checklists completed divided by checklists available) • Checklists updated if selected "No" or "Not yet" for a child who is initially missing a milestone • Use/interact with other non-checklist app features (e.g., tips and activities) • Notification response rate (calculated as the proportion of notifications opened divided by notifications sent)	System usage data [a]
	• Perceived ease of use (e.g., "The app is easy to use.") • Behavioral antecedents (e.g., "I plan to use this app to track my child's development in the future." and "I would use this app to share concerns about my child's development with my doctor or healthcare provider.") • Behavioral outcomes: Act early if concerns about a child's development are identified (e.g., "Who did you share your child's milestone summary with?", "Did CDC's Milestone Tracker app help you identify any possible concerns related to your child's development?" and "Who did you share these possible concerns with?")	In-app survey [b]
Cognitive engagement	• Use/interact with other non-checklist app features (e.g., number of videos within milestone checklists watched)	System usage data [a]
	• Attention (e.g., "What is the most interesting part of the app?") • Trust in the app (e.g., "I trust this app to help me track my child's development." and "I trust this app to provide accurate information about what milestones my child should be reaching for their age")	In-app survey [b]
Affective engagement	• Satisfaction and enjoyment (e.g., "I like using this app to track my child's development." and "What features of the app do you like most?") • Relatability and identity (e.g., "This app was designed for someone like me.") • Recommendation (e.g., "Would you recommend this app to friends and family?")	In-app survey [b]

Note. Survey items are examples and not necessarily comprehensive.
[a] Measures objective engagement. [b] Measures subjective engagement.

and trust in the app to help accomplish the goal of learning about and tracking child development. Affective engagement can be measured by evaluating satisfaction and enjoyment, relatability/identity, and recommendation. If system usage data were to become available at the individual user level, then adding elements of the objective measures may help to strengthen the behavior change app engagement scale (described and developed in Chapter 3) as a tool to measure engagement with *CDC's Milestone Tracker* app. Developers and evaluators of apps with similar app features and BCTs that allow for periods of app engagement, disengagement, and re-engagement, could adapt these measures to evaluate engagement with other behavior change apps.

Theoretical Implications

Based on the systematic review (Chapter 2), surveys with parents and caregivers (Chapter 3), and interviews with pediatric clinicians (Chapter 4), the conceptual definition of engagement with behavior change apps that emerged from this work is as follows: engagement is a multidimensional process—consisting of behavioral (e.g., health behaviors and app use and interaction with app features), cognitive (e.g., attention, trust), and affective (e.g., satisfaction, enjoyment, and relatability/identity) factors—which fluctuates based on the needs of the user and the health behavior(s) the app is designed to change. This conceptual definition of engagement expands the conceptualizations offered by Kelders et al (2020), Perski et al (2017), and others by adopting Cole-Lewis et al's (2019) theory that behavioral engagement is multifaceted, comprised of both engagement with the behavior change components of the app and engagement with the features designed to encourage use.

In conceptualizing engagement as permutable, this conceptual definition recognizes the different stages or phases of engagement (O'Brien & Toms, 2008) that users move between when using behavior change apps, which are dependent on their needs and the purpose of the app (a finding that emerged from the surveys with parents and caregivers). Using *CDC's Milestone Tracker* app as an example, engagement may be ephemeral, with spurts of high levels of engagement during infancy (where there are more milestone checklists available) and when a well-child visit approaches. As long as users download the app (point of engagement or initial engagement), create a child profile and complete the age-appropriate milestone checklists (stay engaged), return to the app to complete future checklists when available (re-engagement), know what to do if they have concerns, are empowered to act early if concerns are identified, continuous app engagement is not necessary for the app to be effective. Periods of disengagement can be expected, whereas apps for tracking sleep or exercise may require a user to stay engaged to produce desired outcomes.

As such, engagement evaluation of behavior change apps must take all three dimensions into account. This requires measuring both objective and subjective engagement, as suggested by Graham et al (2021) and others. This can be accomplished by using methods such as system usage data to capture objective measures of behavioral engagement combined with surveys or questionnaires to capture subjective measures of the behavioral and psychosocial aspects of engagement. The frequency with which to collect these data was not a finding that emerged from this study but should be studied further to identify ideal timing to gather data to assess engagement. Additionally, while there may be a potential interaction between the three dimensions of engagement, examining this interaction was outside of the scope of this dissertation project but should be explored in future studies.

Moreover, this dissertation centered populations often excluded from app development and evaluation research. In accordance with the ConNECT framework (Alcaraz et al., 2017)—

which provides valuable guiding principles to help researchers ensure health equity is at the core of intervention design and evaluation and not an afterthought—the studies described in this dissertation approached engagement research from the perspective of groups experiencing disadvantages (and those who work with them). It is possible that different populations exhibit different patterns of engagement (Kelders, Kip, et al., 2020), but a finding from this dissertation is that engagement must be conceptualized as both multidimensional and temporal. The weight of the three dimensions may differ by population, but additional research is needed to explore engagement dimensions and the individual constructs that constitute each dimension among different populations. The culmination of this research helps to consolidate our understanding of engagement and guide future intervention and evaluation design to develop behavior change apps that are engaging for all.

Societal and Practical Implications

Findings from this applied research project have immediate real-world impact. Findings may have implications for not only improving existing interventions, but also for informing future efforts by offering an analysis of factors that influence engagement from the perspectives of populations often excluded from the intervention development and evaluation process. The conceptual definition used in this research can be modified and used by others, which could help to normalize the inclusion of conceptual definitions when reporting on engagement with behavior change apps. Similarly, measures used to evaluate engagement must be reported to improve replicability and comparability across studies. This research did not attempt to create a single measurement framework to be applied to all types of DBCI platforms or behavior change apps for groups experiencing disadvantages. Given that users engage with various forms of technology in different ways, that platforms produce channel-specific metrics, and require different levels of engagement to facilitate behavior change, it is unrealistic to develop a set of measures that can be used across platforms (Saunders, 2015). Nonetheless, understanding the behavioral, cognitive, and affective factors influencing engagement for one platform may provide insight to how users will engage with other behavior change apps. Conclusions drawn from the evaluation studies described in Chapters 3 and 4 will also be presented to the LTSAE team at CDC to inform the development of an in-app survey to evaluate *CDC's Milestone Tracker* app and other updates to the app. Ultimately, those designing behavior change apps for groups experiencing disadvantages should be able to use my work to inform evaluation practices to improve interventions and positively impact health outcomes.

Limitations

While individual study limitations are discussed within each chapter, there were some collective limitations of this dissertation. The intention was to include a sample of parents and caregivers from groups experiencing disadvantages that was at socioeconomically diverse and least somewhat comparable to the patients served by the pediatric clinicians at FQHCs. Instead, the parents and caregivers who participated in the surveys held high educational attainment, higher incomes, and had private health insurance whereas the pediatric clinicians primarily serve families who are insured by Medicaid. Although falling under the umbrella category of groups experiencing disadvantages, it is possible that these two groups of parents and caregivers use and engage with *CDC's Milestone Tracker* app in different ways, which limits the ability to draw comparisons between clinicians perceptions and parents and caregiver's perceptions and experiences.

The inability to evaluate individual user level app usage data was a major setback. Given that 81% of articles included in the systematic review used objective data to measure engagement, it would have been preferable to explore engagement using a combination of objective and subjective data. Further, the LTSAE team noted several limitations to the quality and reliability of the *CDC's Milestone Tracker* app usage data, including missing data resulting from app updates and issues related to app functionality. In February 2022, the CDC released updated developmental milestones to assign milestones most (at least 75%) of children would reach in an attempt to reduce the "wait and see" approach to taking action on missed milestones (Zubler et al., 2022). Accordingly, *CDC's Milestone Tracker* app was also updated to include new milestone checklists for 15 and 30 months, remove some milestones, recategorize others, and revise other app features (including tips and activities and adding open-ended questions to help facilitate conversations between parents and caregivers, clinicians, and others). While the surveys described in Chapter 3 were implemented prior to this update, some of the findings may be moot after the app updates.

Future Research Directions

This initial exploration of engagement with behavior change apps for groups experiencing disadvantages should be viewed as a starting point. Much more research is needed to fully understand how to evaluate engagement with behavior change apps. While scale validation is outside the scope of this dissertation study, future research should evaluate the internal consistency of the engagement measurement framework developed as part of this dissertation. Additionally, similar to Perski et al (2019) and Kelders, Kip, et al (2020), it will be necessary to evaluate the content, construct, criterion and predictive validity of the engagement framework with *CDC's Milestone Tracker* app users. It would also be beneficial to validate this framework with other priority audiences of the app and to test the engagement framework with non-priority audiences. Finally, this dissertation was never intended to explore the relationship between engagement and behavioral outcomes, but this would be an important next step to understand how engagement with behavior change apps helps to facilitate behavior change, improve health outcomes, and reduce health disparities.

References

Abercrombie, J., Pann, J., Shin, F., Taylor, E., Brisendine, A. E., Swanson-Holm, R., James, C., Viehweg, S., & Chödrön, G. (2021). Evaluation of the feasibility and perceived value of integrating Learn the Signs. Act Early. developmental monitoring resources in Early Head Start. *Early Childhood Education Journal*. https://doi.org/10.1007/s10643-021-01247-5

Adams, R. C., & Tapia, C. (2013). Early intervention, IDEA Part C services, and the medical home: Collaboration for best practice and best outcomes. *Pediatrics, 132*(4), e1073-e1088. https://doi.org/10.1542/peds.2013-2305

Adler, N. E., & Newman, K. (2002). Socioeconomic disparities in health: Pathways and policies. *Health Affairs, 21*(2), 60-76. https://doi.org/10.1377/hlthaff.21.2.60

Aharonovich, E., Stohl, M., Cannizzaro, D., & Hasin, D. (2017). Healthcall delivered via smartphone to reduce co-occurring drug and alcohol use in HIV-infected adults: A randomized pilot trial. *Journal of Substance Abuse Treatment, 83*, 15-26. https://doi.org/10.1016/j.jsat.2017.09.013

Aitken, M., Clancy, B., & Nass, D. (2017). *The growing value of digital health: Evidence and impact on human health and the healthcare system*. IQVIA Institute for Human Data Science. https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health

Aitken, M., & Nass, D. (2021). *Digital health trends 2021: Innovation, evidence, regulation, and adoption*. IQVIA Institute for Human Data Science. https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/digital-health-trends-2021/iqvia-institute-digital-health-trends-2021.pdf

Alcaraz, K. I., Sly, J., Ashing, K., Fleisher, L., Gil-Rivas, V., Ford, S., Yi, J. C., Lu, Q., Meade, C. D., Menon, U., & Gwede, C. K. (2017). The connect framework: A model for advancing behavioral medicine science and practice to foster health equity. *Journal of Behavioral Medicine, 40*(1), 23-38. https://doi.org/10.1007/s10865-016-9780-4

Alkhaldi, G., Hamilton, F. L., Lau, R., Webster, R., Michie, S., & Murray, E. (2016). The effectiveness of prompts to promote engagement with digital interventions: A systematic review. *Journal of Medical Internet Research, 18*(1), e6. https://doi.org/10.2196/jmir.4790

AllianceChicago. (2022). *Alliancechicago: Who we are*. Retrieved February 12, 2022 from https://alliancechicago.org/

American Academy of Pediatrics. (2001a). Developmental surveillance and screening of infants and young children. *Pediatrics, 108*(1), 192-195. https://doi.org/10.1542/peds.108.1.192

American Academy of Pediatrics. (2001b). Role of the pediatrician in family-centered early intervention services. *Pediatrics, 107*(5), 1155-1157. https://doi.org/10.1542/peds.107.5.1155

American Academy of Pediatrics. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics, 118*(1), 405-420. https://doi.org/10.1542/peds.2006-1231

American Psychological Association. (2021). *Inclusive language guidelines*. Retrieved February 4, 2022 from https://www.apa.org/about/apa/equity-diversity-inclusion/language-guidelines.pdf

Anand, S. S., Samaan, Z., Middleton, C., Irvine, J., Desai, D., Schulze, K. M., Sothiratnam, S., Hussain, F., Shah, B. R., Pare, G., Beyene, J., & Lear, S. A. (2016). A digital health intervention to lower cardiovascular risk: A randomized clinical trial. *JAMA Cardiology, 1*(5), 601-606. https://doi.org/10.1001/jamacardio.2016.1035

Armenta, V., Warrell, L., Nazneen, N., Escobedo, L., & Arriaga, R. I. (2019). Actearly: A bi-national evaluation study of a mobile application for tracking developmental milestones. In *Proceedings of the IX Latin American Conference on Human Computer Interaction* (pp. 1-8). Panama City, Panama. https://doi.org/10.1145/3358961.3358963

Arora, V. M., Madison, S., & Simpson, L. (2020). Addressing medical misinformation in the patient-clinician relationship. *JAMA, 324*(23), 2367-2368. https://doi.org/10.1001/jama.2020.4263

Arunyanart, W., Fenick, A., Ukritchon, S., Imjaijitt, W., Northrup, V., & Weitzman, C. (2012). Developmental and autism screening: A survey across six states. *Infants & Young Children, 25*(3), 175-187. https://doi.org/10.1097/IYC.0b013e31825a5a42

Atkin, C. K., & Rice, R. E. (2013). Theory and principles of public communication campaigns. In R. E. Rice & C. K. Atkin (Eds.), *Public communication campaigns* (pp. 3-19). Los Angeles, CA: Sage Publications.

Atkins, K. L., Dolata, J. K., Blasco, P. M., Saxton, S. N., & Duvall, S. W. (2020). Early intervention referral outcomes for children at increased risk of experiencing developmental delays. *Maternal and Child Health Journal, 24*(2), 204-212. https://doi.org/10.1007/s10995-019-02830-4

Badal, H. J., Stryker, J. E., DeLuca, N., & Purcell, D. W. (2018). Swipe right: Dating website and app use among men who have sex with men. *AIDS and Behavior, 22*(4), 1265-1272. https://doi.org/10.1007/s10461-017-1882-7

Ballard, A. M., Cardwell, T., & Young, A. M. (2019). Fraud detection protocol for web-based research among men who have sex with men: Development and descriptive evaluation. *JMIR Public Health and Surveillance, 5*(1), e12344. https://doi.org/10.2196/12344

Baltierra, N. B., Muessig, K. E., Pike, E. C., LeGrand, S., Bull, S. S., & Hightow-Weidman, L. B. (2016). More than just tracking time: Complex measures of user engagement with an internet-based health promotion intervention. *Journal of Biomedical Informatics, 59*, 299-307. https://doi.org/10.1016/j.jbi.2015.12.015

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology, 4*(3), 359-373.

Barger, B., Benevides, T., Rizk, S., Rice, C., Heiman, H., Salmon, A., & Sanchez-Alvarez, S. (2022). Race/ethnic inequities in conjoint monitoring and screening for U.S. children 3 and under. *Disability and Health Journal, 15*(1), 101179. https://doi.org/https://doi.org/10.1016/j.dhjo.2021.101179

Barger, B., Rice, C., Wolf, R., & Roach, A. (2018). Better together: Developmental screening and monitoring best identify children who need early intervention. *Disability and Health Journal, 11*(3), 420-426. https://doi.org/10.1016/j.dhjo.2018.01.002

Barman-Adhikari, A., Rice, E., Bender, K., Lengnick-Hall, R., Yoshioka-Maxwell, A., & Rhoades, H. (2016). Social networking technology use and engagement in HIV-related risk and protective behaviors among homeless youth. *Journal of Health Communication, 21*(7), 809-817. https://doi.org/10.1080/10810730.2016.1177139

Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., & Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. *BMC Psychology, 3*(1), 32. https://doi.org/10.1186/s40359-015-0089-9

Bauermeister, J. A., Golinkoff, J. M., Muessig, K. E., Horvath, K. J., & Hightow-Weidman, L. B. (2017). Addressing engagement in technology-based behavioural HIV interventions through paradata metrics. *Current Opinion in HIV and AIDS, 12*(5), 442-446. https://doi.org/10.1097/COH.0000000000000396

Baumel, A., Edan, S., & Kane, J. M. (2019). Is there a trial bias impacting user engagement with unguided e-mental health interventions? A systematic comparison of published reports and real-world usage of the same programs. *Translational Behavioral Medicine, 9*(6), 1020-1033. https://doi.org/10.1093/tbm/ibz147

Beatty, A. L., Fukuoka, Y., & Whooley, M. A. (2013). Using mobile technology for cardiac rehabilitation: A review and framework for development and evaluation. *Journal of the American Heart Association, 2*(6), e000568. https://doi.org/10.1161/JAHA.113.000568

Bell, L., Garnett, C., Qian, T., Perski, O., Williamson, E., & Potts, H. W. (2020). Engagement with a behavior change app for alcohol reduction: Data visualization for longitudinal observational study. *Journal of Medical Internet Research, 22*(12), e23369. https://doi.org/10.2196/23369

Ben-Zeev, D., Brian, R. M., Jonathan, G., Razzano, L., Pashka, N., Carpenter-Song, E., Drake, R. E., & Scherer, E. A. (2018). Mobile health (mHealth) versus clinic-based group intervention for people with serious mental illness: A randomized controlled trial. *Psychiatric Services, 69*(9), 978-985. https://doi.org/10.1176/appi.ps.201800063

Bennett, G. G., Steinberg, D., Askew, S., Levine, E., Foley, P., Batch, B. C., Svetkey, L. P., Bosworth, H. B., Puleo, E. M., Brewer, A., DeVries, A., & Miranda, H. (2018). Effectiveness of an app and provider counseling for obesity treatment in primary care. *American Journal of Preventive Medicine, 55*(6), 777-786. https://doi.org/10.1016/j.amepre.2018.07.005

Bennett, G. G., Steinberg, D. M., Stoute, C., Lanpher, M., Lane, I., Askew, S., Foley, P. B., & Baskin, M. L. (2014). Electronic health (eHealth) interventions for weight management among racial/ethnic minority adults: A systematic review. *Obesity Reviews, 15 Suppl 4*, 146-158. https://doi.org/10.1111/obr.12218

Bernhardt, J. M., Mays, D., & Kreuter, M. W. (2011). Dissemination 2.0: Closing the gap between knowledge and practice with new media and marketing. *Journal of Health Communication, 16*(sup1), 32-44. https://doi.org/10.1080/10810730.2011.593608

Berry, A. D., Garzon, D. L., Mack, P., Kanwischer, K. Z., & Beck, D. G. (2014). Implementing an early childhood developmental screening and surveillance program in primary care settings: Lessons learned from a project in Illinois. *Journal of Pediatric Health Care, 28*(6), 516-525. https://doi.org/10.1016/j.pedhc.2014.04.008

Bilaver, L. A., Sobotka, S. A., & Mandell, D. S. (2020). Understanding racial and ethnic disparities in autism-related service use among Medicaid-enrolled children. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-020-04797-6

Biviji, R., Vest, J. R., Dixon, B. E., Cullen, T., & Harle, C. A. (2020). Factors related to user ratings and user downloads of mobile apps for maternal and infant health: A cross-sectional study. *JMIR Mhealth Uhealth, 8*(1), e15663. https://doi.org/10.2196/15663

Blok, A. C., Sadasivam, R. S., Amante, D. J., Kamberi, A., Flahive, J., Morley, J., Conigliaro, J., & Houston, T. K. (2019). Gamification to motivate the unmotivated smoker: The "Take a Break" digital health intervention. *Games for Health Journal, 8*(4), 275-284. https://doi.org/10.1089/g4h.2018.0076

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*. https://doi.org/10.3389/fpubh.2018.00149

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., Visser, S., & Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics, 127*(6), 1034-1042. https://doi.org/10.1542/peds.2010-2989

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*(1), 371-399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Bradshaw, J., Steiner, A. M., Gengoux, G., & Koegel, L. K. (2015). Feasibility and effectiveness of very early intervention for infants at-risk for autism spectrum disorder: A systematic review. *Journal of Autism and Developmental Disorders, 45*(3), 778-794. https://doi.org/10.1007/s10803-014-2235-2

Brager, J., Breitenstein, S. M., Miller, H., & Gross, D. (2019). Low-income parents' perceptions of and engagement with a digital behavioral parent training program: A mixed-methods study. *Journal of the American Psychiatric Nurses Association*. https://doi.org/10.1177/1078390319872534

Brewer, L. C., Fortuna, K. L., Jones, C., Walker, R., Hayes, S. N., Patten, C. A., & Cooper, L. A. (2020). Back to the future: Achieving health equity through health informatics and digital health. *JMIR Mhealth Uhealth, 8*(1), e14512. https://doi.org/10.2196/14512

Bright, M. A., Zubler, J., Boothby, C., & Whitaker, T. M. (2019). Improving developmental screening, discussion, and referral in pediatric practice. *Clinical Pediatrics, 58*(9), 941-948. https://doi.org/10.1177/0009922819841017

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). London: Taylor & Francis.

Bruder, M. B. (2010). Early childhood intervention: A promise to children and families for their future. *Exceptional Children, 76*(3), 339-355.

Businelle, M. S., Ma, P., Kendzor, D. E., Frank, S. G., Vidrine, D. J., & Wetter, D. W. (2016). An ecological momentary intervention for smoking cessation: Evaluation of feasibility and effectiveness. *Journal of Medical Internet Research, 18*(12), e321-e321. https://doi.org/10.2196/jmir.6058

Businelle, M. S., Ma, P., Kendzor, D. E., Frank, S. G., Wetter, D. W., & Vidrine, D. J. (2016). Using intensive longitudinal data collected via mobile phone to detect imminent lapse in smokers undergoing a scheduled quit attempt. *Journal of Medical Internet Research, 18*(10), e275. https://doi.org/10.2196/jmir.6307

Campbell, J. M., Greenberg, D., Gallagher, P. A., Stoneman, Z., & Simmons, C. (2019). Knowledge of autism for parents of low income with low literacy: Description and relationship to child development knowledge. *Advances in Neurodevelopmental Disorders, 3*(1), 8-16. https://doi.org/10.1007/s41252-018-0079-z

Cao, B., Gupta, S., Wang, J., Hightow-Weidman, L. B., Muessig, K. E., Tang, W., Pan, S., Pendse, R., & Tucker, J. D. (2017). Social media interventions to promote HIV testing, linkage, adherence, and retention: Systematic review and meta-analysis. *Journal of Medical Internet Research, 19*(11), e394. https://doi.org/10.2196/jmir.7997

Carey, R. N., Connell, L. E., Johnston, M., Rothman, A. J., de Bruin, M., Kelly, M. P., & Michie, S. (2018). Behavior change techniques and their mechanisms of action: A synthesis of links described in published intervention literature. *Annals of Behavioral Medicine, 53*(8), 693-707. https://doi.org/10.1093/abm/kay078

Carlson, M., Donahue, S., & Foster, S. (2011). *Towards an understanding of health equity: Glossary*. Alberta Health Services. https://www.albertahealthservices.ca/poph/hi-poph-surv-shsa-tpgwg-glossary.pdf

Carroll, J. K., Moorhead, A., Bond, R., LeBlanc, W. G., Petrella, R. J., & Fiscella, K. (2017). Who uses mobile phone health apps and does use matter? A secondary data analytics approach. *Journal of Medical Internet Research, 19*(4), e125. https://doi.org/10.2196/jmir.5604

Centers for Disease Control and Prevention. (2021a). *About the "Learn the Signs. Act Early." Program*. Retrieved November 22 from https://www.cdc.gov/ncbddd/actearly/about.html

Centers for Disease Control and Prevention. (2021b). *Health equity guiding principles for inclusive communication*. Retrieved November 14, 2020 from https://www.cdc.gov/healthcommunication/Health_Equity.html

Chaudhry, B. M., Schaefbauer, C., Jelen, B., Siek, K. A., & Connelly, K. (2016). Evaluation of a food portion size estimation interface for a varying literacy population. In *Proceedings of* *the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5645-5657). San Jose, CA, USA. https://doi.org/10.1145/2858036.2858554

Cho, H., Powell, D., Pichon, A., Kuhns, L. M., Garofalo, R., & Schnall, R. (2019). Eye-tracking retrospective think-aloud as a novel approach for a usability evaluation. *International Journal of Medical Informatics, 129*, 366-373. https://doi.org/10.1016/j.ijmedinf.2019.07.010

Chödrön, G., Barger, B., Pizur-Barnekow, K., Viehweg, S., & Puk-Ament, A. (2021). "Watch Me!" training increases knowledge and impacts attitudes related to developmental monitoring and referral among childcare providers. *Maternal and Child Health Journal, 25*(6), 980-990. https://doi.org/10.1007/s10995-020-03097-w

Chödrön, G., Pizur-Barnekow, K., Viehweg, S., Puk-Ament, A., & Barger, B. (2021). Childcare providers' attitudes, knowledge, and practice related to developmental monitoring to promote early identification and referral. *Early Child Development and Care, 191*(4), 520-534. https://doi.org/10.1080/03004430.2019.1626373

Coa, K. I., Wiseman, K. P., Higgins, B., & Augustson, E. (2019). Associations between engagement and outcomes in the smokefreetxt program: A growth mixture modeling analysis. *Nicotine & Tobacco Research, 21*(5), 663–669. https://doi.org/10.1093/ntr/nty073

Cole-Lewis, H., Ezeanochie, N., & Turgiss, J. (2019). Understanding health behavior technology engagement: Pathway to measuring digital behavior change interventions. *JMIR Formative Research, 3*(4), e14052. https://doi.org/10.2196/14052

Couper, M. P., Alexander, G. L., Zhang, N., Little, R. J. A., Maddy, N., Nowak, M. A., McClure, J. B., Calvi, J. J., Rolnick, S. J., Stopponi, M. A., & Cole Johnson, C. (2010).

Engagement and retention: Measuring breadth and depth of participant use of an online intervention. *Journal of Medical Internet Research, 12*(4), e52. https://doi.org/10.2196/jmir.1430

Crutzen, R., de Nooijer, J., Brouwer, W., Oenema, A., Brug, J., & de Vries, N. K. (2010). Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: A systematic review. *Health Education & Behavior, 38*(1), 49-62. https://doi.org/10.1177/1090198110372878

Cugelman, B., Thelwall, M., & Dawes, P. (2011). Online interventions for social marketing health behavior change campaigns: A meta-analysis of psychological architectures and adherence factors. *Journal of Medical Internet Research, 13*(1), e17. https://doi.org/10.2196/jmir.1367

Daniel, K. L., Prue, C., Taylor, M. K., Thomas, J., & Scales, M. (2009). 'Learn the signs. Act early': A campaign to help every child reach his or her full potential. *Public Health, 123*, e11-e16. https://doi.org/10.1016/j.puhe.2009.06.002

Daniels, A. M., & Mandell, D. S. (2013). Explaining differences in age at autism spectrum disorder diagnosis: A critical review. *Autism, 18*(5), 583-597. https://doi.org/10.1177/1362361313480277

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. https://doi.org/10.2307/249008

Delaney, T., Mclaughlin, M., Yoong, S. L., Wyse, R., Sutherland, R., Hollis, J., Brown, A., Barnes, C., Dray, J., Wiggers, J., & Wolfenden, L. (2018). *Associations between digital health intervention engagement and dietary intake: A systematic review*

(CRD42018112189)[Research Protocol]. PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018112189

Devlin, A. M., McGee-Lennon, M., O'Donnell, C. A., Bouamrane, M.-M., Agbakoba, R., O'Connor, S., Grieve, E., Finch, T., Wyke, S., Watson, N., Browne, S., Mair, F. S., & the "Dallas" Evaluation Team. (2015). Delivering digital health and well-being at scale: Lessons learned during the implementation of the dallas program in the United Kingdom. *Journal of the American Medical Informatics Association, 23*(1), 48-59. https://doi.org/10.1093/jamia/ocv097

Dillingham, R., Ingersoll, K., Flickinger, T. E., Waldman, A. L., Grabowski, M., Laurence, C., Wispelwey, E., Reynolds, G., Conaway, M., & Cohn, W. F. (2018). Positivelinks: A mobile health intervention for retention in HIV care and clinical outcomes with 12-month follow-up. *AIDS Patient Care and STDs, 32*(6), 241-250. https://doi.org/10.1089/apc.2017.0303

DiMaggio, P., Hargittai, E., Neuman, W. R., & Robinson, J. P. (2001). Social implications of the internet. *Annual Review of Sociology, 27*(1), 307-336. https://doi.org/10.1146/annurev.soc.27.1.307

Doherty, K., & Doherty, G. (2018). Engagement in HCI: Conception, theory and measurement. *ACM Computing Surveys, 51*(5), Article 99. https://doi.org/10.1145/3234149

Donnelly, G. (2022, February 4). The rise of digital health in five charts. *Emerging Tech Brew*. https://www.morningbrew.com/emerging-tech/stories/2022/02/04/the-rise-of-digital-health-in-five-charts

Donohue, M. R., Childs, A. W., Richards, M., & Robins, D. L. (2017). Race influences parent report of concerns about symptoms of autism spectrum disorder. *Autism, 23*(1), 100-111. https://doi.org/10.1177/1362361317722030

Dunst, C. J. (2007). Early intervention for infants and toddlers with developmental disabilities. In S. L. Odom, R. H. Horner, M. E. Snell, & J. Blacher (Eds.), *Handbook of developmental disabilities* (pp. 161-180). New York, NY: Guilford Press.

Durkin, M. S., Maenner, M. J., Baio, J., Christensen, D., Daniels, J., Fitzgerald, R., Imm, P., Lee, L. C., Schieve, L. A., Braun, K. V. N., Wingate, M. S., & Yeargin-Allsopp, M. (2017). Autism spectrum disorder among US children (2002–2010): Socioeconomic, racial, and ethnic disparities. *American Journal of Public Health, 107*(11), 1818-1826. https://doi.org/10.2105/ajph.2017.304032

Eremita, M., Semancik, E., Lerer, T., & Dworkin, P. H. (2017). Can we identify parents who do not verbally share concerns for their children's development? *Journal of Developmental and Behavioral Pediatrics, 38*(3), 224-227. https://doi.org/10.1097/DBP.0000000000000426

Escriva Boulley, G., Leroy, T., Bernetière, C., Paquienseguy, F., Desfriches-Doria, O., & Préau, M. (2018). Digital health interventions to help living with cancer: A systematic review of participants' engagement and psychosocial effects. *Psycho-Oncology, 27*(12), 2677-2686. https://doi.org/10.1002/pon.4867

Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research, 7*(1), e11. https://doi.org/10.2196/jmir.7.1.e11

Farmer, J. E., Falk, L. W., Clark, M. J., Mayfield, W. A., & Green, K. K. (2022). Developmental monitoring and referral for low-income children served by WIC: Program development and implementation outcomes. *Maternal and Child Health Journal, 26*, 230-241. https://doi.org/10.1007/s10995-021-03319-9

Fefferman, M. L., Andrews, B. L., & Msall, M. E. (2017). Disparities in access to early intervention services for extremely preterm infants by family income. *International Public Health Journal, 9*(2), 155-167.

Feinberg, E., Silverstein, M., Donahue, S., & Bliss, R. (2011). The impact of race on participation in Part C early intervention services. *Journal of Developmental and Behavioral Pediatrics, 32*(4), 284. https://doi.org/10.1097/DBP.0b013e3182142fbd

Ferguson, C. J. (2016). An effect size primer: A guide for clinicians and researchers. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 301–310). American Psychological Association. https://doi.org/10.1037/14805-020

Firth, J., Torous, J., Nicholas, J., Carney, R., Pratap, A., Rosenbaum, S., & Sarris, J. (2017). The efficacy of smartphone-based mental health interventions for depressive symptoms: A meta-analysis of randomized controlled trials. *World Psychiatry, 16*(3), 287-298. https://doi.org/10.1002/wps.20472

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health, 4*(2), e19. https://doi.org/10.2196/mental.7785

Flaherty, S. J., McCarthy, M., Collins, A. M., McCafferty, C., & McAuliffe, F. M. (2021). Exploring engagement with health apps: The emerging importance of situational involvement and individual characteristics. *European Journal of Marketing, 55*(3), 122-147. https://doi.org/10.1108/EJM-06-2019-0531

Flanagin, A., Frey, T., Christiansen, S. L., & AMA Manual of Style Committee. (2021). Updated guidance on the reporting of race and ethnicity in medical and science journals. *JAMA, 326*(7), 621-627. https://doi.org/10.1001/jama.2021.13304

Fountain, C., King, M. D., & Bearman, P. S. (2011). Age of diagnosis for autism: Individual and community factors across 10 birth cohorts. *Journal of Epidemiology and Community Health, 65*(6), 503-510. https://doi.org/10.1136/jech.2009.104588

Fowler Jr, F. (2013). *Survey research methods*. Thousand Oaks, CA: Sage Publications.

Free, C., Phillips, G., Galli, L., Watson, L., Felix, L., Edwards, P., Patel, V., & Haines, A. (2013). The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: A systematic review. *PLoS Medicine, 10*(1), e1001362. https://doi.org/10.1371/journal.pmed.1001362

Freelon, D. G. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science, 5*(1), 20-33.

Gadomski, A. M., Riley, M. R., Scribani, M., & Tallman, N. (2018). Impact of "Learn the Signs. Act Early." materials on parental engagement and doctor interaction regarding child development. *Journal of Developmental and Behavioral Pediatrics, 39*(9), 693-700. https://doi.org/10.1097/dbp.0000000000000604

Gallagher, P. A., Greenberg, D., Campbell, J. M., Stoneman, Z., & Feinberg, I. Z. (2019). Early identification and connection to services among urban parents who have low income and low-literacy skills. *Focus on Autism and Other Developmental Disabilities, 34*(2), 97-106. https://doi.org/10.1177/1088357618794913

Gerndt, K., & Mitchell, N. (Forthcoming 2022). *Understanding developmental surveillance in pediatric practice: Key informant interviews – final report*. American Academy of Pediatrics.

Getachew-Smith, H. B., Brockner, A. E., & Adebayo, N. A. (2020). *Engagement with digital behavior change intervention apps for socioeconomically disadvantaged populations: Systematic review protocol* [Study Design]. DigitalHub. Galter Health Sciences Library & Learning Center. https://doi.org/10.18131/g3-0506-e214

Gibbons, C. M. (2011). Use of health information technology among racial and ethnic underserved communities. *Perspectives in Health Information Management, 8*(Winter), 1f. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3035830/pdf/phim0008-0001f.pdf

Gibbons, C. M., Fleisher, L., Slamon, R. E., Bass, S., Kandadai, V., & Beck, J. R. (2011). Exploring the potential of web 2.0 to address health disparities. *Journal of Health Communication, 16*(sup1), 77-89. https://doi.org/10.1080/10810730.2011.596916

Glasgow, R. E., Christiansen, S. M., Kurz, D., King, D. K., Woolley, T., Faber, A. J., Estabrooks, P. A., Strycker, L., Toobert, D., & Dickman, J. (2011). Engagement in a diabetes self-management website: Usage patterns and generalizability of program use. *Journal of Medical Internet Research, 13*(1), e9. https://doi.org/10.2196/jmir.1391

Glasgow, R. E., Nelson, C. C., Kearney, K. A., Reid, R., Ritzwoller, D. P., Strecher, V. J., Couper, M. P., Green, B., & Wildenhaus, K. (2007). Reach, engagement, and retention in an internet-based weight loss program in a multi-site randomized controlled trial. *Journal of Medical Internet Research, 9*(2), e11. https://doi.org/10.2196/jmir.9.2.e11

Godino, J. G., Merchant, G., Norman, G. J., Donohue, M. C., Marshall, S. J., Fowler, J. H., Calfas, K. J., Huang, J. S., Rock, C. L., Griswold, W. G., Gupta, A., Raab, F., Fogg, B. J., Robinson, T. N., & Patrick, K. (2016). Using social and mobile tools for weight loss in overweight and obese young adults (Project Smart): A 2 year, parallel-group, randomised, controlled trial. *The Lancet Diabetes & Endocrinology, 4*(9), 747-755. https://doi.org/10.1016/S2213-8587(16)30105-X

Gordon, W. J., Coravos, A. R., & Stern, A. D. (2021). Ushering in safe, effective, secure, and ethical medicine in the digital era. *npj Digital Medicine, 4*(1), 56. https://doi.org/10.1038/s41746-021-00424-5

Grady, A., Yoong, S. L., Wolfenden, L., Wyse, R., Sutherland, R., Nathan, N., Mclaughlin, M., Delaney, T., & Hodder, R. (2018). *The effectiveness of strategies to improve user engagement with digital health interventions to improve risk factors for chronic disease: A systematic review.* (CRD42018077333)[Research Protocol]. PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018077333

Graham, A. K., Kwasny, M. J., Lattie, E. G., Greene, C. J., Gupta, N. V., Reddy, M., & Mohr, D. C. (2021). Targeting subjective engagement in experimental therapeutics for digital mental health interventions. *Internet Interventions, 25*, 100403. https://doi.org/https://doi.org/10.1016/j.invent.2021.100403

Graybill, E., Self-Brown, S., Lai, B., Vinoski, E., McGill, T., & Crimmins, D. (2016). Addressing disparities in parent education: Examining the effects of Learn the Signs/Act Early parent education materials on parent outcomes. *Early Childhood Education Journal, 44*(1), 31-38. https://doi.org/10.1007/s10643-014-0680-3

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*(5), 701.

Grimes, A., & Grinter, R. E. (2007). Designing persuasion: Health technology for low-income African American communities. In *International Conference on Persuasive Technology* (pp. 24-35). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-77006-0_4

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough?: An experiment with data saturation and variability. *Field Methods, 18*(1), 59-82. https://doi.org/10.1177/1525822X05279903

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis*. Thousand Oaks, CA: Sage Publications.

Guralnick, M. J. (1991). The next decade of research on the effectiveness of early intervention. *Exceptional Children, 58*(2), 174-183. https://doi.org/10.1177/001440299105800209

Guralnick, M. J., & Bricker, D. (1987). The effectiveness of early intervention for children with cognitive and general developmental delays. In M. J. Guralnick & F. C. Bennett (Eds.), *The effectiveness of early intervention for at-risk and handicapped children* (pp. 115–173). Academic Press.

Hakobyan, L., Lumsden, J., Shaw, R., & O'Sullivan, D. (2016). A longitudinal evaluation of the acceptability and impact of a diet diary app for older adults with age-related macular degeneration. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 124-134). Florence, Italy. https://doi.org/10.1145/2935334.2935356

Hall, A. K., Cole-Lewis, H., & Bernhardt, J. M. (2015). Mobile text messaging for health: A systematic review of reviews. *Annual Review of Public Health, 36*, 393-415. https://doi.org/10.1146/annurev-publhealth-031914-122855

Hanmer, J., & Cherepanov, D. (2016). A single question about a respondent's perceived financial ability to pay monthly bills explains more variance in health utility scores than absolute income and assets questions. *Quality of Life Research, 25*(9), 2233-2237. https://doi.org/10.1007/s11136-016-1269-7

Hartzler, A. L., Venkatakrishnan, A., Mohan, S., Silva, M., Lozano, P., Ralston, J. D., Ludman, E., Rosenberg, D., Newton, K. M., Nelson, L., & Pirolli, P. (2016). Acceptability of a team-based mobile health (mHealth) application for lifestyle self-management in individuals with chronic illnesses. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 3277-3281). IEEE, Orlando, FL, USA. https://doi.org/10.1109/EMBC.2016.7591428

Hayes, G. R., Cheng, K. G., Hirano, S. H., Tang, K. P., Nagel, M. S., & Baker, D. E. (2014). Estrellita: A mobile capture and access tool for the support of preterm infants and their caregivers. *ACM Transactions on Computer-Human Interaction, 21*(3), Article 19, 1-28. https://doi.org/10.1145/2617574

Health Resources and Services Administration. (2021). *Health center program: Impact and growth*. Retrieved February 12, 2022 from https://bphc.hrsa.gov/about/healthcenterprogram/index.html

Herbst, E., McCaslin, S. E., Hassanbeigi Daryani, S., Laird, K. T., Hopkins, L. B., Pennington, D., & Kuhn, E. (2019). A qualitative examination of stay quit coach, a mobile application for veteran smokers with posttraumatic stress disorder. *Nicotine & Tobacco Research, 22*(4), 560-569. https://doi.org/10.1093/ntr/ntz037

Hertzog, M. A. (2008). Considerations in determining sample size for pilot studies. *Research in Nursing and Health, 31*(2), 180-191. https://doi.org/https://doi.org/10.1002/nur.20247

Higgins, J. P. T., Savović, J., Page, M. J., Elbers, R. G., & Sterne, J. A. (2019). Assessing risk of bias in a randomized trial. *Cochrane Handbook for Systematic Reviews of Interventions*, 205-228. https://doi.org/10.1002/9781119536604.ch8

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019)*. Cochrane. www.training.cochrane.org/handbook.

Hirai, A. H., Kogan, M. D., Kandasamy, V., Reuland, C., & Bethell, C. (2018). Prevalence and variation of developmental screening and surveillance in early childhood. *JAMA Pediatrics, 172*(9), 857-866. https://doi.org/10.1001/jamapediatrics.2018.1524

Hoek, J., Gendall, P., Eckert, C., Rolls, K., & Louviere, J. (2016). A comparison of on-pack quitline information formats. *Tobacco Control, 25*(2), 211. https://doi.org/10.1136/tobaccocontrol-2014-051820

Holden, R. J., & Karsh, B. T. (2010). The technology acceptance model: Its past and its future in health care. *Journal of Biomedical Informatics, 43*(1), 159-172. https://doi.org/https://doi.org/10.1016/j.jbi.2009.07.002

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Ipsen, C., Kurth, N., & Hall, J. (2021). Evaluating mturk as a recruitment tool for rural people with disabilities. *Disability and Health Journal, 14*(1), 100991. https://doi.org/https://doi.org/10.1016/j.dhjo.2020.100991

Jaspers, M. W. M. (2009). A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. *International Journal of Medical Informatics, 78*(5), 340-353. https://doi.org/10.1016/j.ijmedinf.2008.10.002

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*(9), 641-661. https://doi.org/10.1016/j.ijhcs.2008.04.004

Jimenez, M. E., Barg, F. K., Guevara, J. P., Gerdes, M., & Fiks, A. G. (2012). Barriers to evaluation for early intervention services: Parent and early intervention employee perspectives. *Academic Pediatrics, 12*(6), 551-557. https://doi.org/10.1016/j.acap.2012.08.006

Jimenez, M. E., DuRivage, N. E., Bezpalko, O., Suh, A., Wade, R., Blum, N. J., & Fiks, A. G. (2017). A pilot randomized trial of a video patient decision aid to facilitate early intervention referrals from primary care. *Clinical Pediatrics, 56*(3), 268-277. https://doi.org/10.1177/0009922816677038

Jimenez, M. E., Fiks, A. G., Shah, L. R., Gerdes, M., Ni, A. Y., Pati, S., & Guevara, J. P. (2014). Factors associated with early intervention referral and evaluation: A mixed methods analysis. *Academic Pediatrics, 14*(3), 315-323. https://doi.org/10.1016/j.acap.2014.01.007

Kawachi, I., Daniels, N., & Robinson, D. E. (2005). Health disparities by race and class: Why both matter. *Health Affairs, 24*(2), 343-352. https://doi.org/10.1377/hlthaff.24.2.343

Kawachi, I., Subramanian, S. V., & Almeida-Filho, N. (2002). A glossary for health inequalities. *Journal of Epidemiology and Community Health, 56*(9), 647. https://doi.org/10.1136/jech.56.9.647

Kelders, S. M., & Kip, H. (2019). Development and initial validation of a scale to measure engagement with eHealth technologies. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-6). Glasgow, Scotland UK. https://doi.org/10.1145/3290607.3312917

Kelders, S. M., Kip, H., & Greeff, J. (2020). Psychometric evaluation of the TWente engagement with ehealth technologies scale (TWEETS): Evaluation study. *Journal of Medical Internet Research, 22*(10), e17757. https://doi.org/10.2196/17757

Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. W. C. (2012). Persuasive system design does matter: A systematic review of adherence to web-based interventions. *Journal of Medical Internet Research, 14*(6), e152. https://doi.org/10.2196/jmir.2104

Kelders, S. M., van Zyl, L. E., & Ludden, G. D. S. (2020). The concept and components of engagement in different domains applied to eHealth: A systematic scoping review. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.00926

Khan, D. U., Ananthanarayan, S., Le, A. T., Schaefbauer, C. L., & Siek, K. A. (2012). Designing mobile snack application for low socioeconomic status families. In *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops* (pp. 57-64). IEEE, San Diego, CA, USA. http://doi.org/10.4108/icst.pervasivehealth.2012.248692

Kiing, J. S. H., Neihart, M., & Chan, Y. H. (2019). Teachers' role in identifying young children at risk for developmental delay and disabilities: Usefulness of the parents evaluation of developmental status tool. *Child: Care, Health and Development, 45*(5), 637-643. https://doi.org/10.1111/cch.12693

Kim, K. K., Logan, H. C., Young, E., & Sabee, C. M. (2015). Youth-centered design and usage results of the in touch mobile self-management program for overweight/obesity. *Personal and Ubiquitous Computing, 19*(1), 59-68. https://doi.org/10.1007/s00779-014-0808-x

King, A. J., Bol, N., Cummins, R. G., & John, K. K. (2019). Improving visual behavior research in communication science: An overview, review, and reporting recommendations for using eye-tracking methods. *Communication Methods and Measures, 13*(3), 149-177. https://doi.org/10.1080/19312458.2018.1558194

Klasnja, P., Consolvo, S., & Pratt, W. (2011). How to evaluate technologies for health behavior change in HCI research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3063-3072). Vancouver, BC, Canada. https://doi.org/10.1145/1978942.1979396

Knotko, L. (2019). *What's a good app converstion rate for an app store page?* SplitMetrics. Retrieved February 2, 2022 from https://splitmetrics.com/blog/good-app-store-conversion-rate/

Koh, S., Lee, M., Brotzman, L. E., & Shelton, R. C. (2018). An orientation for new researchers to key domains, processes, and resources in implementation science. *Translational Behavioral Medicine, 10*(1), 179-185. https://doi.org/10.1093/tbm/iby095

Kontos, E., Blake, K. D., Chou, W. Y. S., & Prestin, A. (2014). Predictors of eHealth usage: Insights on the digital divide from the Health Information National Trends survey 2012.

Journal of Medical Internet Research, 16(7), e172-e172. https://doi.org/10.2196/jmir.3117

Krebs, P., & Duncan, D. T. (2015). Health app use among US mobile phone owners: A national survey. *JMIR Mhealth Uhealth, 3*(4), e101. https://doi.org/10.2196/mhealth.4924

Kreps, G. L., & Neuhauser, L. (2010). New directions in eHealth communication: Opportunities and challenges. *Patient Education and Counseling, 78*(3), 329-336. https://doi.org/10.1016/j.pec.2010.01.013

Kreuter, M. W., Farrell, D. W., Olevitch, L. R., & Brennan, L. K. (2013). *Tailoring health messages: Customizing communication with computer technology*. Routledge.

Kreuter, M. W., & Wray, R. J. (2003). Tailored and targeted health communication: Strategies for enhancing information relevance. *American Journal of Health Behavior, 27*(1), S227-S232. https://doi.org/10.5993/ajhb.27.1.s3.6

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage Publications. https://us.sagepub.com/en-us/nam/content-analysis/book258450

Labrique, A. B., Vasudevan, L., Kochi, E., Fabricant, R., & Mehl, G. (2013). mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. *Global Health: Science and Practice, 1*(2), 160-171. https://doi.org/10.9745/GHSP-D-13-00031

Lalmas, M., O'Brien, H., & Yom-Tov, E. (2014). Measuring user engagement. *Synthesis Lectures on Information Concepts, Retrieval, and Services, 6*(4), 1-132.

Larsen, M. E., Nicholas, J., & Christensen, H. (2016). A systematic assessment of smartphone tools for suicide prevention. *PloS One, 11*(4), e0152285-e0152285. https://doi.org/10.1371/journal.pone.0152285

Lee, E. W. J., McCloud, R. F., & Viswanath, K. (2022). Designing effective eHealth interventions for underserved groups: Five lessons from a decade of eHealth intervention design and deployment. *Journal of Medical Internet Research, 24*(1), e25419. https://doi.org/10.2196/25419

Lee, E. W. J., & Viswanath, K. (2020). Big data in context: Addressing the twin perils of data absenteeism and chauvinism in the context of health disparities research. *Journal of Medical Internet Research, 22*(1), e16377. https://doi.org/10.2196/16377

Lefebvre, C. R., Tada, Y., Hilfiker, S. W., & Baur, C. (2010). The assessment of user engagement with eHealth content: The eHealth engagement scale1. *Journal of Computer-Mediated Communication, 15*(4), 666-681. https://doi.org/10.1111/j.1083-6101.2009.01514.x

Li, Y., Guo, Y., Hong, Y. A., Zeng, Y., Monroe-Wise, A., Zeng, C., Zhu, M., Zhang, H., Qiao, J., Xu, Z., Cai, W., Li, L., & Liu, C. (2022). Dose–response effects of patient engagement on health outcomes in an mHealth intervention: Secondary analysis of a randomized controlled trial. *JMIR Mhealth Uhealth, 10*(1), e25586. https://doi.org/10.2196/25586

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Journal of Clinical Epidemiology, 62*(10), e1-e34. https://doi.org/10.1016/j.jclinepi.2009.06.006

Lie, S. S., Karlsen, B. r., Oord, R. E., Graue, M., & Oftedal, B. (2017). Dropout from an eHealth intervention for adults with type 2 diabetes: A qualitative study. *Journal of Medical Internet Research, 19*(5), e187. https://doi.org/10.2196/jmir.7479

Lipkin, P. H., Macias, M. M., Baer Chen, B., Coury, D., Gottschlich, E. A., Hyman, S. L., Sisk, B., Wolfe, A., & Levy, S. E. (2020). Trends in pediatricians' developmental screening: 2002–2016. *Pediatrics, 145*(4), e20190851. https://doi.org/10.1542/peds.2019-0851

Lipkin, P. H., Macias, M. M., & the AAP Council On Children With Disabilities, Section on Developmental and Behavioral Pediatrics. (2020). Promoting optimal development: Identifying infants and young children with developmental disorders through developmental surveillance and screening. *Pediatrics, 145*(1), e20193449. https://doi.org/10.1542/peds.2019-3449

Little, A. A., Kamholz, K., Corwin, B. K., Barrero-Castillero, A., & Wang, C. J. (2015). Understanding barriers to early intervention services for preterm infants: Lessons from two states. *Academic Pediatrics, 15*(4), 430-438. https://doi.org/10.1016/j.acap.2014.12.006

Lopez, L., Green, A. R., Tan-McGrory, A., King, R., & Betancourt, J. R. (2011). Bridging the digital divide in health care: The role of health information technology in addressing racial and ethnic disparities. *The Joint Commission Journal on Quality Patient Safety, 37* (10), 437-445. https://doi.org/10.1016/S1553-7250(11)37055-9

Lustria, M. L. A., Cortese, J., Noar, S. M., & Glueckauf, R. L. (2009). Computer-tailored health interventions delivered over the web: Review and analysis of key components. *Patient Education and Counseling, 74*(2), 156-173. https://doi.org/10.1016/j.pec.2008.08.023

Maenner, M. J., Shaw, K. A., Bakian, A. V., Bilder, D. A., Durkin, M. S., Esler, A., Furnier, S. M., Hallas, L., Hall-Lande, J., Hudson, A., Hughes, M. M., Patrick, M., Pierce, K., Poynter, J. N., Salinas, A., Shenouda, J., Vehorn, A., Warren, Z., Constantino, J. N., DiRienzo, M., Fitzgerald, R. T., Grzybowski, A., Spivey, M. H., Pettygrove, S., Zahorodny, W., Ali, A., Andrews, J. G., Baroud, T., Gutierrez, J., Hewitt, A., Lee, L.-C., Lopez, M., Mancilla, K. C., McArthur, D., Schwenk, Y. D., Washington, A., Williams, S., & Cogswell, M. E. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years - Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2018. *Morbidity and Mortality Weekly Report. Surveillance Summaries, 70*(11), 1-16. https://doi.org/10.15585/mmwr.ss7011a1

Magnusson, D. M., & Mistry, K. B. (2017). Racial and ethnic disparities in unmet need for pediatric therapy services: The role of family-centered care. *Academic Pediatrics, 17*(1), 27-33. https://doi.org/10.1016/j.acap.2016.06.010

Majnemer, A., O'Donnell, M., Ogourtsova, T., Kasaai, B., Ballantyne, M., Cohen, E., Collet, J.-P., Dewan, T., Elsabbagh, M., Hanlon-Dearman, A., Filliter, J. H., Lach, L., McElroy, T., McGrath, P., McKellin, W., Miller, A., Patel, H., Rempel, G., Shevell, M., Wittmeier, K., & , T. P. P. (2019). Bright coaching: A randomized controlled trial on the effectiveness of a developmental coach system to empower families of children with emerging developmental delay. *Frontiers in Pediatrics, 7*(332). https://doi.org/10.3389/fped.2019.00332

Mandell, D. S., Ittenbach, R. F., Levy, S. E., & Pinto-Martin, J. A. (2007). Disparities in diagnoses received prior to a diagnosis of autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*(9), 1795-1802. https://doi.org/10.1007/s10803-006-0314-8

Mandell, D. S., Novak, M. M., & Zubritsky, C. D. (2005). Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics, 116*(6), 1480. https://doi.org/10.1542/peds.2005-0185

Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuiseppi, C., Durkin, M. S., Giarelli, E., Morrier, M. J., Nicholas, J. S., & Pinto-Martin, J. A. (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health, 99*(3), 493-498. https://doi.org/10.2105/AJPH.2007.131243

Mani, M., Kavanagh, D. J., Hides, L., & Stoyanov, S. R. (2015). Review and evaluation of mindfulness-based iPhone apps. *JMIR Mhealth Uhealth, 3*(3), e82. https://doi.org/10.2196/mhealth.4328

Mann, J. R., Crawford, S., Wilson, L., & McDermott, S. (2008). Does race influence age of diagnosis for children with developmental delay? *Disability and Health Journal, 1*(3), 157-162. https://doi.org/10.1016/j.dhjo.2008.04.002

Marshall, J., Coulter, M. L., Gorski, P. A., & Ewing, A. (2016). Parent recognition and responses to developmental concerns in young children. *Infants & Young Children, 29*(2), 102-115. https://doi.org/10.1097/iyc.0000000000000056

Mauriello, L. M., Van Marter, D. F., Umanzor, C. D., Castle, P. H., & de Aguiar, E. L. (2016). Using mHealth to deliver behavior change interventions within prenatal care at community health centers. *American Journal of Health Promotion, 30*(7), 554-562. https://doi.org/10.4278/ajhp.140530-QUAN-248

Mazurek, M. O., Handen, B. L., Wodka, E. L., Nowinski, L., Butter, E., & Engelhardt, C. R. (2014). Age at first autism spectrum disorder diagnosis: The role of birth cohort, demographic factors, and clinical features. *Journal of Developmental and Behavioral Pediatrics, 35*(9), 561-569. https://doi.org/10.1097/DBP.0000000000000097

McClure, C., Cunningham, M., Bull, S., Berman, S., & Allison, M. A. (2018). Using mobile health to promote early language development: A narrative review. *Academic Pediatrics, 18*(8), 850-854. https://doi.org/10.1016/j.acap.2018.07.010

Mclaughlin, M., Delaney, T., Campbell, L., Hodder, R., Reilly, K., Kingsland, M., Grady, A., Byaruhanga, J., Sutherland, R., Wiggers, J., & Wolfenden, L. (2018). *Associations between digital health intervention engagement and physical activity and sedentary behaviour: A systematic review.* (CRD42018110657)[Research Protocol]. PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018110657

McVay, M. A., Bennett, G. G., Steinberg, D., & Voils, C. I. (2019). Dose-response research in digital health interventions: Concepts, considerations, and challenges. *Health Psychology, 38*(12), 1168-1174. https://doi.org/10.1037/hea0000805

Mendu, S., Boukhechba, M., Gordon, J. R., Datta, D., Molina, E., Arroyo, G., Proctor, S. K., Wells, K. J., & Barnes, L. E. (2018). Design of a culturally-informed virtual human for educating Hispanic women about cervical cancer. In *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare* (pp. 360-366). New York, NY, USA. https://doi.org/10.1145/3240925.3240968

Michie, S., Johnston, M., Rothman, A. J., de Bruin, M., Kelly, M. P., Carey, R. N., Bohlen, L. E. C., Groarke, H. N. K., Anderson, N. C., & Zink, S. (2021). Developing an evidence-based online method of linking behaviour change techniques and theoretical mechanisms of action: A multiple methods study. *Health Services and Delivery Research, 9*(1), 1-168. https://doi.org/10.3310/hsdr09010

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine, 46*(1), 81-95. https://doi.org/10.1007/s12160-013-9486-6

Miller, D. P., Jr., Denizard-Thompson, N., Weaver, K. E., Case, L. D., Troyer, J. L., Spangler, J. G., Lawler, D., & Pignone, M. P. (2018). Effect of a digital health intervention on receipt of colorectal cancer screening in vulnerable patients: A randomized controlled trial. *Annals of Internal Medicine, 168*(8), 550-557. https://doi.org/10.7326/M17-2315

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & and PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264-269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Mohr, D. C., Schueller, S. M., Riley, W. T., Brown, C. H., Cuijpers, P., Duan, N., Kwasny, M. J., Stiles-Shields, C., & Cheung, K. (2015). Trials of intervention principles: Evaluation methods for evolving behavioral intervention technologies. *Journal of Medical Internet Research, 17*(7), e166. https://doi.org/10.2196/jmir.4391

Mohr, D. C., Tomasino, K. N., Lattie, E. G., Palac, H. L., Kwasny, M. J., Weingardt, K., Karr, C. J., Kaiser, S. M., Rossom, R. C., Bardsley, L. R., Caccamo, L., Stiles-Shields, C., & Schueller, S. M. (2017). Intellicare: An eclectic, skills-based app suite for the treatment of depression and anxiety. *Journal of Medical Internet Research, 19*(1), e10. https://doi.org/10.2196/jmir.6645

Molina, M. D., & Sundar, S. S. (2018). Can mobile apps motivate fitness tracking? A study of technological affordances and workout behaviors. *Health Communication*, 1-10. https://doi.org/10.1080/10410236.2018.1536961

Moore, C., Zamora, I., Patel Gera, M., & Williams, M. E. (2017). Developmental screening and referrals: Assessing the influence of provider specialty, training, and interagency communication. *Clinical Pediatrics*, 1040–1047. https://doi.org/doi.org/10.1177/0009922817701174

Morrison, C., & Doherty, G. (2014). Analyzing engagement in a web-based intervention platform through visualizing log-data. *Journal of Medical Internet Research, 16*(11), e252. https://doi.org/10.2196/jmir.3575

Muessig, K. E., LeGrand, S., Horvath, K. J., Bauermeister, J. A., & Hightow-Weidman, L. B. (2017). Recent mobile health interventions to support medication adherence among HIV-positive MSM. *Current Opinion in HIV and AIDS, 12*(5), 432-441. https://doi.org/10.1097/COH.0000000000000401

Muñoz, D., & Arriaga, R. I. (2015). Low-income parents' values involving the use of technology for accessing health information. In: Abascal, J., Barbosa, S., Fetter, M., Gross, T., Palanque, P., Winckler, M. (eds) Human-Computer Interaction – INTERACT 2015. INTERACT 2015. Springer, Cham. https://doi.org/10.1007/978-3-319-22698-9_5

Murray, E., Hekler, E. B., Andersson, G., Collins, L. M., Doherty, A., Hollis, C., Rivera, D. E., West, R., & Wyatt, J. C. (2016). Evaluating digital health interventions: Key questions and approaches. *American Journal of Preventive Medicine, 51*(5), 843-851. https://doi.org/10.1016/j.amepre.2016.06.008

Ng, A., Reddy, M., Zalta, A. K., & Schueller, S. M. (2018). Veterans' perspectives on fitbit use in treatment for post-traumatic stress disorder: An interview study. *JMIR Mental Health, 5*(2), e10415. https://doi.org/10.2196/10415

Nicholl, B. I., Sandal, L. F., Stochkendahl, M. J., McCallum, M., Suresh, N., Vasseljen, O., Hartvigsen, J., Mork, P. J., Kjaer, P., Søgaard, K., & Mair, F. S. (2017). Digital support interventions for the self-management of low back pain: A systematic review. *Journal of Medical Internet Research, 19*(5), e179. https://doi.org/10.2196/jmir.7290

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the Association for Information Science and Technology, 59*(6), 938-955. https://doi.org/10.1002/asi.20801

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the Association for Information Science and Technology, 61*(1), 50-69. https://doi.org/10.1002/asi.21229

O'Brien, H. L., & Toms, E. G. (2013). Examining the generalizability of the user engagement scale (UES) in exploratory search. *Information Processing & Management, 49*(5), 1092-1107. https://doi.org/10.1016/j.ipm.2012.08.005

O'Connor, S., Hanlon, P., O'Donnell, C. A., Garcia, S., Glanville, J., & Mair, F. S. (2016). Understanding factors affecting patient and public engagement and recruitment to digital health interventions: A systematic review of qualitative studies. *BMC Medical Informatics and Decision Making, 16*(1), 120. https://doi.org/10.1186/s12911-016-0359-3

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews, 5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Pagoto, S., Waring, M., Jake-Schoffman, D., Goetz, J., Michaels, Z., Oleski, J., & DiVito, J. (2018). What type of engagement predicts success in a Facebook weight loss group? In

Proceedings of the 51st Hawaii International Conference on System Sciences (pp. 3304-3312). https://doi.org/10.24251/HICSS.2018.419

Pagoto, S., & Waring, M. E. (2016). A call for a science of engagement: Comment on Rus and Cameron. *Annals of Behavioral Medicine, 50*(5), 690-691. https://doi.org/10.1007/s12160-016-9839-z

Pagoto, S., Waring, M. E., & Xu, R. (2019). A call for a public health agenda for social media research. *Journal of Medical Internet Research, 21*(12), e16661. https://doi.org/10.2196/16661

Patel, M. L., Wakayama, L. N., & Bennett, G. G. (2021). Self-monitoring via digital health in weight loss interventions: A systematic review among adults with overweight or obesity. *Obesity, 29*(3), 478-499. https://doi.org/https://doi.org/10.1002/oby.23088

Pearson, S. D., & Raeke, L. H. (2000). Patients' trust in physicians: Many theories, few measures, and little data. *Journal of General Internal Medicine, 15*(7), 509-513. https://doi.org/10.1046/j.1525-1497.2000.11002.x

Perski, O., Baretta, D., Blandford, A., West, R., & Michie, S. (2018). Engagement features judged by excessive drinkers as most important to include in smartphone applications for alcohol reduction: A mixed-methods study. *Digital Health, 4*. https://doi.org/10.1177/2055207618785841

Perski, O., Blandford, A., Garnett, C., Crane, D., West, R., & Michie, S. (2019). A self-report measure of engagement with digital behavior change interventions (DBCIs): Development and psychometric evaluation of the "DBCI Engagement Scale". *Translational Behavioral Medicine, 10*(1), 267-277. https://doi.org/10.1093/tbm/ibz039

Perski, O., Blandford, A., Ubhi, H. K., West, R., & Michie, S. (2017). Smokers' and drinkers' choice of smartphone applications and expectations of engagement: A think aloud and interview study. *BMC Medical Informatics and Decision Making, 17*(1), 25. https://doi.org/10.1186/s12911-017-0422-8

Perski, O., Blandford, A., West, R., & Michie, S. (2017). Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. *Translational Behavioral Medicine, 7*(2), 254-267. https://doi.org/10.1007/s13142-016-0453-1

Perski, O., Lumsden, J., Garnett, C., Blandford, A., West, R., & Michie, S. (2019). Assessing the psychometric properties of the digital behavior change intervention engagement scale in users of an app for reducing alcohol consumption: Evaluation study. *Journal of Medical Internet Research, 21*(11), e16197. https://doi.org/10.2196/16197

Perski, O., & Short, C. E. (2021). Acceptability of digital health interventions: Embracing the complexity. *Translational Behavioral Medicine, 11*(7), 1473-1480. https://doi.org/10.1093/tbm/ibab048

Peterson, E. T., & Carrabis, J. (2008). *Measuring the immeasurable: Visitor engagement*. Web Analytics Demystified. http://www.verticalstudio.com/hs-fs/hub/74398/file-15425996-pdf/docs/web_analytics_demystified_and_nextstage_global_-_measuring_the_immeasurable_-_visitor_engagement.pdf

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion. Springer series in social psychology* (pp. 1-24). New York, NY: Springer. https://doi.org/10.1007/978-1-4612-4964-1_1

Pew Research Center. (2021a). *Internet/broadband fact sheet*. https://www.pewresearch.org/internet/fact-sheet/internet-broadband/

Pew Research Center. (2021b). *Mobile fact sheet*. https://www.pewresearch.org/internet/fact-sheet/mobile/

Pfammatter, A. F., Mitsos, A., Wang, S., Hood, S. H., & Spring, B. (2017). Evaluating and improving recruitment and retention in an mHealth clinical trial: An example of iterating methods during a trial. *mHealth, 3*, 49. https://doi.org/10.21037/mhealth.2017.09.02

Pratap, A., Renn, B. N., Volponi, J., Mooney, S. D., Gazzaley, A., Arean, P. A., & Anguera, J. A. (2018). Using mobile apps to assess and treat depression in Hispanic and Latino populations: Fully remote randomized clinical trial. *Journal of Medical Internet Research, 20*(8), e10130. https://doi.org/10.2196/10130

Pratt-Chapman, M., Moses, J., & Arem, H. (2021). Strategies for the identification and prevention of survey fraud: Data analysis of a web-based survey. *JMIR Cancer, 7*(3), e30730. https://doi.org/10.2196/30730

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology, 51*(3), 390-395. https://doi.org/10.1037/0022-006X.51.3.390

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research, 38*(2), 65-76. https://doi.org/10.1007/s10488-010-0319-7

Quante, M., Khandpur, N., Kontos, E. Z., Bakker, J. P., Owens, J. A., & Redline, S. (2019). A qualitative assessment of the acceptability of smartphone applications for improving sleep behaviors in low-income and minority adolescents. *Behavioral Sleep Medicine, 17*(5), 573-585. https://doi.org/10.1080/15402002.2018.1432483

Raspa, M., Levis, D. M., Kish-Doto, J., Wallace, I., Rice, C., Barger, B., Green, K. K., & Wolf, R. B. (2015). Examining parents' experiences and information needs regarding early identification of developmental delays: Qualitative research to inform a public health campaign. *Journal of Developmental and Behavioral Pediatrics, 36*(8), 575. https://doi.org/10.1097/DBP.0000000000000205

Ratto, A. B., Reznick, J. S., & Turner-Brown, L. (2015). Cultural effects on the diagnosis of autism spectrum disorder among Latinos. *Focus on Autism and Other Developmental Disabilities, 31*(4), 275-283. https://doi.org/10.1177/1088357615587501

Roberts, A. L., Potts, H. W. W., Koutoukidis, D. A., Smith, L., & Fisher, A. (2019). Breast, prostate, and colorectal cancer survivors' experiences of using publicly available physical activity mobile apps: Qualitative study. *JMIR mHealth and uHealth, 7*(1), Article e10918. https://doi.org/10.2196/10918

Roberts, M. Y., & Kaiser, A. P. (2015). Early intervention for toddlers with language delays: A randomized controlled trial. *Pediatrics, 135*(4), 686-693. https://doi.org/10.1542/peds.2014-2134

Rosenberg, S. A., Robinson, C. C., Shaw, E. F., & Ellison, M. C. (2013). Part C early intervention for infants and toddlers: Percentage eligible versus served. *Pediatrics, 131*(1), 38-46. https://doi.org/10.1542/peds.2012-1662

Rosenberg, S. A., Zhang, D., & Robinson, C. C. (2008). Prevalence of developmental delays and participation in early intervention services for young children. *Pediatrics, 121*(6), e1503-e1509. https://doi.org/10.1542/peds.2007-1680

Rung, A. L., Oral, E., Berghammer, L., & Peters, E. S. (2020). Feasibility and acceptability of a mobile mindfulness meditation intervention among women: Intervention study. *JMIR mHealth and uHealth, 8*(6), Article e15943. https://doi.org/10.2196/15943

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68. https://doi.org/10.1037/0003-066X.55.1.68

Rydz, D., Srour, M., Oskoui, M., Marget, N., Shiller, M., Birnbaum, R., Majnemer, A., & Shevell, M. I. (2006). Screening for developmental delay in the setting of a community pediatric clinic: A prospective assessment of parent-report questionnaires. *Pediatrics, 118*(4), e1178-e1186. https://doi.org/10.1542/peds.2006-0466

Saksono, H., Castaneda-Sceppa, C., Hoffman, J., Morris, V., El-Nasr, M. S., & Parker, A. G. (2020). Storywell: Designing for family fitness app motivation by using social rewards and reflection. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). Honolulu, HI, USA. https://doi.org/10.1145/3313831.3376686

Sardi, L., Idri, A., Redman, L. M., Alami, H., Bezad, R., & Fernández-Alemán, J. L. (2020). Mobile health applications for postnatal care: Review and analysis of functionalities and technical features. *Computer Methods and Programs in Biomedicine, 184*, 105114. https://doi.org/https://doi.org/10.1016/j.cmpb.2019.105114

Saunders, R. P. (2015). *Implementation monitoring and process evaluation*. Thousand Oaks, CA: Sage Publications.

Sawesi, S., Rashrash, M., Phalakornkule, K., Carpenter, J. S., & Jones, J. F. (2016). The impact of information technology on patient engagement and health behavior change: A systematic review of the literature. *JMIR Medical Informatics, 4*(1), e1. https://doi.org/10.2196/medinform.4514

Schaefbauer, C. L., Khan, D. U., Le, A., Sczechowski, G., & Siek, K. A. (2015). Snack Buddy: Supporting healthy snacking in low socioeconomic status families. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 1045-1057. Vancouver, BC, Canada. https://doi.org/10.1145/2675133.2675180

Scherr, C. L., Getachew-Smith, H., Smith, J. D., & Wakschlag, L. S. (In preparation). Strategies for assessing social-emotional concerns in toddlers.

Scherr, C. L., Getachew-Smith, H. J., Sudec, L., Brooks, J. J., & Roberts, M. (2020). Parents' sensemaking processes in the identification of developmental delays and engagement with early intervention services. *Social Science and Medicine, 255*, 112941. https://doi.org/https://doi.org/10.1016/j.socscimed.2020.112941

Sepah, S. C., Jiang, L., Ellis, R. J., McDermott, K., & Peters, A. L. (2017). Engagement and outcomes in a digital diabetes prevention program: 3-year update. *BMJ Open Diabetes Research and Care, 5*(1), e000422. https://doi.org/10.1136/bmjdrc-2017-000422

Sharpe, E. E., Karasouli, E., & Meyer, C. (2017). Examining factors of engagement with digital interventions for weight management: Rapid review. *JMIR Research Protocols, 6*(10), e205. https://doi.org/10.2196/resprot.6059

Shaw, K. A., McArthur, D., Hughes, M. M., Bakian, A. V., Lee, L.-C., Pettygrove, S., & Maenner, M. J. (2021). Progress and disparities in early identification of autism spectrum disorder: Autism and Developmental Disabilities Monitoring Network, 2002-2016. *Journal of the American Academy of Child and Adolescent Psychiatry*. https://doi.org/https://doi.org/10.1016/j.jaac.2021.11.019

Short, C. E., DeSmet, A., Woods, C., Williams, S. L., Maher, C., Middelweerd, A., Müller, A. M., Wark, P. A., Vandelanotte, C., Poppe, L., Hingle, M. D., & Crutzen, R. (2018). Measuring engagement in eHealth and mHealth behavior change interventions: Viewpoint of methodologies. *Journal of Medical Internet Research, 20*(11), e292. https://doi.org/10.2196/jmir.9397

Short, C. E., Vandelanotte, C., Dixon, M. W., Rosenkranz, R., Caperchione, C., Hooker, C., Karunanithi, M., Kolt, G. S., Maeder, A., Ding, H., Taylor, P., & Duncan, M. J. (2014). Examining participant engagement in an information technology-based physical activity and nutrition intervention for men: The Manup randomized controlled trial. *JMIR Research Protocols, 3*(1), e2. https://doi.org/10.2196/resprot.2776

Sices, L., Feudtner, C., McLaughlin, J., Drotar, D., & Williams, M. (2003). How do primary care physicians identify young children with developmental delays? A national survey. *Journal of Developmental and Behavioral Pediatrics, 24*(6), 409-417. https://doi.org/10.1097/00004703-200312000-00002

Sieverink, F., Kelders, S. M., & van Gemert-Pijnen, J. E. W. C. (2017). Clarifying the concept of adherence to eHealth technology: Systematic review on when usage becomes adherence. *Journal of Medical Internet Research, 19*(12), e402. https://doi.org/10.2196/jmir.8578

Simblett, S., Greer, B., Matcham, F., Curtis, H., Polhemus, A., Ferrao, J., Gamble, P., & Wykes, T. (2018). Barriers to and facilitators of engagement with remote measurement technology for managing health: Systematic review and content analysis of findings. *Journal of Medical Internet Research, 20*(7), e10480. https://doi.org/10.2196/10480

Sinnenberg, L., Buttenheim, A. M., Padrez, K., Mancheno, C., Ungar, L., & Merchant, R. M. (2017). Twitter as a tool for health research: A systematic review. *American Journal of Public Health, 107*(1), e1-e8. https://doi.org/10.2105/AJPH.2016.303512

Solar, O., & Irwin, A. (2007). *A conceptual framework for action on the social determinants of health*. World Health Organization. https://www.who.int/sdhconference/resources/ConceptualframeworkforactiononSDH_eng.pdf

Storozuk, A., Ashley, M., Delage, V., & Maloney, E. A. (2020). Got bots? Practical recommendations to protect online survey data from bot attacks. *The Quantitative Methods for Psychology, 16*(5), 472-481. https://doi.org/10.20982/tqmp.16.5.p472

Stow, C., & Dodd, B. (2003). Providing an equitable service to bilingual children in the UK: A review. *International Journal of Language and Communication Disorders, 38*(4), 351-377. https://doi.org/10.1080/1368282031000156888

Stowell, E., Lyson, M. C., Saksono, H., Wurth, R. C., Jimison, H., Pavel, M., & Parker, A. G. (2018). Designing and evaluating mHealth interventions for vulnerable populations: A systematic review. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-17). Montreal, QC, Canada. https://doi.org/10.1145/3173574.3173589

Stoyanov, S. R., Hides, L., Kavanagh, D. J., & Wilson, H. (2016). Development and validation of the user version of the mobile application rating scale (uMARS). *JMIR Mhealth Uhealth, 4*(2), e72. https://doi.org/10.2196/mhealth.5849

Suggs, L. S. (2006). A 10-year retrospective of research in new technologies for health communication. *Journal of Health Communication, 11*(1), 61-74. https://doi.org/10.1080/10810730500461083

Suh, H., Shahriaree, N., Hekler, E. B., & Kientz, J. A. (2016). Developing and validating the user burden scale: A tool for assessing user burden in computing systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3988-3999). San Jose, CA, USA. https://doi.org/10.1145/2858036.2858448

Swindle, T. M., Ward, W. L., Whiteside-Mansell, L., Bokony, P., & Pettit, D. (2014). Technology use and interest among low-income parents of young children: Differences by age group and ethnicity. *Journal of Nutrition Education and Behavior, 46*(6), 484-490. https://doi.org/10.1016/j.jneb.2014.06.004

Szinay, D., Jones, A., Chadborn, T., Brown, J., & Naughton, F. (2020). Influences on the uptake of and engagement with health and well-being smartphone apps: Systematic review. *Journal of Medical Internet Research, 22*(5), e17572. https://doi.org/10.2196/17572

Taki, S., Lymer, S., Russell, C. G., Campbell, K., Laws, R., Ong, K. L., Elliott, R., & Denney-Wilson, E. (2017). Assessing user engagement of an mHealth intervention: Development and implementation of the Growing Healthy app engagement index. *JMIR mHealth and uHealth, 5*(6), e89. https://doi.org/10.2196/mhealth.7236

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thies, K., Anderson, D., & Cramer, B. (2017). Lack of adoption of a mobile app to support patient self-management of diabetes and hypertension in a federally qualified health center: Interview analysis of staff and patients in a failed randomized trial. *JMIR Human Factors, 4*(4), e24. https://doi.org/10.2196/humanfactors.7709

Thomson, J., Hass, C., Horn, I., Kleine, E., Mitchell, S., Gary, K., Ahmed, I., Hamel, D., & Amresh, A. (2017). Aspira: Employing a serious game in an mHealth app to improve asthma outcomes. In *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1-7). IEEE, Perth, WA, Austrailia. https://doi.org/10.1109/SeGAH.2017.7939268

Valente, T. W. (2002). *Evaluating health promotion programs*. Oxford, UK: Oxford University Press.

Valicenti-McDermott, M., Hottinger, K., Seijo, R., & Shulman, L. (2012). Age at diagnosis of autism spectrum disorders. *The Journal of Pediatrics, 161*(3), 554-556. https://doi.org/10.1016/j.jpeds.2012.05.012

van ’t Hof, M., Tisseur, C., van Berckelear-Onnes, I., van Nieuwenhuyzen, A., Daniels, A. M., Deen, M., Hoek, H. W., & Ester, W. A. (2020). Age at autism spectrum disorder diagnosis: A systematic review and meta-analysis from 2012 to 2019. *Autism, 25*(4), 862-873. https://doi.org/10.1177/1362361320971107

Van den Broeck, J., Argeseanu Cunningham, S., Eeckels, R., & Herbst, K. (2005). Data cleaning: Detecting, diagnosing, and editing data abnormalities. *PLoS Medicine, 2*(10), e267. https://doi.org/10.1371/journal.pmed.0020267

Vandelanotte, C., Müller, A. M., Short, C. E., Hingle, M., Nathan, N., Williams, S. L., Lopez, M. L., Parekh, S., & Maher, C. A. (2016). Past, present, and future of eHealth and mHealth research to improve physical activity and dietary behaviors. *Journal of Nutrition Education and Behavior, 48*(3), 219-228.e211. https://doi.org/10.1016/j.jneb.2015.12.006

Walker, A. K., & Fox, E. L. (2018). Why marginalization, not vulnerability, can best identify people in need of special medical and nutrition care. *AMA Journal of Ethics, 20*(10), E941. https://doi.org/10.1001/amajethics.2018.941

Wang, J., Cai, C., Padhye, N., Orlander, P., & Zare, M. (2018). A behavioral lifestyle intervention enhanced with multiple-behavior self-monitoring using mobile and connected tools for underserved individuals with type 2 diabetes and comorbid overweight or obesity: Pilot comparative effectiveness trial. *JMIR Mhealth Uhealth, 6*(4), e92. https://doi.org/10.2196/mhealth.4478

Wellman, S., & Davis, P. (2018). Developmental delay: When to suspect it and when to refer. *InnovAiT, 12*(2), 72-78. https://doi.org/10.1177/1755738018813662

Williams, D. R., Priest, N., & Anderson, N. B. (2016). Understanding associations among race, socioeconomic status, and health: Patterns and prospects. *Health Psychology, 35*(4), 407-411. https://doi.org/10.1037/hea0000242

Wilson, D. B. (2009). Systematic coding. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 159–176). New York, NY: Russell Sage Foundation.

Wittke, K., & Spaulding, T. J. (2018). Which preschool children with specific language impairment receive language intervention? *Language, Speech & Hearing Services in Schools, 49*(1), 59-71. https://doi.org/10.1044/2017_LSHSS-17-0024

Womack, J. J., Anderson, L. N., & Ledford, C. J. W. (2018). Presence of complex and potentially conflicting information in prenatal mobile apps. *Health Promotion Practice, 21*(2), 238-245. https://doi.org/10.1177/1524839918796216

Wong, C. A., Ming, D., Maslow, G., & Gifford, E. J. (2020). Mitigating the impacts of the COVID-19 pandemic response on at-risk children. *Pediatrics, 146*(1). https://doi.org/10.1542/peds.2020-0973

Woolfenden, S., Posada, N., Krchnakova, R., Crawford, J., Gilbert, J., Jursik, B., Sarkozy, V., Perkins, D., & Kemp, L. (2015). Equitable access to developmental surveillance and early intervention – understanding the barriers for children from culturally and linguistically diverse (CALD) backgrounds. *Health Expectations, 18*(6), 3286-3301. https://doi.org/10.1111/hex.12318

Wrigley, A., & Dawson, A. (2016). Vulnerability and marginalized populations. In H. D. Barrett, W. L. Ortmann, A. Dawson, C. Saenz, A. Reis, & G. Bolan (Eds.), *Public health ethics: Cases spanning the globe* (Vol. 3, pp. 203-240). Springer, Cham. https://doi.org/https://doi.org/10.1007/978-3-319-23847-0_7

Yang, Q., Millette, D., Zhou, C., Beatty, M., Carcioppolo, N., & Wilson, G. (2019). The effectiveness of interactivity in improving mediating variables, behaviors and outcomes of web-based health interventions: A meta-analytic review. *Health Communication*, 1-15. https://doi.org/10.1080/10410236.2019.1631992

Yardley, L., Choudhury, T., Patrick, K., & Michie, S. (2016). Current issues and future directions for research into digital behavior change interventions. *American Journal of Preventive Medicine, 51*(5), 814-815. https://doi.org/https://doi.org/10.1016/j.amepre.2016.07.019

Yardley, L., Spring, B. J., Riper, H., Morrison, L. G., Crane, D. H., Curtis, K., Merchant, G. C., Naughton, F., & Blandford, A. (2016). Understanding and promoting effective engagement with digital behavior change interventions. *American Journal of Preventive Medicine, 51*(5), 833-842. https://doi.org/10.1016/j.amepre.2016.06.015

Yen, P. Y., Wantland, D., & Bakken, S. (2010). Development of a customizable health it usability evaluation scale. In *AMIA…Annual Symposium Proceedings. AMIA Symposium 2010* (pp. 917-921). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3041285/pdf/amia-2010_sympproc_0917.pdf

Yoshikawa, H., Wuermli, A. J., Britto, P. R., Dreyer, B., Leckman, J. F., Lye, S. J., Ponguta, L. A., Richter, L. M., & Stein, A. (2020). Effects of the global coronavirus disease-2019 pandemic on early childhood development: Short- and long-term risks and mitigating program and policy actions. *The Journal of Pediatrics, 223*, 188-193. https://doi.org/10.1016/j.jpeds.2020.05.020

Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., Danielson, M. L., Bitsko, R. H., Blumberg, S. J., Kogan, M. D., & Boyle, C. A. (2019). Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics, 144*(4), e20190811. https://doi.org/10.1542/peds.2019-0811

Zubler, J. M., Wiggins, L. D., Macias, M. M., Whitaker, T. M., Shaw, J. S., Squires, J. K., Pajek, J. A., Wolf, R. B., Slaughter, K. S., Broughton, A. S., Gerndt, K. L., Mlodoch, B. J., & Lipkin, P. H. (2022). Evidence-informed milestones for developmental surveillance tools. *Pediatrics*, e2021052138. https://doi.org/10.1542/peds.2021-052138

Zuckerman, K. E., Mattox, K. M., Sinche, B. K., Blaschke, G. S., & Bethell, C. (2014). Racial, ethnic, and language disparities in early childhood developmental/behavioral evaluations:

A narrative review. *Clinical Pediatrics, 53*(7), 619-631.

https://doi.org/10.1177/0009922813501378

Appendices

Appendix A

Example Search Strategy: Embase

1. mobile:ab,ti
2. 'mobile application':ab,ti
3. 'mobile app':ab,ti
4. 'smartphone application':ab,ti
5. 'smartphone app':ab,ti
6. 'mobile application'/exp
7. 'smartphone application'/exp
8. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7
9. 'health'/exp
10. health:ti,ab
11. 9 OR 10
12. intervention:ti,ab
13. 'engagement'/exp
14. interactive:ti,ab
15. 13 OR 14
16. diverse:ti,ab
17. disadvantaged:ti,ab
18. marginalized:ti,ab
19. 'vulnerable population'/exp
20. underserved:ti,ab
21. 'social status'/exp
22. 'economically disadvantaged':ti,ab
23. 'economically vulnerable':ti,ab
24. 'socioeconomic status':ti,ab
25. 'socio-economic status':ti,ab
26. 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25

27. 'social media'/exp
28. 'facebook'/exp
29. 'twitter'/exp
30. 'instagram'/exp
31. snapchat:ti,ab
32. 'whatsapp'/exp
33. 'text message'/exp
34. groupme:ti,ab
35. 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34
36. 8 AND 11 AND 12 AND 15 AND 26 NOT 35

Appendix B

Recruitment Materials

This appendix includes the recruitment materials used to recruit parents and caregivers of young children to complete the two surveys (as reported in Chapter 3). The introductory recruitment blurb for email and social media and recruitment flyer are below.

Research Survey Opportunity (up to $50 gift card) for Parents and Caregivers of Young Children:

The Health Communication Interaction Design Lab at Northwestern University is looking for families with young children (ages 0-5) to share their opinions and experiences with children's health mobile apps. The study includes **2 online surveys** which will take no more than 25 minutes each to complete. You will receive a **$25 gift card** for completing each survey **($50 total**). To participate in this study and complete the online survey, **please click this link: https://bit.ly/kidsappsurvey.** If you have any questions about this study, please email hcid@northwestern.edu.

Are you a parent or caregiver?

We want to hear about your opinions and experiences!

The Health Communication Interaction Design Lab at Northwestern University is looking for families with young children to share their opinions and experiences with children's health apps.

Earn up to $50 for completing 2 surveys

Click here to learn more

For more info contact :

hcid@northwestern.edu

Northwestern

IRB STU#00215896

https://bit.ly/kidsappsurvey

Appendix C

Detecting Fraudulent Responses

A combination of strategies to dissuade fraud used successfully in other online surveys (Ballard et al., 2019; Pratt-Chapman et al., 2021; Storozuk et al., 2020) were built into the survey prior to launch. Fraud detection strategies are outlined in Table C1 and included system-generated tactics (e.g., selecting the fraud deterrent options in Qualtrics built into the survey design) and user-generated tactics (i.e., required participants to enter pairs of items that could be compared for consistency). A fraud detection protocol was created and the survey was continuously monitored for signs of unusual responses indicating fraudulent responses. Embedded in the study compensation protocol (Figure C1), the fraud detection checklist (Table C2) was completed in real-time for every eligible survey response. Responses received points and were categorized as "no fraud," "low fraud," or "high fraud," where both low and high indicated suspected fraud. Those identified as suspicious received emails (Figure C2) that their responses did not pass quality checks. Responses flagged as low fraud received the low fraud email and responses flagged as high fraud received the high fraud email, which differed in tone, the amount of information provided, and contact information to confirm their response. Respondents who replied and confirmed personal information to validate their survey were included in the dataset and received the incentive. Respondents who did not respond were considered fraudulent (invalid) and did not receive the incentive. Despite these efforts to prevent survey fraud, two different versions of the baseline survey were still attacked by bots, however both were caught early due to continuous monitoring.

Table C1

Strategies to Deter Fraud

System-Generated	User-Generated
1. Included language in the consent informing potential participants that duplicate responses would forfeit compensation: “You can only take this survey **one time**. If you take this survey more than one time you **will not** receive any incentive.” 2. Included language in the consent informing potential participants that the investigators reserve the right to confirm eligibility: “Our study team reserves the right to confirm eligibility by email, telephone, or text.” 3. Added Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) question 4. Selected “prevent ballot box stuffing” option in Qualtrics 5. Selected “prevent indexing” option in Qualtrics 6. Added time stamps, calculate total duration as embedded data items 7. Collected geolocation to compare with response to state or territory question (within-subjects comparison) 8. Collected IP address (between-subjects comparison) 9. Collected contact information at the end of the survey to make the process to receive the incentive more time consuming	10. Required eligible participants to confirm their understanding that fraudulent responses would not be compensated 11. Required eligible participants to retype a word: “By entering “Yes” below and continuing this survey, you are acknowledging that responses from ineligible respondents or those who respond multiple times will not receive the incentive. Please type “Yes” in the box below.” 12. Required respondents to select their state or territory to compare with geolocation (within-subjects comparison) 13. Collected personal information e.g., name and email address (both within-subjects and between-subjects comparison) 14. Included question asking how the respondent heard about the survey: “How did you hear about this study?”

Note. Fraud deterrent strategies adapted from Ballard et al., 2019; Pratt-Chapman et al., 2021; Storozuk et al., 2020. Some strategies were used in tandem to detect fraudulent responses and invalidate responses.

Figure C1

Compensation Protocol for Surveys with Parents and Caregivers

Compensation Checklist

1. For completed and eligible responses, enter the following into the tracker [variable name in brackets]
    a. ResponseID [ResponseId]
    b. Name [name]
    c. Email [email]
    d. Phone [phone]
    e. Date baseline complete [RecordedDate]
2. Complete fraud detection checklist
    a. If 3 or more points, categorize as HIGH fraud
    b. If less 1-2 points, categorize as LOW fraud
    c. If 0 points, categorize as NO fraud
3. Update tracker with date fraud checklist complete and fraud type (HIGH, LOW, or NO)
4. Update tracker for estimated date to send follow-up
5. For LOW fraud
    a. Check phone numbers using whitepages.com to see if invalid phone number or business. If invalid or business phone number, add 1 point for possible fraud
6. After survey closes, compare name, email, address, phone, age variables. If 2 personal information items match, add 1 point for possible fraud
7. For LOW fraud send LOW fraud email
    a. If participant calls, confirm personal information (name, age, number of and/or ages of children, or email address) to validate their survey. If personal information is confirmed, then include response in dataset and process compensation
    b. If participant does not respond, survey should be considered fraudulent (invalid) and they will not receive compensation
8. For HIGH fraud send HIGH fraud email

    a. If participant responds, ask them to call (404) XXX-XXXX and confirm personal information (name, age, number of and/or ages of children, or email address) to validate their survey. If personal information is confirmed, then include response in dataset and process compensation
    b. If participant does not respond, survey should be considered fraudulent (invalid) and they will not receive compensation

9. If emails bounce back for either HIGH fraud or LOW fraud, send text message (using Google number)
10. Update tracker with validity (Valid or Invalid)
11. If valid (= NO fraud or those who confirm personal information), order Visa egift card
12. Update tracker with date baseline gift card sent
13. Update Family Survey Compensation Orders sheet with date, order number, number of gift cards, and amount
14. Send baseline survey thank you email to respondents
15. Generate personalized link in Qualtrics
16. Send follow-up survey invitation email with personalized link
17. Once follow-up survey sent, update tracker with date follow-up link sent
18. Send reminder email if follow-up survey not completed by 1 week after invite sent, update tracker
19. Once follow-up survey complete, update tracker with date follow-up complete
20. Repeat process to order Visa egift card and update tracker

Table C2

Fraud Detection Checklist

	Criteria	0 point	1 points
1	If flagged as bot by Qualtrics, 1 point		
2	If survey duration less than 5 minutes, 1 point		
3	If entered anything other than "Yes" on fraudcheck, 1 point		
4	If response to referral implausible, 1 point		
5	If geolocation (IP address) and state selected are different, 1 point		
6	If email address is suspicious (e.g., alternating letters and numbers a12bcd34e@email.com), 1 point		
7	If email address contains name and is different from name entered, 1 point		
8	If responses to children_age not entered in order (e.g., 1 child entered for Child 4), 1 point		
9	If responses to open-ended Heat Map questions are exactly the same, 1 point		
10	If responses to matrix items exactly the same (straight-line), 1 point		
11	If odd responses to open-ended items, 1 point		
12	If missing data on more than 35% of survey questions, categorize as HIGH fraud		
	TOTAL		

0 points = NO fraud
1-2 points = LOW fraud
3+ points = HIGH fraud

Figure C2

Email and Text Templates for Possible Fraud

HIGH Fraud Email

SUBJECT: Health Study Follow-up

Hello,

You recently completed a survey for a health study online. However, we detected that your survey entry was fraudulent. If you think this is a mistake, please contact us at xxxx@northwestern.edu

Thank you,
Health Communication Interaction Design Lab

LOW Fraud Email

SUBJECT: Health Study Follow-up

Hello,

Thank you for completing the children’s health app study online survey. We have been experiencing fraud in the study and your survey entry did not pass a quality check. We sincerely apologize for the inconvenience if this was an error. Please contact us by calling (404) XXX-XXXX to confirm that you did indeed complete a survey and we will send your $25 incentive.

Thank you,
Health Communication Interaction Design Lab

Appendix D

Baseline Survey Instrument

Consent

The Health Communication Interaction Design Lab at Northwestern University invites you to join a research study to share your opinions and experiences with children's health mobile apps. The following is a brief summary of what you need to know about this study:

- The purpose of this study is to understand how parents and caregivers use a smartphone app to learn about developmental milestones and monitor their children's health.
- You will be asked to complete two online surveys. The first online survey will take **no more than 25 minutes** to complete and the **second** online survey in about **one month** will take **no more than 20 minutes** to complete.
- Your participation in this study is voluntary.
- There are no risks from participating in this study. If you participate, all of your responses will be kept confidential and all identifying information will be removed. De-identified data from this study may be shared with the research community (for example journals and professional conferences).
- You will receive a prepaid gift card for a total of $50 for your participation in this study. You will receive $25 for completing the first survey and $25 for completing the second survey. If you choose not to complete the second survey, you will only receive the $25 gift card for completing the first survey.
- You can only take this survey **one time**. If you take this survey more than one time you **will not** receive any incentive. Our study team reserves the right to confirm eligibility by email, telephone, or text.
- This study is supported by the Department of Communication Studies and the Presidential Fellowship at Northwestern University.
- This study was deemed exempt by the Institutional Review Board at Northwestern University.
- If you want a copy of this consent for your records, you can print it from the screen.
- If you have questions or concerns, please contact the Health Communication Interaction Design Lab at Northwestern University at hcid@northwestern.edu.

If you wish to participate, please click the "I Agree" button and you will be taken to the survey. If you do not wish to participate in this study, please select X in the corner of your browser.

Introduction

Thank you for your interest in our study. Please complete the following survey. Please note that once you start the survey **you cannot go back to questions you already answered**. Please know that there are no right or wrong answers, we are simply looking for your opinions and perspectives based on your experiences.

Please check below to continue.

Screener

We would like to start by asking you a few questions about yourself.

How old are you?

Do you own a smartphone?

○ No
○ Yes

What language do you feel most comfortable speaking and reading?

○ English
○ Spanish
○ Both English and Spanish
○ Another language

How many children do you have?

Do you have children younger than 6 years old?

○ No
○ Yes

Have you ever been told that any of your children has a developmental delay or disability?

○ No
○ Yes

Do any of your children currently receive (or have they received in the past) any of the following therapies or services? Select all that apply.

☐ Speech, occupational, physical, or other therapy provided by a county or state early intervention or special education program (such as Part C or early intervention services or therapy)
☐ Speech, occupational, physical, or other therapy provided by other agencies or organizations
☐ Other therapies or services for developmental delays or disabilities
☐ None of the above

In what state or territory do you currently live?

How would you describe your racial background?

○ American Indian or Alaska Native
○ Asian
○ Black or African American
○ Native Hawaiian or Other Pacific Islander
○ White
○ More than one race
○ Other
○ Prefer not to answer

How would you describe your ethnicity?

○ Hispanic/Latino
○ Not Hispanic/Latino
○ Prefer not to answer

How difficult is it for you to meet the monthly payments on your bills?

Not difficult at all
○

Slightly difficult
○

Somewhat difficult
○

Very difficult
○

Extremely difficult
○

Prefer not to answer
○

Do you have a physical condition, mental illness, impairment, disability or chronic health condition that can affect your daily activities OR that requires you to use special equipment or devices, such as a wheelchair, walker, TDD, or communication device?

○ No
○ Yes
○ Prefer not to answer

Invitation for Eligible Participants

Thank you for answering these questions and for agreeing to participate in our study. This survey should take about **25 minutes** to complete. As a reminder, any responses you provide will be confidential. It is your choice to answer the questions, and you can stop at any time.

By entering "Yes" below and continuing this survey, you are acknowledging that responses from ineligible respondents or those who respond multiple times will not receive the incentive. **Please type "Yes" in the box below.**

☐

Media Usage and Information Sources

The following questions ask about how you use your phone.

How frequently do you use the apps on your smartphone, if at all?

○ Daily
○ Weekly
○ Monthly
○ Never

I use my smartphone to: Select all that apply.

☐ Check my email
☐ Connect with friends via social media
☐ Get news and information via social media
☐ Navigate using Google Maps or similar tools
☐ Read the news
☐ Research things to purchase
☐ Download and play games
☐ Download and use health/fitness apps
☐ Other (please specify)

Have you ever used a smartphone app for your health?

○ No
○ Yes

Please describe the health apps you have used and how you have used them.

What did you like about the health apps?

What did you dislike about the health apps?

Have you ever used a smartphone app for your child’s health?

○ No

○ Yes

Please describe the health apps you have used and how you have used them.

What did you like about the health apps?

What did you dislike about the health apps?

What do you think an app focusing on children's health for families and caregivers should provide or do?

How important is it for you to be able to relate to an app's content?

Very important
○

Important
○

Moderately important
○

Slightly important
○

Not important
○

Child Development

This next section asks about your child/children. If you have more than one child, for the purposes of this survey, please think about your child/children younger than 6 years old.

Please indicate how familiar you are with each of the following:

	Not familiar at all	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar
Developmental milestones	○	○	○	○	○
Early Intervention services (EI) or therapy and other services provided by the county or state at a low cost for young children with developmental delays or disabilities	○	○	○	○	○
Special education services for children older than 3 years old who have developmental delays or disabilities	○	○	○	○	○
Milestone Checklists offered by the Centers for Disease Control and Prevention's (CDC) or other "Learn the Signs. Act Early." program resources	○	○	○	○	○
CDC's Milestone Tracker app	○	○	○	○	○

How concerned are you about your child's development (the way your child plays, learns, speaks, acts, and moves for their age)?

Extremely concerned
○

Moderately concerned
○

Somewhat concerned
○

Slightly concerned
○

Not at all concerned
○

Please explain why you responded this way.

What did (or would) you do about your concerns about how your own child was growing and learning?

What would you do if you were concerned about how your own child was growing and learning?

Who have you talked to, or who would you talk to, about your concerns about how your child gets along or is learning or growing?

Who would you talk to if you had concerns about how your child gets along or is learning or growing?

How do people you know (such as family and friends) typically figure out if their child is behind in how they are growing and learning?

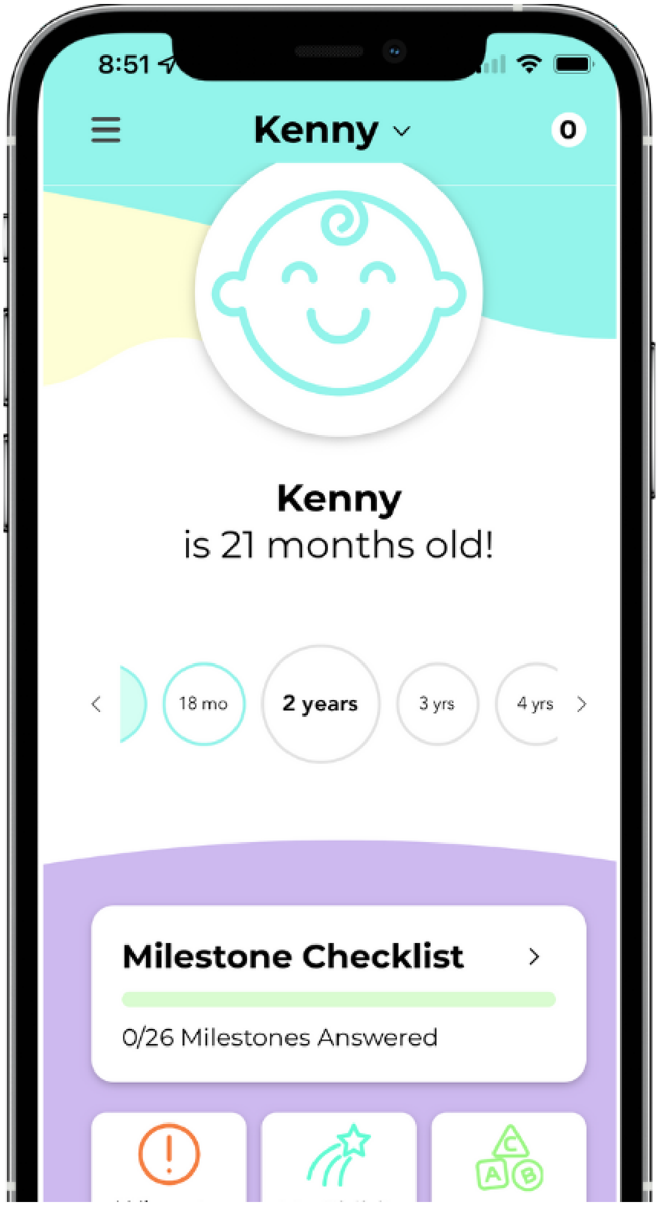
What tools or resources do they use?

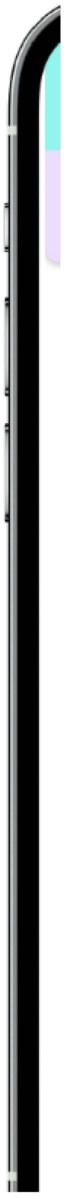
Milestone Tracker App - Introduction

In this next section, you will see images of an app and then will be asked questions about the app. The app is called the *CDC's Milestone Tracker*, an app created by the Centers for

Disease Control and Prevention (CDC) for families and caregivers to learn about and track developmental milestones, learn tips and activities, and learn what to do if they have concerns about their child's development.

Click as fast as you can on the image of the app that **catches your attention**. Please click only one spot.

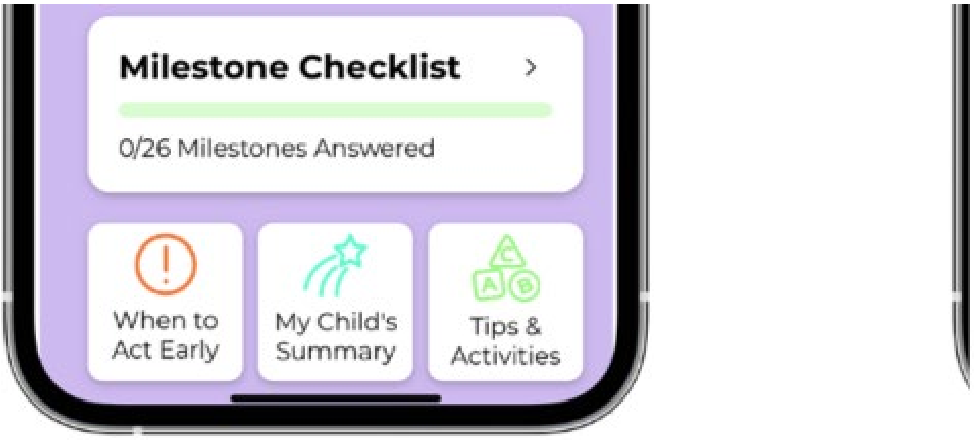




Please explain why you clicked on that spot.

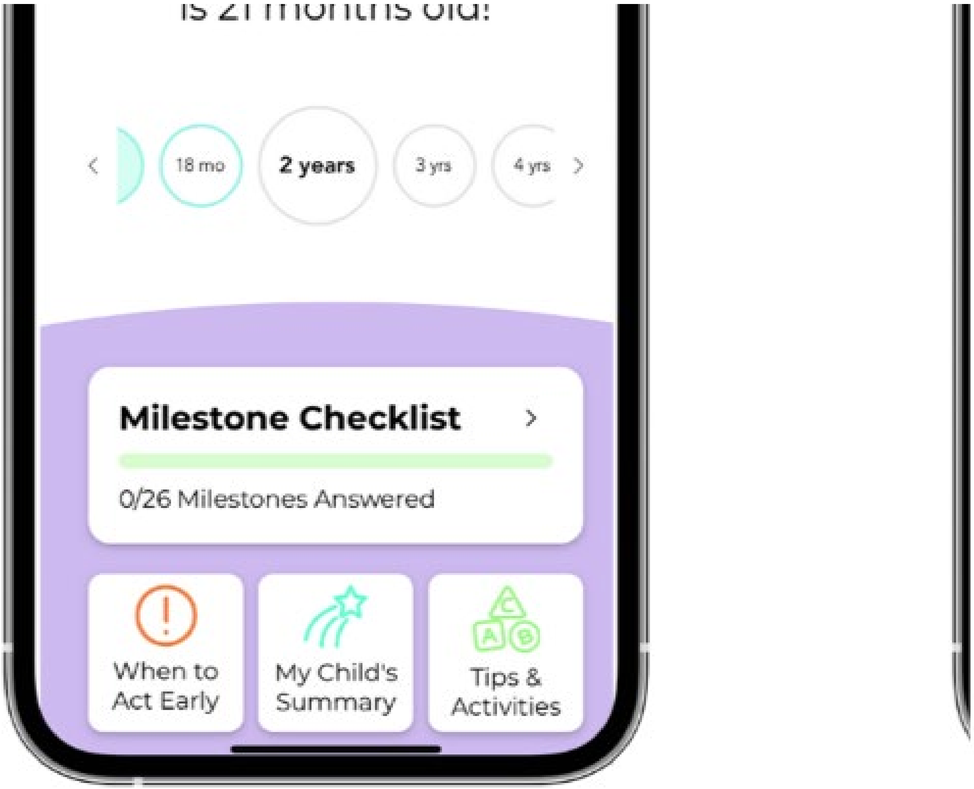
☐

Click on the **most interesting** image of the app. Please click only one spot.



Please explain why you clicked on that spot.

Click on the image of the app that you would **go to first** if you were using the app for the first time. Please click only one spot.



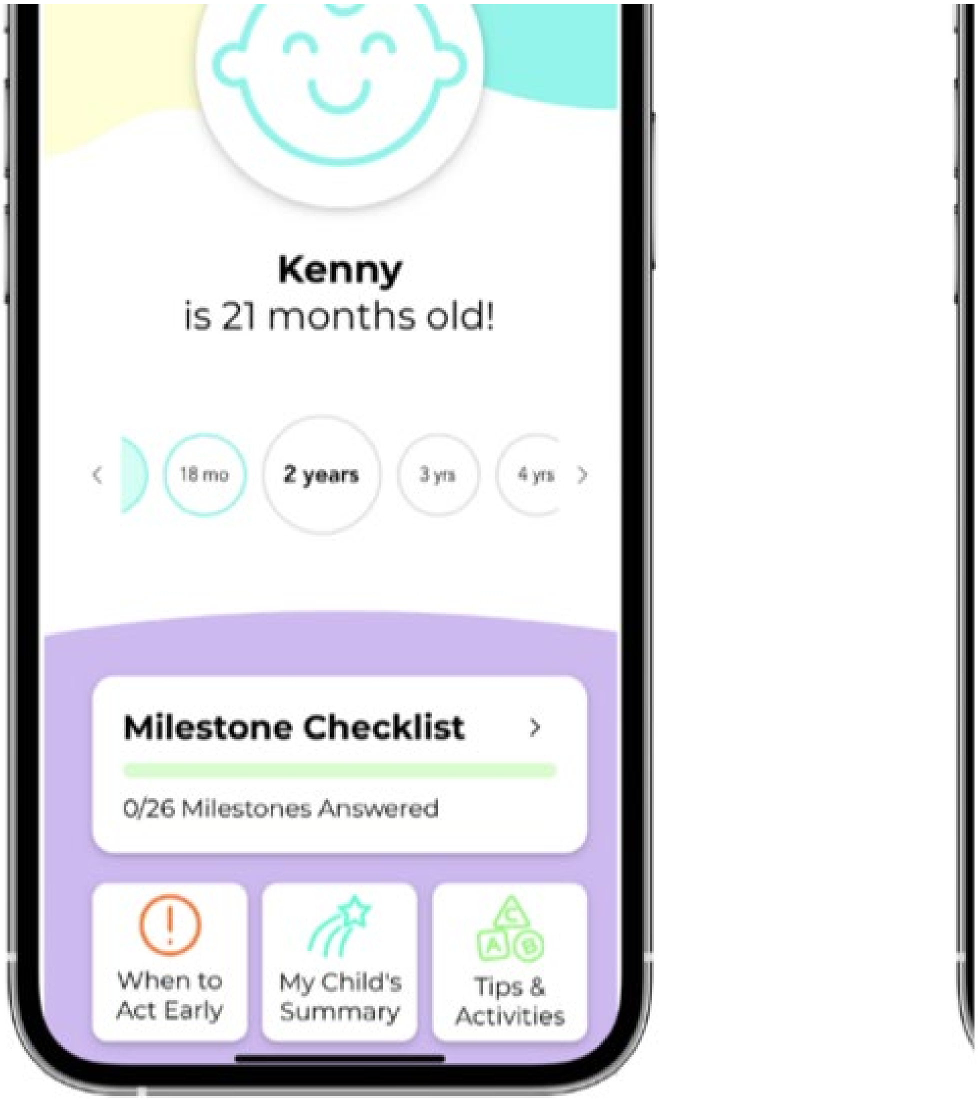
is 21 months old!

Milestone Checklist

0/26 Milestones Answered

Please explain why you clicked on that spot.

Click on the images of the app that you **like**. Please click as many spots as you'd like.



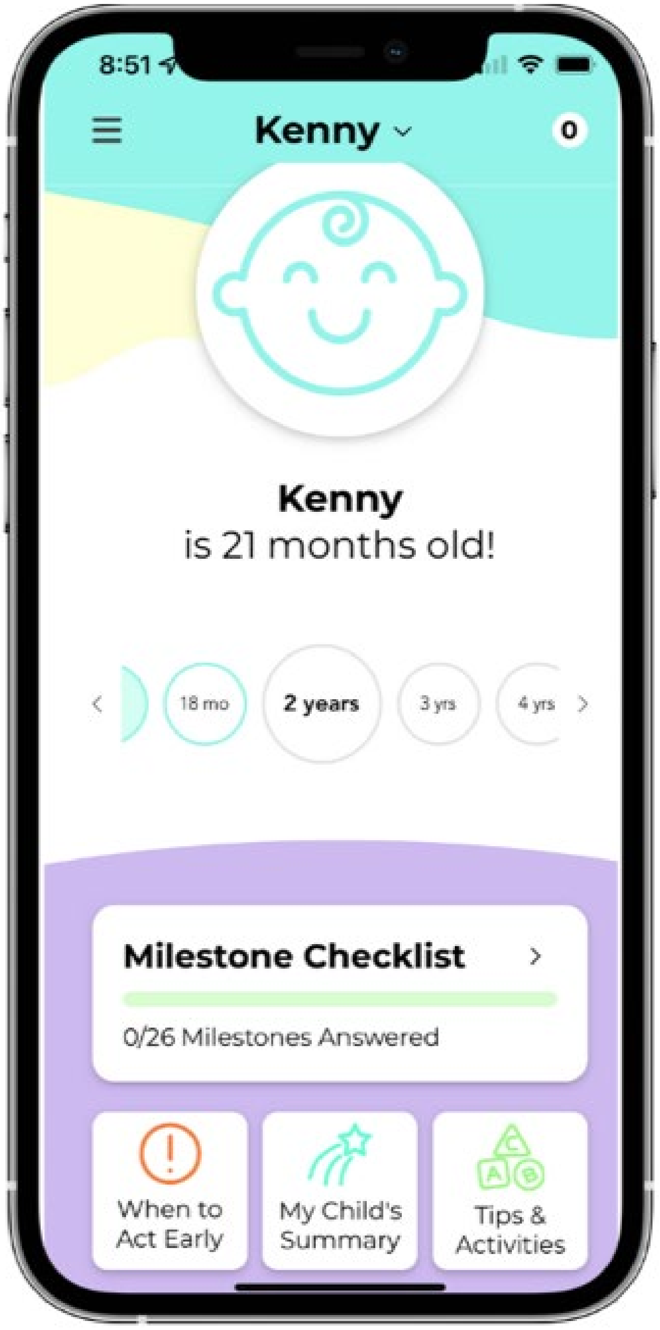
Kenny
is 21 months old!

Milestone Checklist

0/26 Milestones Answered

Please explain why you clicked on those spots.

Click on the images of the app that you **do not like**. Please click as many spots as you'd like.

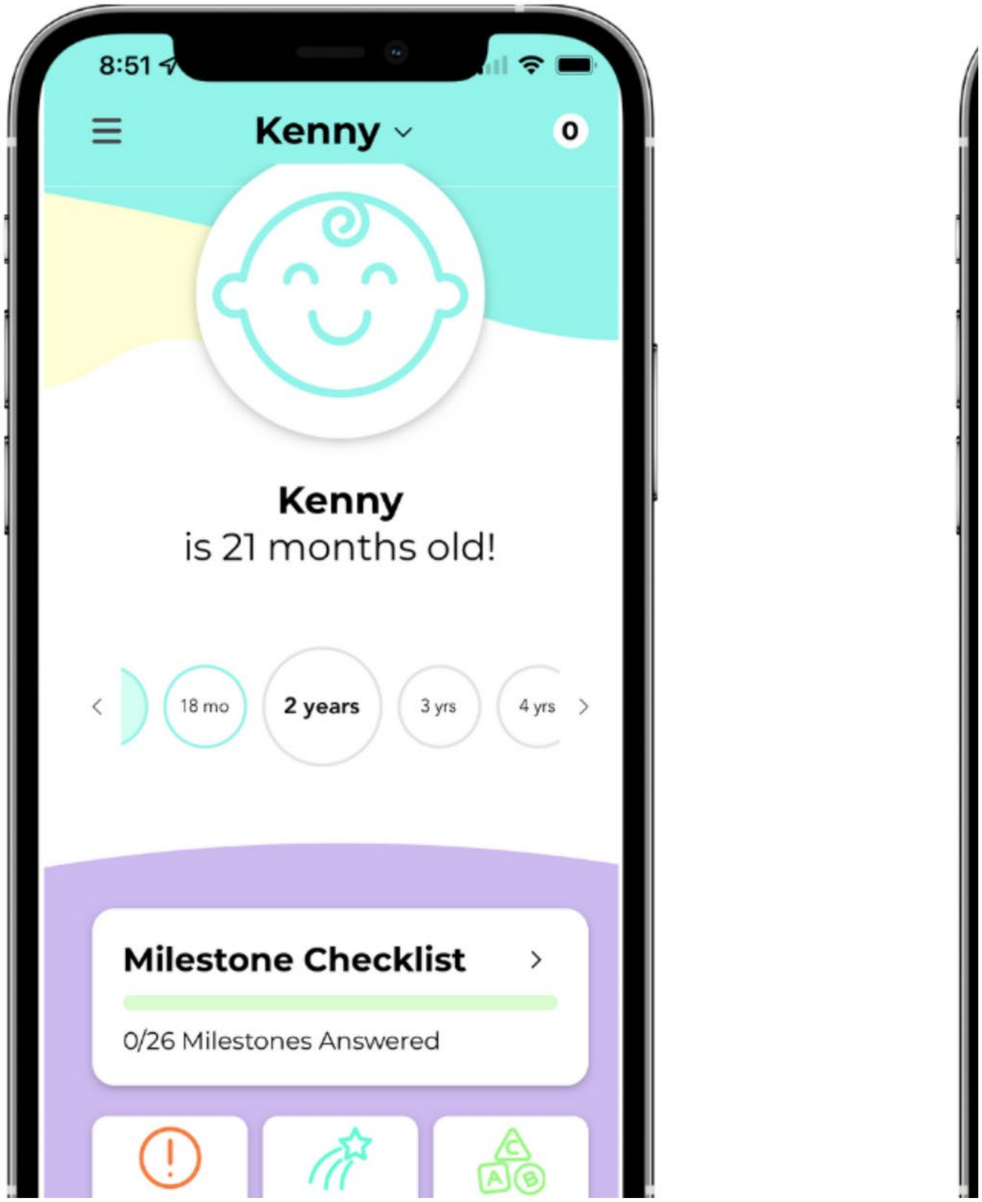


Kenny
is 21 months old!

0/26 Milestones Answered

Please explain why you clicked on those spots.

Click on the images of the app that are **unclear**. Please click as many spots as you'd like.



Please explain why you clicked on those spots.

Milestone Tracker App - Perceived Use, Acceptability, and Behavioral Intentions

This section asks questions about the *CDC's Milestone Tracker* app you just saw.

What are your initial thoughts about the app?

How did the app make you feel?

Did you learn anything new from the *CDC's Milestone Tracker* app?

○ No

○ Yes

What did you learn?

What questions do you have about the app or the content you just saw?

How would you rate this app?

1-5 stars (5 being the highest)

Please explain your answer

Please indicate how much you agree or disagree with the following statements about the *CDC's Milestone Tracker* app:

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I am confident I can track my child's development (the way your child plays, learns, speaks, acts, and moves for their age).	○	○	○	○	○
I plan to use this app to track my child's development in the future.	○	○	○	○	○
I would use this app to learn more about tips and activities that can help my child learn and grow.	○	○	○	○	○
I would use this app to help me talk about my child's development with a doctor or healthcare provider.	○	○	○	○	○
I would use this app to help me know what to do next if I am concerned about my child's development.	○	○	○	○	○

Neither

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I would use this app to share concerns about my child's development with my doctor or healthcare provider.	○	○	○	○	○
I trust this app to help me track my child's development.	○	○	○	○	○
I trust this app to help me identify concerns about my child's development.	○	○	○	○	○
I trust this app to provide accurate information about what milestones my child should be reaching for their age.	○	○	○	○	○
This app was designed for someone like me.	○	○	○	○	○

Which of the following best describes how often you might use the *CDC's Milestone Tracker* app?

○ Daily
○ Weekly
○ Monthly
○ A few times per year (i.e., 3-5 times)
○ Yearly
○ N/A – I do not plan to use the *CDC's Milestone Tracker* app

Would you recommend this app to friends and family?

○ No
○ Yes
○ Maybe

Why or why not?

Demographics

This last set of questions asks you questions about yourself.

What is your gender?

○ Man
○ Woman
○ Non-binary (please specify)

○ Prefer not to answer

How old are your children? Please enter their age and use the dropdown to select months or years.

	Months or years?
Child 1 age	
Child 2 age	
Child 3 age	
Child 4 age	
Child 5 age	

What is your current relationship status?

- ○ Married/Domestic Partner/Civil Union
- ○ Separated or Divorced
- ○ Widowed
- ○ Living with Partner
- ○ Single
- ○ Prefer not to answer

What is the highest grade or level of formal education you have completed?

- ○ Some high school
- ○ High school graduate (or GED)
- ○ Some college (or certificate)
- ○ Associate’s degree (2 year college graduate)
- ○ Bachelor’s degree (4 year college graduate)
- ○ Post graduate degree (MA, MS, JD, PhD, MD, etc.)

○ Prefer not to answer

What is your current household income?

○ Less than $15,000
○ $15,000 to $29,999
○ $30,000 to $44,999
○ $45,000 to $59,999
○ $60,000 to $89,999
○ $90,000 to $149,999
○ $150,000 to $199,999
○ $200,000 or above
○ Prefer not to answer

What is your current employment status?

○ Work full-time
○ Work part-time
○ Unemployed
○ Prefer not to answer
○ Retired and not employed
○ Full-time stay at home parent
○ Self-employed

Do you currently have health insurance or health care coverage?

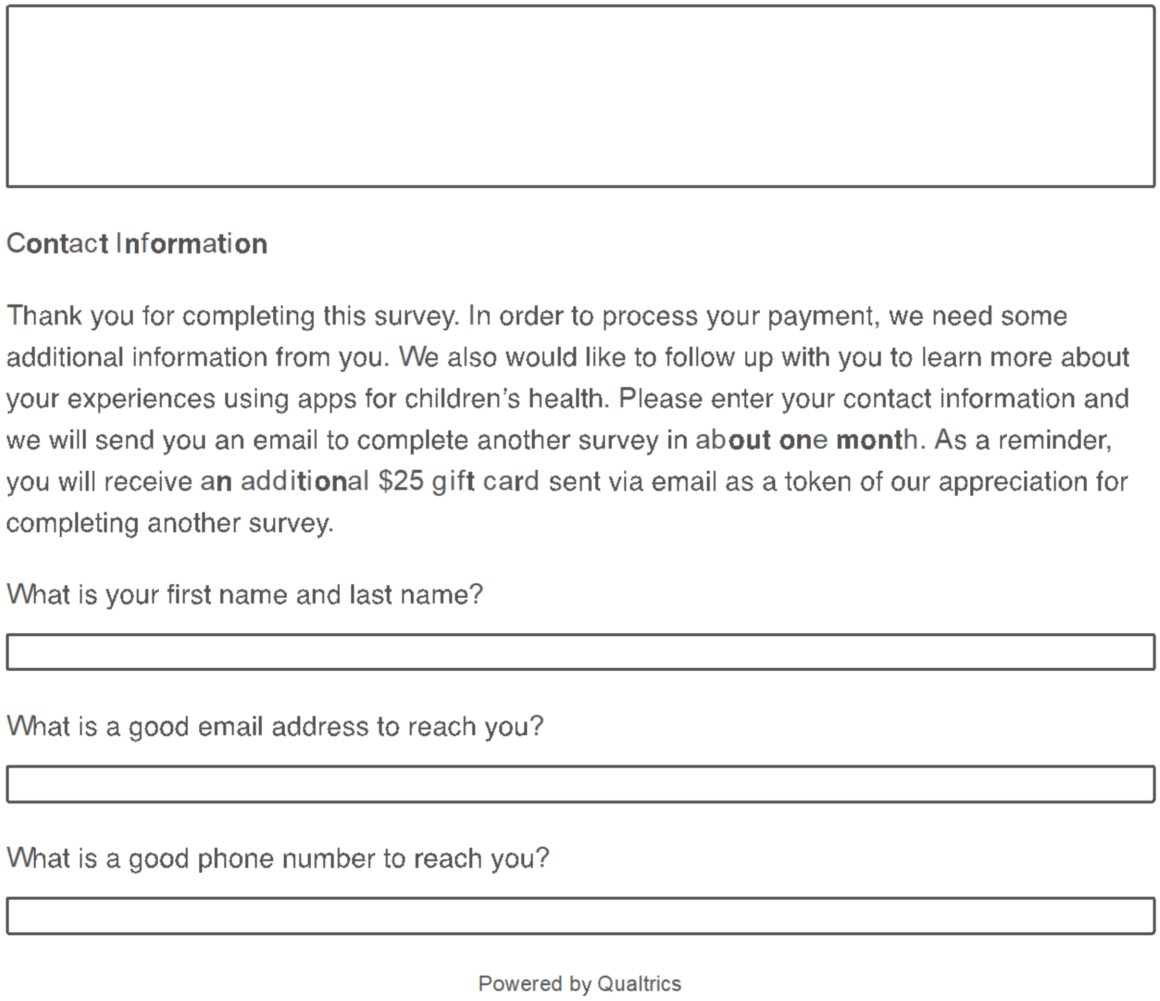
○ No
○ Yes
○ Don't know

What kind of coverage?

○ Private through workplace/spouse or parent’s workplace
○ Private through the Health Insurance Marketplace (Affordable Care Act or HealthCare.gov)
○ Medicare
○ Medicaid
○ Prefer not to answer

How did you hear about this study?

Is there anything else you would like to share?



Contact Information

Thank you for completing this survey. In order to process your payment, we need some additional information from you. We also would like to follow up with you to learn more about your experiences using apps for children’s health. Please enter your contact information and we will send you an email to complete another survey in about one month. As a reminder, you will receive an additional $25 gift card sent via email as a token of our appreciation for completing another survey.

What is your first name and last name?

What is a good email address to reach you?

What is a good phone number to reach you?

Appendix E

Follow-up Survey Instrument

Consent

[illegible]

- [illegible]
- [illegible] second [illegible] no more than 20 minutes [illegible]
- [illegible]
- [illegible]
- [illegible] $50 [illegible] $25 [illegible] $25 [illegible]
- [illegible] one time [illegible] will not [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]

If you wish to participate, please click the "I Agree" button and you will be taken to the survey. If you do not wish to participate in this study, please select X in the corner of your browser.

Introduction

Thank you for continuing to participate in our study. As a reminder, any responses you provide will be confidential. It is your choice to answer the questions, and you can stop at any time. Please note that once you start the survey **you cannot go back to questions you already answered**. This survey should take about 20 minutes to complete. Please know that there are no right or wrong answers, we are simply looking for your opinions and perspectives based on your experiences.

Please check below to continue.

Child Development

This first section asks about your child/children. If you have more than one child, for the purposes of this survey, please think about your child/children younger than 6 years old.

Please indicate how familiar you are with each of the following:

	Not familiar at all	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar
Developmental milestones	○	○	○	○	○
Early Intervention services (EI) or therapy and other services provided by the county or state at a low cost for young children with developmental delays or disabilities	○	○	○	○	○
Special education services for children older than 3 years old who have developmental delays or disabilities	○	○	○	○	○
Milestone Checklists offered by the Centers for Disease Control and Prevention's (CDC) or other "Learn the Signs. Act Early." program resources	○	○	○	○	○
CDC's Milestone Tracker app	○	○	○	○	○

How concerned are you about your child's development (the way your child plays, learns, speaks, acts, and moves for their age)?

Extremely concerned
○

Moderately concerned
○

Somewhat concerned
○

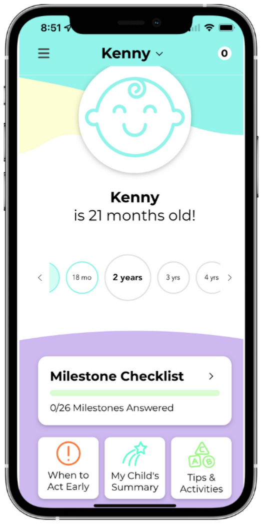
Slightly concerned
○

Not at all concerned
○

Please explain why you responded this way.

Milestone Tracker App - Perceived Use and Acceptability

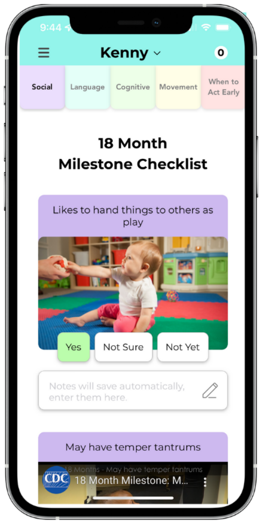
This section asks questions about the Centers for Disease Control and Prevention (or CDC) *Milestone Tracker* app you saw on the first survey. As a reminder, here are some images of the *CDC's Milestone Tracker* app.



Kenny
is 21 months old!

Milestone Checklist

0/26 Milestones Answered



18 Month
Milestone Checklist

Likes to hand things to others as play

Notes will save automatically, enter them here.

May have temper tantrums

18 Months - May have temper tantrums

18 Month Milestone: M...

18 month Summary

Below are the responses you entered for Kenny.

- Click "Email Summary" below to email a summary of your child's milestones to yourself or other care providers.

- Click "Show Doctor" to share a summary during your child's doctor visit.

If you have noted any concerns or selected "not yet" or "not sure" for any milestones, talk to your child's doctor, share the summary, ask for developmental screening, and go to **cdc.gov/FindEI** to learn more about services that can help. For more information, visit **cdc.gov/Concerned.**

How often did you think of the *CDC's Milestone Tracker* app during the past month?

Daily
○

2-3 times per week
○

Once a week
○

2-3 times per month
○

Monthly
○

Never
○

What are your thoughts about the *CDC's Milestone Tracker* app?

How did the app make you feel?

Which components caught your attention the most?

Did you learn anything new from the *CDC's Milestone Tracker* app?

○ No

○ Yes

What did you learn?

What questions do you have about the app?

How would you rate the *CDC's Milestone Tracker* app?

1-5 stars (5 being the highest)

Please explain your answer

Please indicate how much you agree or disagree with the following statements about the *CDC's Milestone Tracker* app:

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I like using this app to track my child's development.	○	○	○	○	○
I have learned more about my child's development using this app.	○	○	○	○	○
The app is easy to use.	○	○	○	○	○
I couldn't find all of the answers I needed in the app.	○	○	○	○	○
I am confident I can track my child's development (the way your child plays, learns, speaks, acts, and moves for their age).	○	○	○	○	○
I plan to use this app to track my child's development in the future.	○	○	○	○	○
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I would use this app to learn more about tips and activities that can help my child learn and grow.	○	○	○	○	○
I would use this app to help me talk about my child's development with a doctor or healthcare provider.	○	○	○	○	○
I would use this app to help me know what to do next if I am concerned about my child's development.	○	○	○	○	○
I would use this app to share concerns about my child's development with my doctor or healthcare provider.	○	○	○	○	○
I trust this app to help me track my child's development.	○	○	○	○	○
I trust this app to help me identify concerns about my child's development.	○	○	○	○	○
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I trust this app to provide accurate information about what milestones my child should be reaching for their age.	○	○	○	○	○
This app was designed for someone like me.	○	○	○	○	○
The language used in the app was easy to understand.	○	○	○	○	○
I was concerned about the Internet data usage on my phone when using the app.	○	○	○	○	○

Milestone Tracker App - Behavioral Outcomes and Behavioral Intentions

This next section asks about your experiences with the *CDC's Milestone Tracker* app since you completed the first survey.

Did you download the *CDC's Milestone Tracker* app to your smartphone or tablet?

○ No

○ Yes

How much time (in minutes) do you think you spent on the app over the past 2 weeks?

☐

Please explain why not

☐

Which of the app's components do you remember visiting? Select all that apply.

☐ Children & Add Child

☐ Milestone Checklist

☐ When to Act Early

☐ My Child's Summary

☐ Tips & Activities

☐ Milestone Overview

☐ Appointments

☐ Other (please explain) ☐

☐ Can't remember

Did you complete any milestone checklists for your child using the *CDC's Milestone Tracker* app?

○ No

○ Yes

How many milestone checklists did you complete?

Which age milestone checklists did you complete?

Please explain why not

Which of the following best describes how often you use the *CDC's Milestone Tracker* app?

○ Daily
○ 2-3 times per week
○ Once a week
○ 2-3 times per month
○ Monthly
○ N/A – I do not use the *CDC's Milestone Tracker* app

Please explain why you used the *CDC's Milestone Tracker* when you did?

What typically prompts you to use the *CDC's Milestone Tracker* app? Select all that apply.

☐ Notifications from the app
☐ When I have a possible concern
☐ Preparing for a doctor's appointment (e.g., well-child visit)
☐ Something else (please explain)
☐ N/A – I do not use the *CDC's Milestone Tracker* app

Who did you share your child's milestone summary with? Select all that apply.

☐ Your child's doctor or health care provider
☐ Your child's childcare provider or teacher
☐ Your family or friends
☐ N/A – I did not share my child's summary with anyone (please explain why not)

☐ Another early childhood professional (please specify)

Did the *CDC's Milestone Tracker* app help you identify any possible concerns related to your child's development (the way your child plays, learns, speaks, acts, and moves for their age)?

○ No
○ Yes

Who did you share these possible concerns with? Select all that apply

- ☐ Your child's doctor or health care provider
- ☐ Your child's childcare provider or teacher
- ☐ Another early childhood professional (please specify)
- ☐ Your family or friends
- ☐ N/A – I have not shared my child's summary with anyone yet but I plan to (please explain who you plan to share concerns with)
- ☐ N/A – I did not share my child's summary with anyone and I do not plan to (please explain why not)

Did your child's doctor or health care provider do any of the following? Select all that apply.

- ☐ Perform a developmental screening (typically involves the parent completing a list of questions about the child's skills and abilities)
- ☐ Refer you to another professional to learn more about your child's development
- ☐ Refer you to intervention services (e.g., speech or another type of therapy)
- ☐ Discuss your child's development and next steps
- ☐ Told you to "wait and see"
- ☐ Something else (please explain)

Did your child receive services to support their development? Select all that apply.

- ☐ Yes, speech, occupational, physical, or other therapy (not provided by a county or state early intervention program)
- ☐ Yes, services from my county or state early intervention program
- ☐ Yes, other services not listed above (please specify)
- ☐ No, my child did not receive services but needed them
- ☐ No, my child did not need or receive services

How engaging is the *CDC's Milestone Tracker* app?

- ○ Extremely engaging
- ○ Moderately engaging
- ○ Somewhat engaging
- ○ Slightly engaging
- ○ Not engaging at all

Please explain your answer

What would make it more engaging?

Which of the following best describes how often you plan to use the *CDC's Milestone Tracker* app in the future?

- ○ Daily
- ○ Weekly
- ○ Monthly
- ○ A few times per year (such as 3-5 times)
- ○ Yearly
- ○ N/A – I do not plan to use the *CDC's Milestone Tracker* app

Please explain why not

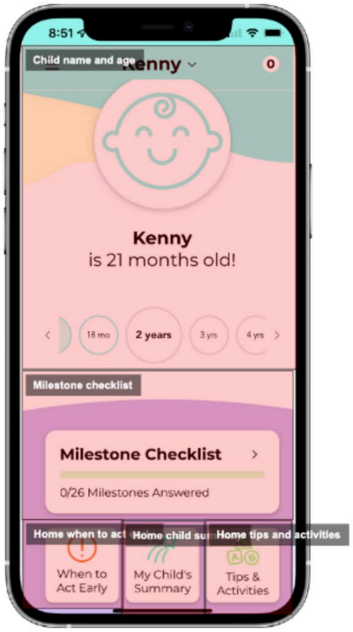
Would you recommend this app to friends and family?

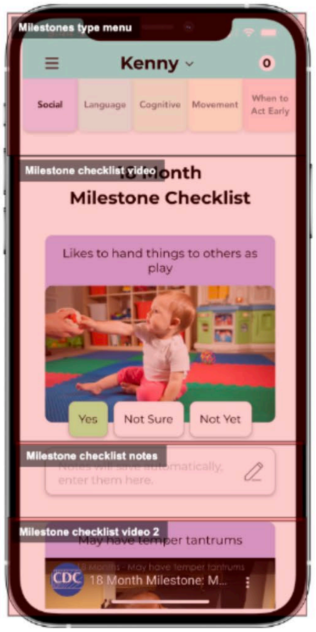
○ No
○ Yes
○ Maybe

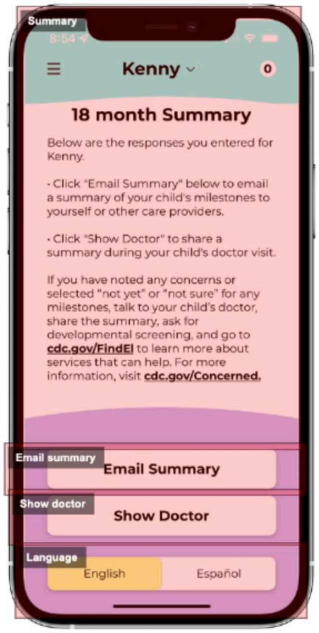
Why or why not?

Is there anything else you would like to share?

Appendix F







18 month Summary

Below are the responses you entered for Kenny.

- Click "Email Summary" below to email a summary of your child's milestones to yourself or other care providers.
- Click "Show Doctor" to share a summary during your child's doctor visit.

If you have noted any concerns or selected "not yet" or "not sure" for any milestones, talk to your child's doctor, share the summary, ask for developmental screening, and go to **cdc.gov/FindEI** to learn more about services that can help. For more information, visit **cdc.gov/Concerned.**

Note. In addition to the 13 designated heat map regions, the Qualtrics survey platform also captured "other" as any click or touch outside of the 13 pre-specified areas.

Appendix G

Comparison of Survey Items to Other Engagement Scales and Items

Table G1

Comparison of Short Measures of Engagement to Survey Items

Short Measure of Engagement [a]	Survey Item [b]
How much did you like the app?	I like using this app to track my child's development. [c]
How engaging was the app?	How engaging is the *CDC's Milestone Tracker app*? [d]

Note. Short measures of engagement developed by Perki et al (2019a) and Perski et al (2019b).
[a] Items asked on 7-point scale ranging from 1 (not at all) to 3 (moderately) to 7 (extremely). [b] Items asked on follow-up survey. [c] Respondents asked to indicate their level of agreement on a scale ranging from 1 (strongly disagree) to 5 (strongly agree). [d] How engaging is the *CDC's Milestone Tracker* app is on a scale ranging from 1 (not engaging at all) to 5 (extremely engaging).

Table G2

Comparison of TWente Engagement with Ehealth Technologies Scale (TWEETS) to Survey Items

Construct	TWEETS Item [a]	Survey Item [b, c]
Behavioral engagement	1. [this technology] is part of my daily routine	1. I plan to use this app to track my child's development in the future.
	2. [this technology] is easy to use	2. The app is easy to use. [d]
	3. I'm able to use [this technology] as often as needed (to achieve my goals)	3a. I would use this app to share concerns about my child's development with my doctor or healthcare provider. 3b. I would use this app to learn more about tips and activities that can help my child learn and grow.
Cognitive engagement	4. [this technology] makes it easier for me to work on [my goal]	4. I trust this app to help me track my child's development.
	5. [this technology] motivates me to [reach my goal]	5. N/A
	6. [this technology] helps me to get more insight into [my behavior relating to the goal]	6a. I trust this app to help me identify concerns about my child's development. 6b. I trust this app to provide accurate information about what milestones my child should be reaching for their age.
Affective engagement	7. I enjoy using [this technology]	7. N/A
	8. I enjoy seeing the progress I make in [this technology]	8. I like using this app to track my 9. child's development. [d]
	9. [This technology] fits me as a person	9. This app was designed for someone like me.

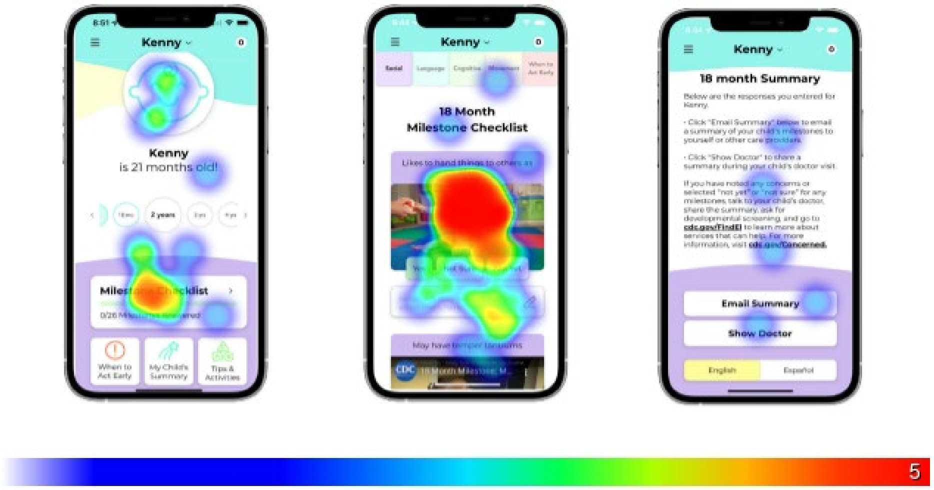
Note. TWEETS developed by Kelders and Kip (2019) and Kelders et al (2020).

[a] TWEETS on a 5-point scale ranging from 0 (strongly disagree) to 4 (strongly agree). [b] Respondents asked to indicate their level of agreement on a scale ranging from 1 (strongly disagree) to 5 (strongly agree). [c] Items asked on baseline and one-month follow-up surveys unless otherwise noted. [d] Item only asked on follow-up survey.

Appendix H

Figure H1

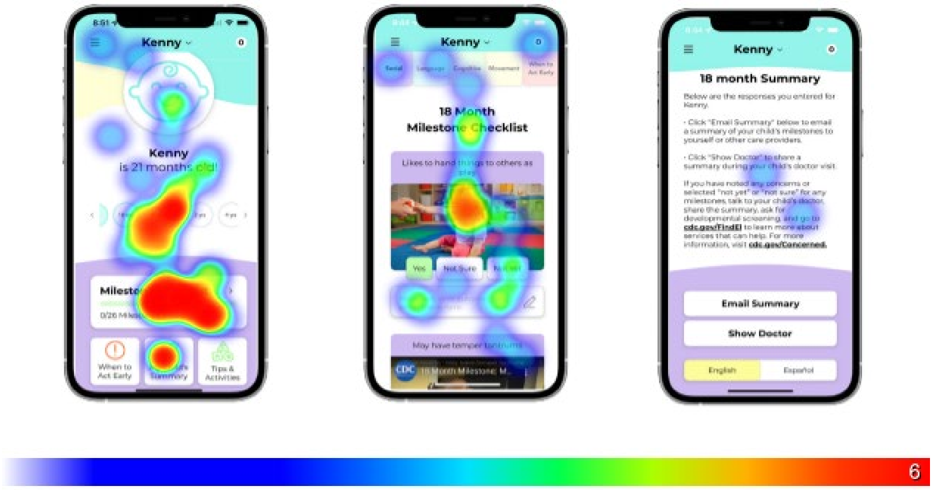
Heat Map of Clicks or Touches on MT App Assessing Most Interesting Image



Note. Respondents were asked “Click on the most interesting image of the app. Please click only one spot.” The bright spots of the heat map represent the relative intensity of total clicks or touches on each part of the MT app image. Red indicates the part of the app that received the highest number of clicks or touches.

Figure H2

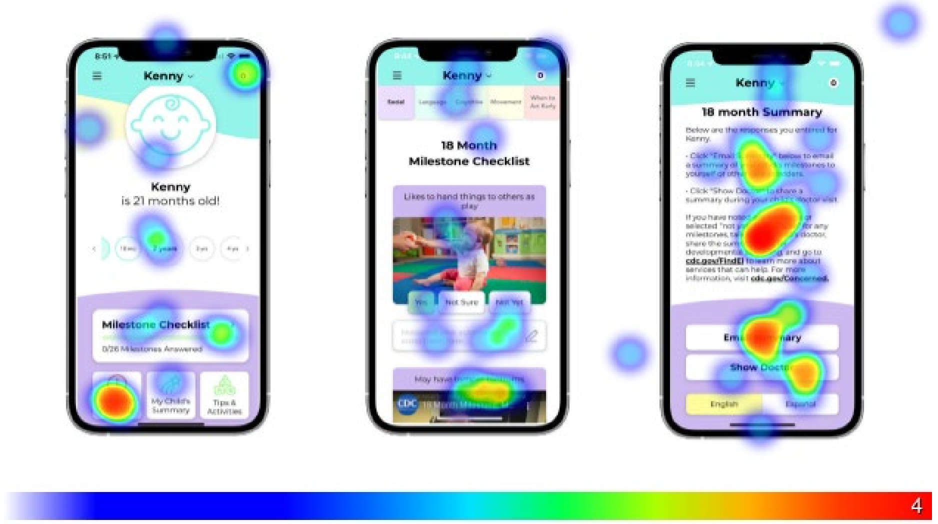
Heat Map of Clicks or Touches on MT App Assessing First Use



Note. Respondents were asked “Click on the image of the app that you would go to first if you were using the app for the first time. Please click only one spot.” The bright spots of the heat map represent the relative intensity of total clicks or touches on each part of the MT app image. Red indicates the part of the app that received the highest number of clicks or touches.

Figure H3

Heat Map of Clicks or Touches on MT App Assessing Confusion



Note. Respondents were asked "Click on the images of the app that are unclear. Please click as many spots as you'd like." The bright spots of the heat map represent the relative intensity of total clicks or touches on each part of the MT app image. Red indicates the part of the app that received the highest number of clicks or touches.

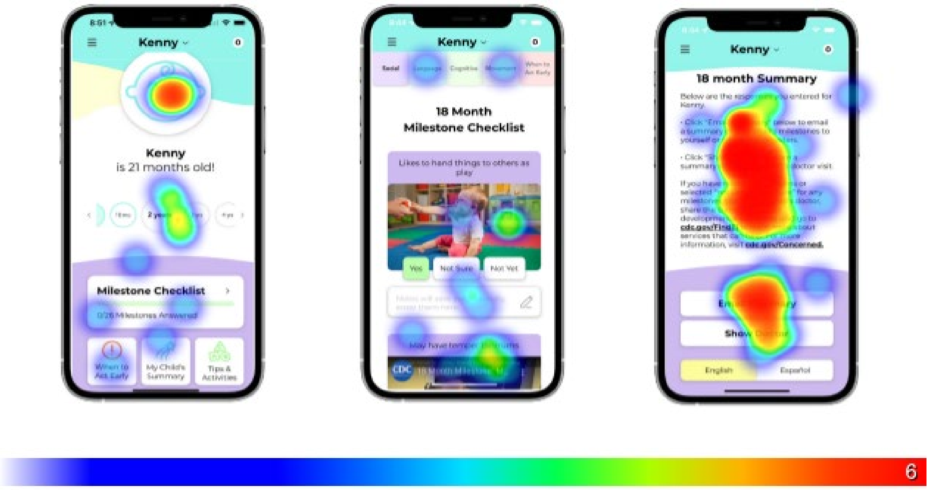
Figure H4

Heat Map of Clicks or Touches on MT App Assessing Satisfaction (Like)

Note. Respondents were asked "Click on the images of the app that you like. Please click as many spots as you'd like." The bright spots of the heat map represent the relative intensity of total clicks or touches on each part of the MT app image. Red indicates the part of the app that received the highest number of clicks or touches.

Figure H5

Heat Map of Clicks or Touches on MT App Assessing Dissatisfaction (Dislike)



Note. Respondents were asked “Click on the images of the app that you do not like. Please click as many spots as you’d like.” The bright spots of the heat map represent the relative intensity of total clicks or touches on each part of the MT app image. Red indicates the part of the app that received the highest number of clicks or touches.

Appendix I

Interview Guide

Semi-Structured Individual Interview Guide Version #2 Date/Time of Interview: Interviewer:	Study ID#:

Semi-Structured Individual Interview Guide: Pediatric Primary Care

Welcome and overview of the interview: Thank you for agreeing to participate in our research study today. My name is [interviewer name], and I want to confirm that now is still a good time to talk for about 60 minutes [offer to reschedule if this is no longer a good time]. Great, I appreciate your time. Before we get started, I would like to review the informed consent document with you to make sure that you understand the study and to confirm your interest in participating. **[Review verbal informed consent document. If consent, start recording**].

To begin, I will explain a little bit about what we hope to accomplish today. The goal of this interview is for us to learn more about how you conduct well-child office visits with toddlers ages 2-3, and hear more about your interactions with their parents or caregivers. The most important thing to keep in mind is that there is not a single right answer to our questions, we are most interested in understanding your process and strategies. Please feel free to share information as it comes to mind. Before we get started, do you have any questions for me?

SECTION I: GENERAL DESCRIPTION OF WELL-CHILD VISITS

1. To begin, I would like you to think about a well-child office visit with a toddler (ages 2-3). Can you please briefly walk me through your process for a routine visit?

2. I know that you have limited time during an office visit. Can you tell me about your priorities during a well-child visit with a toddler (ages 2-3)?
    a. What things are the most important for you to assess? (Why?)
    b. What things are less important for you to assess? (Why?)

SECTION II: SOCIAL-EMOTIONAL WELLBEING

Next, I would like you to think specifically about toddler's development.

3. Do you typically perform developmental surveillance (i.e., monitoring) or screening during routine well-child visits?
    a. If yes, how do you practice developmental surveillance or screening?
        i. What tools/resources do you use?
    b. If no, why not?

4. How would you define or describe what social-emotional wellbeing in a toddler is to a student or resident?

   a. How would you define or describe social-emotional challenges in a toddler to a student or resident?

To make sure we are on the same page, for this interview, we define social-emotional wellbeing as toddler's ability to manage emotions and behaviors in a manner that allows them to engage in age-appropriate behaviors like sharing, going to the park, participating in daily routines. The opposite of that is what we consider challenges. In other words, social-emotional wellbeing is how they are doing in terms of emotions and behavior.

5. Based on the definition we provided above, would you say that you conduct social-emotional surveillance (monitoring) or screening?

SECTION III: ASSESSING CHILD SOCIAL-EMOTIONAL DEVELOPMENT

[If the interviewee does NOT perform social-emotional surveillance (monitoring) or screening, skip to Section IV]

6. How do you identify toddlers who are experiencing a challenge with social-emotional development?

7. Do you typically perform social-emotional surveillance (i.e., monitoring) or screening during routine well-child visits?
   a. If yes, how do you practice social-emotional surveillance or screening?
      i. What tools/resources do you use?
   b. If no, why not?

8. What do you do if you are concerned about a child's social-emotional wellbeing?

9. Do you talk with parents/caregivers about everything that you observe related to their toddler's development, including social-emotional well-being, or do you only talk with parents/caregivers about some things that you observe during their visit? In other words, is there a threshold for what you bring up and when? (If Yes) Can you tell me more about that?
   a. What do you say to parents/caregivers when you have concerns about their child's social-emotional wellbeing?
      i. How do those conversations typically go?
   b. What strategies do you use if they do not share your concern or disagree with you?

10. What concerns do you have about identifying and/or labeling a child who needs extra behavioral or social-emotional support or resources?
   a. Is stigma a potential concern when identifying and/or labeling a toddler who need support?
      i. If so, in what context? (i.e., certain communities or for certain challenges or supports?)

11. We know that you acquire a lot of information during a single visit. How do you integrate all the information to inform your assessment of a toddler's social-emotional wellbeing and make decisions about what to do?

SECTION IV: ASSESSING CHILD DEVELOPMENT

[If the interviewee DOES perform social-emotional surveillance (monitoring) or screening, skip to Section V]

12. How do you identify toddlers who are experiencing a challenge with development?

13. What do you do if you are concerned about a child's development?

14. Do you talk with parents/caregivers about everything that you observe related to their toddler's development, or do you only talk with parents/caregivers about some things that you observe during their visit? In other words, is there a threshold for what you bring up and when? (If Yes) Can you tell me more about that?
    a. What do you say to parents/caregivers when you have concerns about their child's development?
        i. How do those conversations typically go?
    b. What strategies do you use if they do not share your concern or disagree with you?

15. What concerns do you have about identifying and/or labeling a child who needs extra developmental support or resources?
    a. Is stigma a potential concern when identifying and/or labeling a toddler who need support?
        i. If so, in what context? (i.e., certain communities or for certain challenges or supports?)

16. We know that you acquire a lot of information during a single visit. How do you integrate all the information to inform your assessment of a toddler's development and make decisions about what to do?

SECTION V: RESOURCES

Thank you for that discussion. Now, I would like to shift gears a little bit. This next set of questions asks about available tools and resources to help with developmental surveillance.

17. Are you familiar with the Centers for Disease Control and Prevention's (or CDC) *Learn the Signs. Act Early*. materials/resources?
    a. If yes, do you use any of these materials/resources during well-child visits? Which ones?

18. **[Show *Milestone Tracker* app slides]** One CDC resource is the *Milestone Tracker*, a mobile application for parents/caregivers. On your screen, you can see a snapshot of the app. The app's features include interactive milestone checklists for ages 2 months through 5 years, photos and videos to help parents/caregivers recognize milestones, tips

and activities for supporting early development, and appointment and developmental screening reminders. During visits, parents/caregivers can use the app to quickly share their child's progress with healthcare providers. Another feature of the app is that a parent/caregiver can email or show a summary of their child's progress to a provider for any milestones a child is missing or where they have concerns.

   a. I'd like you to please share your initial reactions to this feature and the app in general.
      i. Do you see any components that might catch your attention? If so, which components?
      ii. What would you do if a parent or caregiver shared this information with you?
   b. What impact do you think this might have on how you conduct developmental surveillance during well-child visits?
   c. How do you think this app would help you discuss a parent or caregiver's concerns about their child's development?
   d. Would you recommend this app to parents and caregivers? Why or why not?

19. **[Skip to 19.a. if interviewee mentions Bright Futures earlier in the interview]** Are you familiar with Bright Futures from the American Academy of Pediatrics (AAP)/Health Resources and Services Administration (HRSA)?
   a. If yes, do you use any Bright Futures resources during well-child visits? Which ones?
      i. How do you use the information from Bright Futures during visits?
      ii. Do you find it useful? Why or why not?
      How do you integrate Bright Futures with other sources of information about a child's developmental or social-emotional wellbeing?

20. The purpose of this interview was to gather information about how clinicians integrate social-emotional screening and surveillance in practice. Is there anything else you would like to share or that might be helpful for me to know before we wrap our interview today? Do you have any comments or anything else to add?

SECTION VI: SOCIODEMOGRAPHICS

Thank you for answering the questions so far. This last set of questions asks about you and your practice.

1. To which gender identity do you most identify?
   a. Woman
   b. Man
   c. Transgender man/trans man
   d. Transgender woman/trans woman
   e. Genderqueer
   f. Don't know
   g. Prefer not to answer

    h. Other (please specify):
2. Are you of Hispanic, Latino, or of Spanish origin?
    a. Yes
    b. No
3. How would you describe yourself?
    a. American Indian or Alaska Native
    b. Asian
    c. Black or African American
    d. Native Hawaiian or Other Pacific Islander
    e. White
    f. Other
4. How many years have you practiced medicine? _________
5. Practice type (select all that apply)
    a. Independent practice
    b. Hospital-affiliated practice
    c. Affiliated with university or medical school
    d. Federally Qualified Health Center (FQHC)
    e. Other (please specify)
6. Percent patients served (best estimate)
    a. % patients Medicaid ________
    b. % patients Medicare ________
    c. % patients private ________
    d. % patients self-pay or uninsured ________
7. What geographic area does your practice primarily serve?
    a. Urban
    b. Suburban
    c. Rural

CLOSING

Thank you so much, those are all of the questions that I have for you. Do you have any other questions or comments for me?
If I have any follow-up up questions in the future, would it be okay for me to re-contact you?

Thank you for taking the time to participate in the interview. I am going to stop the recording now. **[Stop recording interview]**

In terms of compensation, please provide your name, preferred email, and postal address. We will use this information to provide you with $100 virtual gift card in appreciation of your time, which will be emailed to you within 2 weeks.

Thank you again, and hope you have a wonderful day.