HHS Public Access
Author manuscript
Dev Psychobiol. Author manuscript; available in PMC 2020 July 08.

Published in final edited form as:
Dev Psychobiol. 2019 April ; 61(3): 444–464. doi:10.1002/dev.21831.

Automated sensing of daily activity: A new lens into development

Kaya de Barbaro
Department of Psychology, The University of Texas at Austin, Austin, Texas

Abstract

Rapidly maturing technologies for sensing and activity recognition can provide unprecedented access to the complex structure daily activity and interaction, promising new insight into the mechanisms by which experience shapes developmental outcomes. Motion data, autonomic activity, and “snippets” of audio and video recordings can be conveniently logged by wearable sensors (Lazer et al., 2009). Machine learning algorithms can process these signals into meaningful markers, from child and parent behavior to outcomes such as depression or teenage drinking. Theoretically motivated aspects of daily activity can be combined and synchronized to examine reciprocal effects between children’s behaviors and their environments or internal processes. Captured over longitudinal time, such data provide a new opportunity to study the processes by which individual differences emerge and stabilize. This paper introduces the reader to developments in sensing and activity recognition with implications for developmental phenomena across the lifespan, sketching a framework for leveraging mobile sensors for transactional analyses that bridge micro- and longitudinal- time-scales of development. It finishes by detailing resources and best practices to facilitate the next generation of developmentalists to contribute to this emerging area.

Keywords

daily interactions; dynamical systems theory; ecological validity; machine learning; wearable and mobile sensors

1 | INTRODUCTION

Across individuals, development proceeds with remarkable regularity. By and large, most children become experts in language, motor control, and social interaction, with recognizable milestones along the way. At the same time, there are meaningful differences between individuals. Some differences appear to be systematically related to past experiences, whereas others exist despite ostensible similarities in circumstances. We know variation in genes is related to variation in many outcomes (Plomin, DeFries, Knopik, & Neiderhiser, 2016), and at the same time, we know that even from the first stages of

Correspondence: Kaya de Barbaro, Department of Psychology, The University of Texas at Austin, Austin, TX., kaya@austin.utexas.edu.

CONFLICT OF INTEREST
The author has no conflicts of interest to declare.

embryonic development, the structure of an organism is not prespecified by the genes, but rather is determined by interactions between prior structure and what is present in the local cellular environment (Gottlieb, 1998; Stiles, 2008). Studies indicate that experiences across the lifespan matter for long term outcomes (Sroufe, Egeland, Carlson, & Collins, 2005), but at the same time, even extreme factors like trauma, abuse, and neglect do not affect all individuals in the same way.

It is no longer contentious that development is the product of complex interactions between organism-internal and organism-external factors. However, moving beyond debates about nature versus nurture to characterize “the way in which biology and experience work together throughout thick and thin” (Gottesman & Hanson, 2005; p. 263) continues to be a challenge for developmental science.

In this paper, we argue that the key to this puzzle lies in characterizing daily activity: the mundane day-in and- day out of what we perceive and do, who and what we interact with, and the internal states that occupy the minutes, hours, and days of our lives. Technological advances now make capturing and analyzing large volumes of daily activity occurring in natural environments possible. From wearable cameras and step trackers to smart homes and smart thread, sensors to capture daily activity are literally being woven into the fabrics of our daily lives. The widespread adoption and presence of these sensors, paired with the coming of age of powerful algorithms to automatically extract meaningful activities from raw sensor data, allows unprecedented access to dynamics of activity in the daily contexts in which development happens. Captured over longitudinal time, such data provide a radical new opportunity to tackle the oldest puzzles of development; namely, the mechanisms by which experience shapes developmental outcomes.

1.1 | Sensing daily activity: A new lens into development

Broadly, wearable or mobile sensors are computer chips small enough to be carried or worn (Starner et al., 1997). Developments in micro-processing have led to increasingly smaller devices without compromising storage, battery life or computing power. Devices embedded with mobile sensors have ushered in an era of “ubiquitous” computing, referring to the fact that such devices can now be found everywhere: in our pockets, on our bodies, our homes and communities (Abowd & Mynatt, 2000; Weiser, 1991).

The use of mobile sensors to collect objective, ecologically valid data to gain insight into human behavior began in the late 1990s (Pentland, 2000). The pioneering studies of Sandy Pentland (Eagle & Pentland, 2006) used Bluetooth scans to log proximity between office workers’ cell phones. With simple measures of physical presence and distance between individuals, they were able to characterize numerous aspects of social behavior, from friendships to work habits and places of informal gathering. This paper was the first suggestion that engineering techniques could provide insight into the complexity of human interactions.

The initial efforts of this new “computational social science” (Lazer et al., 2009) were led by computer scientists and engineers. However, the field has matured such that it is now increasingly feasible and common for social scientists to be involved with and even lead mobile sensing studies. Indeed, it has been nearly 20 years since the first mobile sensing publications geared toward social scientists were published (Goodwin, Velicer, & Intille, 2008; Healey, 2000; Miller, 2012), with high-profile calls-to-action continuing to emerge (Harari, Müller, Aung, & Rentfrow, 2017; Schmid Mast, Gatica-Perez, Frauendorfer, Nguyen, & Choudhury, 2015; Timmons et al., 2017). Empirical findings from mobile sensing research are no longer limited to engineering and computer science journals, but now span a range of disciplinary perspectives, from clinical psychology to personality research (Ben-Zeev, Scherer et al., 2017; Chin, Goodwin, Vosoughi, Roy, & Naigles, 2017; Lathia, Sandstrom, Mascolo, & Rentfrow, 2017).

This paper will begin with a review and synthesis of existing research in psychology and developmental science focused on analyzing daily activity collected via mobile and wearable sensors. This work indicates the feasibility and unique potential of such data to contribute new theories and insights into developmental processes. Next, we introduce research from the field of ubiquitous computing. Advances in this field provide new opportunities to capture daily activity across the nested layers of the developmental system, with implications for a range of phenomena (outlined in Table 1). In the fourth section ("Pushing the envelope") we sketch a framework for leveraging mobile sensors to access the reciprocal transactions between participants' behaviors and their internal, social and physical environments as well as the accumulation of these feedback cycles over longitudinal time. This is the holy grail of dynamical systems approaches to development (Thelen & Smith, 1994) and represents the ultimate promise of what mobile sensors could bring to developmental science research. The fifth and final section of the paper provides practical advice for developmental scientists wishing to leverage these tools within their own work, including advice for collaborations and training for students.

2 | DAILY ACTIVITY: A UNIQUE CONTRIBUTION

The study of free-flowing behavior is not common in psychology research, which has been dubbed the science of "self-reports and finger-movements" (Baumeister, Vohs, & Funder, 2007). The psychological is largely considered to reside within the skull (Hutchins, 1995), and activity has generally been treated as an indirect way to access the internal workings of the mind. Developmental science is unique within psychology for its long-standing focus on observations of interactions. This is likely due to the difficulties in administering typical standardized measures to preverbal children, as well as the consideration of the impacts of social partners and embodied interactions to developmental outcomes.

Historically, however, developmental research has rarely involved studies of natural daily activity (Bronfenbrenner, 1977; Lee, Cole, Golenia, & Adolph, 2018). Video recordings of scripted laboratory interactions are the data of choice in observational studies across development (Adolph, 2016), owing to the fact that they provide high fidelity access to free-flowing behaviors that can be endlessly observed and systematically characterized. Scripted interactions allow researchers to elicit specific phenomena that would only rarely occur during the short time of a laboratory visit. Video records thus function as a microscope for human activity (Goguen, 1997), allowing researchers to carefully and methodically capture both emergent qualities of interactions as well as the precise micro-dynamics of children's activity and interactions (Bakeman & Gottman, 1997). The vast benefits of video in laboratory settings are such that it is irreplaceable in developmental research. However, leveraging mobile and wearable sensors to capture activity in daily life can provide unique access to structure and variability unlikely to be captured or present in the laboratory.

2.1 | Existing research

A rapidly growing body of research points to the value of characterizing truly ecological daily activity within developmental science. This research indicates that daily activity, far from being simply “noisy” or “random,” has consequences for child outcomes across domains.

Wearable sensors have provided key insights into the relationship between sensorimotor development and early cognition. Embodied cognition perspectives emphasize the fundamental link between perception and action (Gibson, 1988). The particulars of a learner’s body, and the actions available to a learner, determine what is perceivable and therefore, what is learnable (Kretch, Franchak, & Adolph, 2014; Smith & Gasser, 2005). In turn, these changing experiences can elicit novel inputs for social and cognitive development (Campos et al., 2000; Soska, Adolph, & Johnson, 2010). Sensors that can capture the rapid changes in sensorimotor experience across children’s first months and years of life have provided new insight into these early processes.

Much of this work stems from the use of lightweight cameras positioned on children’s foreheads to capture “what’s in view” for infants and toddlers (Borjon et al., 2018). Using such techniques, Fausey and colleagues (Fausey, Jayaraman, & Smith, 2016) were able to record six hours of daily visual experience of individual infant participants. Analyses of the footage indicated that the distribution of faces in infants’ view is highest within the first months of life and then decreases in the first 2 years, whereas the distribution of hands in infants’ view shows the reverse trajectory. These striking regularities have implications for what is learned and learnable across the first 2 years of life (Smith, Jayaraman, Clerkin, & Yu, 2018). For example, the frequency with which particular objects appeared in view aligned with early word learning, suggesting the importance of these daily exposures for learning outcomes (Clerkin, Hart, Rehg, Yu, & Smith, 2017).

A wide variety of regularities in daily exposure matter for children’s outcomes. Traditional measures indicate that the presence of household chaos (Vernon-Feagans, Willoughby, & Garrett-Peters, 2016), TV and electronics (Christakis et al., 2009), books (Raikes et al., 2006) and neighborhood institutions (Leventhal & Brooks-Gunn, 2000) all have implications for cognitive development. However, questionnaire measures and even objective scans of the home (Caldwell & Bradley, 1984) are an imperfect proxy of lived experience: just because there are books in the home does not mean a child is engaging with them. Emerging sensor research has begun linking some dynamic measures of experience to developmental outcomes. For example, research with teenagers has found that objective GPS-based (but not subjective) measures of physical proximity to liquor stores are associated with an increased likelihood of underage drinking (Byrnes et al., 2017). Such objective measures of the structure of daily experience will likely continue to reveal new insights into developmental trajectories.

Perhaps the most extensive focus of developmental sensing research is on children's early language environments. In the Human Speechome Project, Deb Roy famously outfitted his own home with enough cameras to capture a near-continuous record of his own child's experiences from birth to 3 years of age. A combination of automated analyses and extensive human annotation was used to create a massive dataset characterizing his child's language exposures: including the words, physical locations, and motion of multiple speakers in the house. Analyses revealed that the distinctness of the context in which a word was uttered–from its physical location to the other words around it–helped to predict the age of production of individual words (Roy, Frank, DeCamp, Miller, & Roy, 2015). However, the vast resources required to create this dataset are arguably yet to outweigh its overall contributions to the field. Mobile and wearable sensors are a creative counterpoint to such a relatively heavy-weight setup. By virtue of being able to be placed on the body, a single camera or audio recorder can be used in place of an integrated multi-camera system to capture various aspects of a child's daily exposures.

In this vein, a sizeable body of research has emerged using the LENA, a lightweight wearable audio recorder developed to study language development (Zimmerman, 2009). The LENA includes specialized algorithms developed to identify onset and offset of individual parent and child vocalizations, which are then used to identify patterns of contingent speech (i.e., conversations) between children and their caregivers. Research with the LENA has demonstrated that, over and above parental speech, it is the volume of *conversations* that predicts children's vocabulary size and school achievement (Chin et al., 2017; Weisleder & Fernald, 2013). The LENA has been used to study language environments across many populations (reviewed by Wang, Chen et al., 2017). While additional algorithms for processing ambulatory audio recordings continue to be developed (e.g., Ludwig, 2018), developmentalists have also begun to supplement LENA'S automated detection algorithms with human coding to pursue additional questions, such as the impacts of parent-ese speech style (Ramírez-Esparza, García-Sierra, & Kuhl, 2014) or how speech varies according to activity (Soderstrom & Wittebolle, 2013).

Audio data from wearable devices has also been successfully used to characterize aspects of caregiving quality and family interactions across the lifespan. Daily conflict behaviors between parents and children have been associated with parental trait emotionality (Slatcher & Trentacosta, 2012) and depression symptoms (Slatcher & Trentacosta, 2011). More daily family conflict has also been associated with children's lower cortisol at awakening and flatter diurnal slope over the day, indicating more maladaptive responses (Slatcher & Robles, 2012). Conversely, more daily maternal responsiveness, characterized by warmth, emotional support, and expressions of pride, is associated with better immune response in teenagers with asthma (Tobin et al., 2015). We are unaware of any work that compares objective caregiving measures of daily activity with gold-standard lab situations, or that examines variability in caregiving within or across days and its potential associations with child outcomes; both key directions for future research. However, these studies indicate the feasibility and predictive validity of objective measures of daily family interactions across a wide age range.

Physical activity is another readily available marker of daily activity with implications for many domains of development (Pellegrini & Smith, 1998). In a pioneering series of studies in the 1980s, Buss and colleagues compared the amount of free-flowing physical movement as captured by an early motion sensor, the actometer, in a longitudinal sample of 3-, 4-, and 7-year-olds. Findings indicated strong stability in children's activity across time, as well as strong correlations with concurrent and future measures of children's personality measures. More active children were independently judged by teachers as more outgoing; additionally they were more likely to be resistant to adult demands (Buss, Block, & Block, 1980). A follow-up study indicated that mothers of children identified as more active via actometer were more likely to be impatient, hostile, and physically intrusive in their interactions with their children (Buss, 1981). The longitudinal stability of these measures, and their strong associations with personality and parenting measures is striking. These associations lead us back to the main puzzles of developmental science: how do relationships between children's daily experiences and their outcomes become established? While child activity may lead to parental aggression, it is also feasible that the relationship is bidirectional, that is, that parental aggression could also contribute to child hyperactivity, perhaps mediated by difficulties regulating emotions. Additionally, while the observed correlations were strong, a complete developmental account would provide some insight into both lawful continuity and lawful change (Sroufe et al., 2005), or why certain children showed stable patterns of activity across time, whereas others varied across these three timepoints. Advances in sensing mean that future research could begin to disentangle parent and child characteristics to quantify how they interact and influence one another over time, a proposition we will return to below.

2.2 | Existing research: Lessons learned

Existing research indicates the feasibility and value of sensing daily activity to provide insight into development. Below, we synthesize this research to highlight four potential contributions of these emerging techniques.

First, while some research indicates a strong correspondence between measures derived from structured laboratory paradigms and unstructured home activity, other research indicates key differences between these measures. For example, the total volume of speech in structured interactions is strongly ($r = 0.7$) predicted by speech in unstructured interactions (Tamis-LeMonda, Kuchirko, Luo, Escobar, & Bornstein, 2017). However, as reported by the NICHD Early Child Care Research Network, sensitivity in free play and unstructured home interactions showed only moderate correlations ($r = 0.4$), and only sensitivity in unstructured interactions showed a main effect on attachment outcomes (NICHD Early Child Care Research Network, 1997). Additionally, negative impacts of caregivers exhibiting less sensitive care were shown to be reduced when children spent more time with their caregivers, suggesting that both the quality and *quantity* of care are distinct and interacting factors relevant for understanding outcomes (Ibid).

Second, other aspects of natural activity are simply not present in structured laboratory interactions. These include extended periods of non-interaction between parents and their children (Tamis-LeMonda et al., 2017), the propensity to walk in circles rather than in a straight line (Lee et al., 2018), or the low proportion of faces present in the view of toddlers (Jayaraman, Fausey, & Smith, 2015). These common but seldom observed-in-laboratory features of experience may have critical developmental functions (Smith et al., 2018): children may learn to gaze follow not from looking at eyes but looking at hands (Deak, Krasno, Triesch, Lewis, & Sepeta, 2014); extended silences between interactions may be critical for consolidation and marking of semantic boundaries (Tamis-LeMonda et al., 2017). Sensing daily activity can provide new access to these previously unmeasured aspects of children's experiences.

Third, sensors allow for the collection of much larger volumes of data than would be possible to obtain in the lab. This means that we have a chance to capture activities that occur rarely but may have developmental import (VanDam, 2016). Additionally, large volumes of data allow researchers to access temporal structure and variability of activity, which can predict outcomes over and above traditional summary measures (Li & Lansford, 2018; Sherman, Mumford, & Schnyer, 2015).

Finally, the possibility to capture activity over long periods of time means sensors can provide access to phenomena which may take hours, days, or weeks to unfold, such as the cumulative impact of multiple extended tantrums on parental responses to their children, and the resulting impacts on children's behavior. The ability to capture such transactive reciprocal processes across time holds perhaps the greatest promise of sensors for developmental science, a proposition we explore further in Section 4 ("The future of mobile sensing in developmental science") following a review of ongoing research and developments from the ubiquitous computing community.

3 | ADVANCES IN SENSING AND UBIQUITOUS COMPUTING

Developmental science stands to benefit from ongoing developments from the field of ubiquitous computing. These include technological advances for capturing novel datastreams as well as powerful "activity recognition" algorithms which transform raw sensor data into meaningful markers of activity. In the next section, we provide a primer on activity recognition followed by an overview of mobile sensing research within ubiquitous computing, focusing on relatively stable solutions that hold the most potential for immediate incorporation into research. We organize this research according to its potential to provide access to nested levels of the developmental system, considering aspects of individual psychobiological activity, interpersonal activity, and daily ecological contexts.

3.1 | Introduction to activity recognition

Much of raw sensor data is uninterpretable or meaningless to humans: a log of physical changes in patterns of light, motion, or pressure. Activity recognition is the process by which these signals are transformed into markers of meaningful activity, such as steps or changes in affect (Bulling, Blanke, & Schiele, 2014). The basis of activity recognition algorithms is to identify distinctive physical signatures of meaningful activities, which, once learned by a model, can be used to predict new instances of behavior.

These algorithms are developed via a "training" process, during which labels of identified behaviors or activities ("outputs", known in the field as "ground truth") are paired with raw sensor data (the "inputs"). Training consists of the learning of optimal "betas" for predicting ground truth labels from sensor inputs, akin to learning the weights on a regression (indeed, regressions are a commonly used machine learning model). Via this process, smart watches, for example, use changes in light reflected at the wrist to detect changes in the blood flow corresponding to heart beats. Similarly, the algorithms in a Fitbit can identify steps (and distinguish them from bike riding or driving in a car) via characteristic patterns in motion, using features such as the typical speed and direction of acceleration of the arm as it swings with each step. For research grounded in theory, an important question is *whether* and *how reliably* raw sensor data can be translated into meaningful markers of individual, interpersonal, and ecological activity.

3.2 | Sensing individual activity

3.2.1 | Physiological activity—A variety of wearable sensors can capture continuous markers of autonomic nervous system activity while participants are fully ambulatory, including markers of heartbeat, skin conductance, respiration, and temperature. While typically requiring direct contact with the skin, the current generation of ambulatory physiological sensors are much more comfortable than older generations, such that it is feasible to collect 11+ hours/day of usable data for upwards of 4 weeks (Rahman et al., 2014). The quality of ambulatory psychophysiological recordings is highly variable. Electrode-based and chest-strap monitors should be used when high accuracy data are required, as it is known that wrist-worn smart-bands provide relatively low-quality indicators of heart rate (Wang, Blackburn et al., 2017). Most wrist-based sensors do not provide the resolution necessary to reliably detect short-lasting or subtle physiological changes, or calculate high-resolution physiological indices such as heart rate variability (i.e., vagal tone). Multiple companies now also provide physiological monitoring solutions that include wearable amplifiers (biopac.com; mindwaretech.com). While larger and more cumbersome, these amplified setups allow for accurate identification of complex markers of blood flow such as PEP (Cacioppo, Uchino, & Berntson, 1994).

While typically developed for adults, our pilot testing of a variety of commercially available and research-grade devices–both chest-worn straps and wrist-bands (worn on children's ankles)–indicates that they appear to provide valid data with young subjects, although systematic testing of these claims is warranted.

A major issue with even the highest quality physiological sensors is that the magnitude of effects of physical activity typically overwhelms the subtle physiological signals associated with affect and emotion. Thus, it is rare for these signals to be directly used on their own as markers of stress, for example. One common solution to this is to restrict analyses of ANS data to periods of reduced motion (e.g., Hovsepian et al., 2015). Interestingly, this may be less so an issue with infants in particular, as motor activity appears to be a more precisely valid indicator of arousal level for pre-locomotor infants (Wass, de Barbara, & Clackson, 2015) relative to older (self-locomoting) participants.

3.2.2 | Affect—Active-sensing techniques which require participants to actively log data or responses are the gold standard for ambulatory assessment of affect and internal experiences (Ebner-Priemer & Trull, 2009; Mehl & Conner, 2011). Repeated surveying of subjective experience, also referred to as ecological momentary assessment (EMA), was pioneered in the 1980s by Csikszentmihalyi and colleagues (Csikczentmihalyi & Larson, 1983; Shiffman, Stone, & Hufford, 2008) and is now a well-established method within psychology (Bolger & Laurenceau, 2013). Today, EMAs are typically administered on participants' mobile phones via a variety of open-source and commercially available applications (Kubiak & Krog, 2011; Wilhelm, Perrez, & Pawlik, 2012). Many common mood assessments are based on the PANAS (Watson, Clark, &Tellegen, 1988) or similar affect scales (Peeters, Berkhof, Delespaul, Rottenberg, & Nicolson, 2006), and it is common in EMA studies to administer approximately 30 items 4–6 times a day for periods ranging from 1 week to 1 month (as reviewed by aan het Rot, Hogenelst, & Schoevers, 2012).

Within ubiquitous computing research, novel designs for collection of EMA data are being explored, including picture selection (Poliak, Adams, & Gay, 2011), lock-screen (Zhang, Pina, & Fogarty, 2016) and smart-watch designs (Intille, Haynes, Maniar, Ponnada, & Manjourides, 2016) as well as methods to record images, audio, or video in lieu of survey-based responses. The latter can provide much higher dimensional data and may be less burdensome than traditional methods (see, e.g., the use of images in place of food diaries; Cordeiro, Bales, Cherry, & Fogarty, 2015). With such novel methods, certain caution is required when addressing data validity (Chan et al., 2018). While EMA has historically not been subject to the same standards as traditional questionnaires within psychology, experts in EMA design now emphasize these concerns (Stone, 2017; Stone & Shiftman, 2002), recommending, for example, that a minimum of three items be used to address individual constructs when examining within-person across-time variation (Bolger & Laurenceau, 2013; Cranford et al., 2006).

Ambulatory affect has also been characterized through a variety of "passive" methods where sensor data are logged without active intervention by participants. Passive-sensing affect models have been developed using physiological, audio, and motion data (Sharma & Gedeon, 2012), as well as social media data (De Choudhury, Counts, & Gamon, 2012). An array of algorithms that map physiological parameters to affective states have been published in the computer science literature (Guo et al., 2013; Kim, Bang, & Kim, 2004). However, given the complex theoretical and conceptual issues surrounding affect (Frijda, 1986), combined with the relatively lax standards of validation within the ubiquitous computing community, passive affect sensing is of variable quality and many published models would likely not pass more thorough quality control standards. This is not surprising given that much laboratory research indicates that physiological signals alone cannot differentiate a wide spectrum of emotions (Wilhelm & Grossman, 2010). However, some promising models distinguishing a more restricted subset of affective states have been developed. The best combine physiological and motion signals, such as a model developed by the MD2K consortium which uses respiration, heart rate and motion to predict stress levels (Hovsepian et al., 2015). Note that this model was designed with both social scientists and engineers, is based on training data from a widely recognized stressor (Trier social stress task, Kirschbaum, Pirke, & Hellhammer, 1993), and has been validated with concurrent EMA in a completely naturalistic situation. We emphasize such critical quality control measures further in the practical considerations section.

The use of high-fidelity audio or linguistic content is perhaps the most promising avenue for automated assessment of affect. Ambulatory audio recordings have been used to identify a variety of paralinguistic vocalizations, such as laughing or crying, in both adults and children (Rao, Kim, Clements, Rozga, & Messinger, 2014; Yatani & Truong, 2012). Acoustic features have also been used to identify affect and tone within speech, including identification of basic emotions (Basu, 2002; Rachuri et al., 2010), stress (Lu et al., 2012) and emotional arousal (Juslin & Scherer, 2005) as well as clinical symptoms, such as depressive (i.e., flat) tone (Moore II, Clements, Peifer, & Weisser, 2008; Taguchi et al., 2018). While promising, as above, many of these models have been developed with small samples and require additional validation to ensure their robustness. Given the importance of affect for social and emotional developmental outcomes as well as, e.g., assessments of well-being this is a key area for future research.

Analyses of textual content, known as natural language processing (NLP), have also been successfully used to identify mood as well as other individual state and trait characteristics. NLP methods have been widely implemented in ambulatory psychological research with content transcribed from audio snippets (Mehl, 2017, as detailed above). More recently, NLP techniques have been used on data collected from social media sites or messaging platforms (De Choudhury et al., 2012). Both audio snippet and social media data have been used successfully to assess mood and affect (De Choudhury et al., 2012; Gill, French, Gergle, & Oberlander, 2008), depressive symptoms (De Choudhury, Gamon, Counts, & Horvitz, 2013; Rude, Gortner, & Pennebaker, 2004), and suicidality (De Choudhury, Kiciman, Dredze, Coppersmith, & Kumar, 2016; Stirman & Pennebaker, 2001), among other states. While promising, there are some challenges with obtaining sufficient text content for NLP analyses. In particular, while audio snippets are now feasible to collect at large scale via cell-phones, automated transcription of ambulatory recordings is still not feasible, meaning these recordings require time-consuming human transcription (Mehl, 2017). Additionally, many people do not post frequently enough on social media to derive multiple daily markers of affect, and they may be unwilling to share content from higher-density but more private chat or messaging platforms, meaning that researchers are often relegated to use of platforms with relatively volumes of data. Advances in automated audio transcription (LeCun, Bengio, & Hinton, 2015) as well as privacy-preserving techniques may increase the feasibility of these techniques for large-scale high-density research.

3.2.3 | Embodied activity—Reports of activity can be collected via active (via calendars or EMAs) or passive methods. The most common passive activity recognition models use inertial sensors (i.e., accelerometer and gyrometer, which measure linear and angular acceleration of motion, respectively) relative position sensors (barometric pressure, magnetometer, geolocation data) and ambient light and audio sensors. These sensors are present in all off-the-shelf smart phones and many inexpensive wearables (Harari et al., 2016; Lara & Labrador, 2013).

The most robust and commonly used markers of activity in both social and computer science are related to locomotion and sleep (Bussmann & Ebner-Priemer, 2012). Movement can be accurately calculated via GPS by simply summing travelled distance, or via algorithms that consider speed and patterns of motion to summarize type or quantity of motion. Movement patterns have been used in a number of large-scale longitudinal studies (Brocklebank, Falconer, Page, Perry, & Cooper, 2015; McConnell et al., 2017).

Sleep is another common and robustly detected activity used in various large-scale studies (e.g., Robillard et al., 2015). Actigraphy cannot reliably detect certain subtle sleep measures such as sleep cycles, night wakings, or sleep efficiency. (Sadeh, 2011). However, when compared with gold-standard polysomnography measures, sleep-wake cycles and total sleep time show over 90% agreement in both children and adults (Sadeh, Sharkey, & Carskadon, 1994; Sitnick, Goodlin-Jones, & Anders, 2008). Another developmental^ relevant and highly robust measure is mobile phone use, commonly logged by apps used in ubiquitous computing research (Harari et al., 2017).

In smaller scale studies, wearable motion sensors have also been used to identify a wide variety of activities of daily living, for example, distinguishing between eating, drinking, reading, watching TV, and vacuuming (Lara & Labrador, 2013; Thomaz, Essa, & Abowd, 2015), as well as subtle behaviors such as aggressiveness or self-injury (Ploetz, Hammerla et al., 2012). Indeed, two studies have developed models that can distinguish children's locomotor states, including rolling, walking, sitting, crawling, and climbing (Busser, Ott, Van Lummel, Uiterwaal, & Blank, 1997; Nam & Park, 2013). Generally, recognition of activity appears to be imminently possible. However, distinguishing between activities which are physically similar may require users to wear multiple sensors (Arif & Kattan, 2015). A powerful alternative is to combine sensors from multimodal sensing platforms, such as cellphones (as reviewed by Harari et al., 2017; Miller, 2012). For example, Campbell et al. (Wang, Chen et al., 2014) developed algorithms to infer likely studying and partying behaviors by combining multiple cellphone data streams, including ambient noise (loud vs. quiet), activity level (high vs. low), and geo-coded location (fraternity vs. library) via GPS data. For younger children, it may be possible to use a similar approach to access moments of focused attention, for example. Namely, cell phone use data combined with motion sensors embedded in a headband may be able to identify a steady focus of gaze while differentiating it from sleeping or focus on the phone itself. Similar markers have been used to identify quality of attention while driving (Tawari, Martin, & Trivedi, 2014).

There is already widespread interest in ambulatory markers of risk- and health-promoting behaviors (Bertz, Epstein, & Preston, 2018), of particular relevance for research in adolescence and early adulthood. The use of alcohol and other drugs of abuse can be captured via transdermal patches or other physiological signals (Simons, Wills, Emery, & Marks, 2015), including, for example, the use of wrist-worn sensors and respiration belts to detect smoking behaviors (Saleheen et al., 2015). Considering less obtrusive platforms, binge drinking has been detected via cell phone features such as changes in speed of typing (Bae, Chung, Ferreira, Dey, & Suffoletto, 2017).

Critically however, the validity of such less-common activity models must be carefully assessed as they are often developed with small samples using laboratory-based data. For these reasons they are best considered a proof of concept requiring thorough quality control and additional development prior to use in studies.

3.3 | Sensing interpersonal activity

Within the ambulatory assessment community in psychology, the gold-standard for objective markers of social interaction are qualitative ratings of audio snippets. In contrast, within the ubiquitous computing community, "social sensing" has primarily focused on markers of proximity and basic acoustic features of audio recordings.

Proximity (i.e., relative distance) between individuals can be detected through a variety of methods (Barrat, Cattuto, Tozzi, Vanhems, & Voirin, 2014). For example, radio-frequency signals and infrared beams can be used to can detect close contact (1–5 feet) as well as directionality, which can be used to detect face-to-face contact between individuals (Olguin et al., 2009). In contrast, Bluetooth radio waves are relatively far-ranging but provide relatively coarse spatial resolution (e.g., distinguishing whether participants are 1–2, 3–4 vs. 5–10 feet from one another) (Boonstra, Larsen, Townsend, & Christensen, 2017; Montanari, Nawaz, Mascolo, & Sailer, 2017; Osmani, Carreras, Matic, & Saar, 2014). However, Bluetooth is an omnidirectional signal and thus cannot be used to detect face-to-face contact. While it is technically possible to obtain Bluetooth scans every 5–20 s, on most cell phones it is only feasible to log scans once every five minutes or so due to battery constraints (Boonstra et al., 2017).

Beyond proximity, algorithms have been developed to detect the presence or absence of conversation from audio snippets as a marker of social interaction (Berke, Choudhury, Ali, & Rabbi, 2011; Wyatt, Choudhury, & Bilmes, 2007; Wyatt, Choudhury, Bilmes, & Kitts, 2011). For behavioral scientists used to coding caregiver warmth or sensitivity, these simple audio and proximity features fall painfully short of the complex qualities of a parent-child bond or peer relationships. However, it is worth considering the insights into daily interactions they may be able to provide.

Algorithms to detect qualitative features of interactions from ambulatory audio recordings have not yet been developed. However, data from laboratory settings suggest this is feasible. Models using features of vocal arousal have been used to automatically identify empathy and conflict as well as distinguish between blame and acceptance in adult-adult problem-solving interactions (Black et al., 2013; Imel et al., 2014; Weusthoff, Baucom, & Hahlweg, 2013). This is another area where collaborations between developmentalists and engineers could provide novel contributions.

3.4 | Sensing the broader ecology

Sensors also hold possibility for capturing aspects of the broader contexts of daily activity. For example, specialized sensors can measure chemical and pollutant exposure (Mead et al., 2013). GPS can be combined with publicly available data to characterize neighborhood quality scores or momentary environmental exposures (Chaix et al., 2013), as well as likely activities at locations such as libraries, parks, museums, or fast-food establishments (Hariharan, Krumm, & Horvitz, 2005).

Basic characteristics of the environment can also be automatically determined via analysis of audio snippets or sporadic images (Hodges et al., 2006), including whether an individual is near a road, on a bus, in a library or in a lecture (Eronen et al., 2006; Pirsiavash & Ramanan, 2012).

Finally, a number of techniques have been developed for recognition of household activities with static sensors placed in the environment, leveraging Bluetooth and Infrared signals (Cook, Duncan, Sprint, & Fritz, 2018), or water and electricity use by appliances (Froehlich et al., 2009; Gupta, Reynolds, & Patel, 2010; Olguin, Gloor, & Pentland, 2009). These techniques could be used to unobtrusively capture a number of developmentally interesting characteristics, from the number of people in the home to the timing and regularity of meals. Out-of-the-box commercial options for such sensors are not yet widely available but active research in this area suggests they may soon be, as we further discuss in Section 5 (“Practical considerations”).

3.5 | Summary and challenges

The possibilities for sensing ambulatory activity are vast, and novel technologies continue to be developed. We summarize opportunities to leverage these techniques to study developmental questions in Table 1, which pairs developmental constructs and outcomes with theoretically relevant daily activities as well as potential sensing solutions for accessing those activities. Admittedly, some of the links are somewhat tenuous. To take full advantage of the promise of mobile sensors, developmentalists must collaborate with engineers and computer scientists to obtain valid and high-quality markers of daily activity. Translating traditional constructs into automatically sense-able activity will require creativity, as well as some openness to considering simpler physical analogues of subtle behaviors, such as in the case of focused attention or caregiver sensitivity. While sensed activities may not always map directly onto known laboratory behaviors or constructs, they may still provide valuable insight into the developmental process. Ultimately, this is an area with many avenues to explore.

4 | PUSHING THE ENVELOPE: THE FUTURE OF MOBILE SENSING IN DEVELOPMENTAL SCIENCE

Dynamical systems theories emphasize that development is an accumulation of decentralized and local interactions occuring in real time (Smith & Thelen, 2003, p. 343). Factors internal and external to the individual–from genetic, to neural, behavioral and environmental–are completely bidirectional over the course of individual development (Gottlieb, 2000, p. 96). The complexity and scale of this system of reciprocal influences means that it has been all but impossible to objectively characterize, leaving developmental processes largely obscured. However, advances in sensing mean we could now collect multiple precisely time-locked and synchronized dimensions of organism-internal and -external factors repeatedly and at scale. Such data will allow developmentalists to begin to disentangle complex reciprocal interactions between these factors and quantify how they drive developmental trajectories, a truly radical proposition for developmental science.

4.1 | An ecological approach: Capturing synchronized activity across the developmental system

Today, it is possible to collect multiple synchronized datastreams that capture not just the actions of a child, but simultaneously, a vast array of possible determinants of their actions: from their perceived environments and internal states, to aspects of caregiving behaviors, and even ecological factors such as household chaos or access to museums.

Datasets which can speak to this ecological approach are relatively rare. In the most impressive example to date, Andrew Campbell and colleagues (Wang, Chen et al., 2014) examined how first-year students' activity over the course of a semester would predict their end-of-year academic performance as well as daily measures of self-reported stress and mental health symptoms. Students' phones were outfitted with an app that collected widely available sensor datastreams, such as screen unlockings, motion patterns, light, and GPS data, as well as snippets of ambient sound. Algorithms were developed to transform these raw datastreams to access various aspects of students' activity, from internal states such as mood and fatigue, to daily care activities, such as time spent sleeping and exercising, and proxies for social interactions and studying. Analyses revealed that while well-being indicators and self-care activities begin high at the start of the semester, there was a reliable drop leading up to exam period (Wang, Chen et al., 2014). Additionally, changes in early versus late semester activities successfully predicted students end of year GPA (Wang, Harari, Hao, Zhou, & Campbell, 2015; Wang, Hartman et al., 2017). Within the mobile sensing community, there is a growing interest in such comprehensive ecological approaches, with efforts to predict and intervene upon physical and mental health outcomes such as psychosis onset and failures in smoking cessation (Ben-Zeev, Brian et al., 2017; Chatterjee et al.., 2016; Insel, 2017).

However, these efforts have not yet leveraged the full potential of these datasets for gaining insight into developmental processes. To do so, analyses of such comprehensive ecological datasets need to provide interpretable access to theoretically meaningful mechanisms of stability and change across longitudinal time. At first blush, this is at odds with typical computer science and engineering approaches which have historically led the majority of mobile sensing efforts. In these fields, the goal is often simply accurate prediction of outcomes rather than insight into mechanisms. This can lead to a "kitchen sink" approach– where sensor streams are selected only by virtue of being available on a given platform, or, even when theoretically relevant activities are considered, little attention is paid to how these may influence one another over time.

Additionally, the models used to analyze these data are often uninterpretable. Today, data science is increasingly "hands off": models are provided raw or lightly preprocessed data as input and left to independently discover regularities in the data that can be used to predict outcomes of interest. These algorithms are powerful: recent innovations in network-based modeling have found solutions to problems that have been considered unsolvable for decades (LeCun et al., 2015). However, there are tradeoffs between computational power and interpretability. At the heart of the issue is the fact that the more parameters included in a model, and the more complex the transformations between inputs and outputs, the more powerful the computational advantage for identifying patterns in the data. However, these same features obscure the relations between inputs and outputs, making the models themselves less interpretable. From an engineering standpoint, when the goal is to solve a problem, be it to predict the weather or identify a face within a crowd, model interpretability is not required (Breiman, 2001). Similarly, within developmental science, if the goal is intervention, including prediction of children who may be at greatest risk or identification of who may benefit most from a certain intervention, such complex models may be the direction of choice.

However, if the goal of developmental science is to learn *why and how* outcomes emerge, these “black box” solutions are problematic (but see, e.g., Bongard & Lipson, 2007). Insight into developmental processes will require analyses that provide interpretable and meaningful descriptions about what matters for development.

4.2 | Characterizing transactional dynamics within the developmental system to bridge micro and longitudinal timescales

Sensor data provide a novel opportunity to characterize and quantify the transactional processes by which daily activity accumulates to drive longitudinal trajectories.

Current gold-standard transactional analyses typically examine lagged temporal relationships across data collected at discrete time points, often months apart. We know, however, that individuals dynamically affect one another at much shorter timescales, minute-by-minute, hour-by-hour, and day-by-day (Granic, Hollenstein, & Lichtwarck-Aschoff, 2016). Sensor data are precisely time stamped, meaning that they could be leveraged to study the micro-dynamic transactions between factors driving developmental outcomes. For example, real-time child activity data could be synchronized with audio recordings characterizing the content or quality of parent-child interactions to examine whether variation in a child’s activity across the day predicts subsequent changes in parenting quality or vice versa, or, as may be expected in a complex bidirectional system: both, with the possibility of feedback processes which amplify initial differences over time. This is critical as recent research suggests, for example, that children with more reactive temperaments may be more sensitive to parenting practices (Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2007) while also being more likely to stress and overwhelm their parents (Belsky, 1984), thereby potentially exacerbating their own early biological predispositions. Similar approaches could be used to examine reciprocal processes theorized to drive individual functioning, such as proposed feedforward relationships between difficulties focusing and household chaos (Sroufe, 2012), or mood and cellphone use (Twenge, 2017).

We know of two LENA studies that have begun to move in this direction. In the first, Warlaumont and colleagues observed that parents of children with and without autism were more likely to respond to children’s vocalizations if they were speech-like rather than non-speech like (e.g., laughter, coughing, etc.). In turn, children’s vocalizations were more likely to be speech like if parents had contingently responded to their previous utterance (Warlaumont, Richards, Gilkerson, & Oiler, 2014). These data suggest a “snowballing” feedforward system where more contingent responding begets more speechlike vocalizations, begetting more contingent responding. Related analyses by Gordon Ramsey have indicated how such feedback loops appear to drive longitudinal differences between vocalization rates of children at high-risk of autism relative to typically developing controls (Ramsay, Bailey, & Ghai, 2016). However, in the first study, the compounded effects of these micro-transactions were not examined at the longitudinal timescale; in the second, the resolution of the analytic technique did not allow for examination of the micro-dynamics. Together though, these studies indicate the potential for transactional analyses that bridge micro- and longitudinal timescales.

Empirically characterizing such developmental trajectories is not simply a matter of searching for traditional (A->B) causal relationships. In accord with systems thinkers across disciplines (Forster, 2002; Isabela Granic & Patterson, 2006; Hutchins, 1995; Oyama, 1985/2000; Taylor, 2000), success in characterizing such complex dynamic system will require a shift in empirical goals as well as models of causality. In particular, rather than searching to identify the direction of causality between characteristics of parent and child behaviors, for example, the goal of a DST approach may be to identify factors which differentiate dyads who frequently lock into such amplifying or feedforward patterns from those who do not. Alternatively, it may be possible to identify conditions which lead to a shift in the typical trajectory of an interaction, such as what might lead a troubled dyad to adaptively respond, or what might destabilize typically adaptive functioning.

Previous efforts with high-density data indicate that there are few if any “off the shelf” analyses, and no single analytic tool that can characterize the dense, multi-modal dynamics of interaction (Gnisci, Bakeman,&Quera, 2008). However, methods developed to study the micro-dynamics of activity in laboratory settings are readily utilized on daily sensor data (Gottman & Roy, 1990). Additionally, methods papers by Granic and colleagues and others (Granic & Hollenstein, 2003; Granic et al., 2016; de Barbara, Johnson, Forster, & Deak, 2013) detail the discovery and quantification of theoretically meaningful structure in high-density multimodal repeated measures data with consideration of longitudinal outcomes. Finally, sophisticated tools have been developed for longitudinal analyses of high-density repeated-measures data (Bolger & Laurenceau, 2013; Kirn-Spoon & Grimm, 2016; Krull, Cheong, Fritz, & MacKinnon, 2016). A combination of such analysis techniques will likely be required to quantitively characterize complex longitudinal trajectories.

Sensor data provide an opportunity to access repeated measures of objective high-density daily activity at a large scale. This proposition is at once exciting and terrifying: exciting because we can now collect unprecedented datasets to investigate complex developmental trajectories, and terrifying because analyses of such complex processes is largely uncharted territory.

5 | PRACTICAL CONSIDERATIONS

In the final section of the paper, we lay out a number of practical considerations for developmental scientists interested in incorporating mobile sensors into their research. These considerations fall into three categories: collection of raw data, translation of raw data into markers of behavior, and analyses of behavioral markers. We finish with a section on collaborations and advice for training the next generation of students to skillfully leverage sensing technology to advance the future of developmental science.

5.1 | Collection of raw data

The main considerations for collection of raw sensor data include form factor and placement, commercial versus bespoke solutions, single versus multi-system platforms, and protection of privacy.

5.1.1 | Sensor selection—Currently, the main options for form factor and placement are mobile phones, wearable on-body sensors and static (i.e., environmental) sensors. Mobile phones are perhaps the most widely used in contemporary computational social science efforts, owing to their ubiquity and the extensive range of raw datastreams collected by a single platform (Harari et al., 2017). The ubiquity of smart phones means there is a potential to engage diverse and hard-to reach populations (Sandstrom, Lathia, Mascolo, & Rentfrow, 2016). Given that mobile phones are already being charged and cared for by participants, they do not pose much burden beyond privacy concerns. Both open source and commercial applications to collect sensor data from phones exist, and we point the interested reader to recently published existing guides for more information (Harari et al., 2016).

In contrast, on-body sensors are more inconvenient, in that they are an additional device that requires active care by participants (e.g., charging, diligence with taking on and off, and correct positioning). However, as they are positioned directly on the body, wearable sensors typically provide much higher quality markers of individual activity. To reduce participant burden in extended recording sessions, comfortable devices with long battery life and quick-charging capabilities are critical. Alternatively, multiple devices can be provided to be alternated while charging. Another relevant consideration is the availability and size of on board storage. Devices that stream data require an additional device and may be more prone to data loss. Additionally, access to high-density raw data as well as tools for visualization or analysis of data may should be considered when selecting sensors.

While our sensing review above focused on data collected from wearable and mobile sensors, static location sensors can capture a variety of interesting datastreams, from physiological data (Adib, Mao, Kabelac, Katabi, & Miller, 2015) and sleep (Rahman et al., 2015) to physical presence (Olguin et al., 2009). Such static sensor systems are not currently available commercially for research purposes but their active development within the ubicomp community suggests that they may soon be. The main benefits of static-device sensing is that they beyond privacy concerns they are very low burden for participants, as they require no charging or active care, and they can be much more powerful as they can be directly connected to wifi and power sources. However, as it is expensive to instrument many spaces with such sensors, data from static sensors will likely not be continuous. Thus, one consideration when considering static sensors is whether interactions of interest are spatially localized or could be adequately characterized in a single spatial area, for example, activity surrounding dinner or bed times.

Several platforms exist for each of these form factors, with consumer market, do-it-yourself (DIY), or bespoke (customizable) options. While consumer-market sensors are typically the most user-friendly, the challenge with using them for research is that raw, high-temporal resolution signals may be inaccessible. Raw sensor data are almost always heavily pre-processed via a proprietary algorithm–which may, for example, average over much data to make up for noisy or low-quality sensors. Whether this is a concern depends on the nature of the phenomenon under study. As with all other sensors, it is advisable to test data validity against a gold-standard in a setting analogous to the true protocol.

It is not difficult to find engineers and computer scientists at local universities interested in collaborating on device construction or in modifying a commercially available open-source sensor platform (e.g., Arduino or Raspberry Pi). The benefit of DIY-sensors is in the direct control over the configuration of the device, including sensed modalities and data precision. While a novel sensor can be developed in a semester or two by a good graduate student, a major consideration is the ultimate usability of the device. Commercial products have teams of designers to ensure comfort, look, and ease of use of devices. Without significant investments there is danger of a functional but practically unusable device. Also, depending on the complexity of the request, for example, whether circuit boards can be purchased or need to be assembled, these devices can quickly become physically unwieldly.

Customizable wearable devices (e.g., mbientlabs.com, move-sense.com) or those built specifically for the research community (e.g., axivity.com, movisens.com, empatica.com) combine many of the benefits of commercial and DIY options. These companies employ both engineers and designers, typically resulting in a functional and usable device. Additionally, because they are catering to researchers, they typically provide access to raw data and potentially relevant activity detection algorithms (e.g., heart beats or sleep detection). While customization may be possible, it will likely be more expensive than working with someone local, and the company may not be willing to work on specialized datastreams if they are not of financial gain. Also, companies may go out of business, or go in a different direction and stop supporting a certain device.

5.1.2 | Single versus multi-sensor platforms—Characterizing complex activity may require multiple datastreams. For example, examining parental responses to child tantrums might involve combining audio, motion and proximity sensing. Some off-the shelf options exist that combine wearable sensors with mobile phone apps that collect additional data. However, existing options may be limited in terms of available datastreams, meaning a single-platform solution may not be feasible. When considering multi-sensor platforms, a major concern is synchronization. Even if each device logs time information, the clocks of individual devices may not be synchronized. Additionally, clocks may drift over the course of data collection, meaning some datapoints (i.e., time-value pairs) are lost or skipped and time is effectively “compressed.” This may not be an issue if high-temporal precision is not required for expected analyses. For example, if hourly summaries of motion data are going to be compared with EMA responses, then a few seconds of drift is not cause for concern. However, if planned analyses require synchronization to within five or even 10 minutes, testing and potentially correcting for time lags drift is advised.

As drift can irregularly affect individual devices, it may be necessary to provide synchronization signals throughout the course of the recordings. One solution is to build a central "timekeeping" device that pings multiple devices. However, this may be impossible if your platform includes commercial devices that cannot be reprogrammed, such as the LENA. An alternative solution is to devise a manual synchronization protocol. The key to developing a synchronization routine is to identify a signal that will be registered by at least one datastream on each device: a clap, for example, can be both heard and seen (via audio or video) and has a unique motion signal (detectable via accelerometer), and thus functions as a good synching cue between these devices. "Synch signals" may not be easy to detect in the milieu of a long (1 hr–2 week) session, especially if they occur sporadically throughout an extended session and other activities could mask the synch cue. In this case, something that will stand out more may be necessary. Our favorites include: a very rapid motion (12–16 g "slam" onto a soft surface), especially if in a sequence of motions (e.g., two slams, a zigzag on the table, and then two slams; see also Ploetz, Chen, Hammerla, & Abowd, 2012). Additionally, adding event-markers around the synch routine is useful for identifying roughly when synch signals will occur. Given these challenges, complete synchronization pipelines should be piloted and assessed before study data are collected.

5.1.3 | Privacy issues—The possibility to capture activity in high-resolution begs the question of participant privacy (Klasnja, Consolvo, Choudhury, Beckwith, & Hightower, 2009). Several high-profile cases indicate that sensor data can have legal consequences, including a recent case in which audio-recordings stored by a voice-controlled speaker were used as evidence in a trial (Sweetland Edwards, 2017; Tokson, 2017). These issues may be disproportionately faced by vulnerable populations such as undocumented immigrants, activists, and people of color, all of whom are more likely to be targets of police violence and intimidation. For high-risk individuals it may be preferable to simply not collect high-density data.

Ultimately, we need to educate potential participants about what might happen with their data if it is breached or legally requested (e.g., whether it can be linked back to them, and what it might be able to tell others about them) before they decide to participate. The ubiquity of consumer sensors means that potential participants are likely already sharing various datastreams with companies. However, consumers may be unaware of this. For example, in a recent survey we completed, over 40% of participants who marked they were unwilling to collect random samples of audio data for research purposes also indicated that they have or would use a voice-controlled device if it were free (Levine et al, in preparation). Thus, while educating participants appears proper, it remains unclear whether providing individuals with information about their existing devices will make them more or less comfortable participating in sensing research. However, even individuals who are somewhat uncomfortable with data use and sharing policies are often willing to share personal data when they perceive concrete benefits to doing so (Matthews, Abdullah, Gay, & Choudhury, 2014; Zhang, De Choudhury, & Grudin, 2014).

There are a number of ways to build in privacy protections for research participants. The basis of these techniques is that the collection and storage of low-resolution data will suffice for many analyses (e.g., storing GPS coordinates at the block rather than address level). The more that the scale of data obscures identification of an individual participant, the more privacy protection is afforded. Similarly, audio and image data can be subsampled to obscure continuous content. Another technique is to store features of interest rather than raw data, a technique that has been deployed both with audio (Wyatt et al., 2011) and image data (Thomaz, Parnami, Bidwell, Essa, & Abowd, 2013). Computations can be done on the collection device, meaning that it is not necessary to store or transfer data to devices or study servers. Note that at preliminary stages of research it may be necessary to collect and store continuous raw datasets which is often critical for the development of novel activity recognition algorithms.

5.2 | Translating raw data into markers of meaningful activity

High-quality tutorials for activity recognition are available (Bulling et al., 2014). Additionally, many machine learning models commonly used for activity recognition are freely available online (e.g., Python sci-kit). Thus, even with no prior experience with activity recognition, advanced undergraduate or masters level computer science students can train and test activity-recognition models.

When working to develop novel activity-recognition models, or when simply evaluating an existing model, there are a number of considerations. First, it is necessary to carefully plan and evaluate the dataset used to train the algorithm. Training datasets collected under "ideal" laboratory conditions may be vastly different than the real-world conditions in which they will be used, ultimately reducing the validity of the detected markers. For example, while developing a model to identify holding and carrying behaviors, early model performance went from 90% accuracy to less than chance after transitioning from a cued paradigm to a naturalistic, free-flowing recording. In the former case, caregivers always picked infants up off the ground beginning from a standing position. In the latter case, there was much more variation in pick-up behaviors–infants were lifted from standing and sitting positions on couches, changing tables, and the floor, which proved to be much more challenging to detect. Our naturalistic recording also included many more negative instances of our behaviors of interest, such as picking a toy up off the ground. The presence of such negative training examples is critical for developing a robust algorithm that will succeed in real-world conditions (for final model, see Yao, Ploetz, Johnson & de Barbara, under review). For these reasons, training data should map onto the true study data as closely as possible, and performance should be assessed in real-world conditions whenever possible.

Next, while an algorithm may have high accuracy overall, it is necessary to consider accuracy for each individual distinction interest within the class, considering both precision (i.e. positive predictive value) as well as recall (i.e. sensitivity). Additionally, it is advisable to assess the performance of algorithms qualitatively via visualizations in addition to quantitatively. Algorithms are designed only to optimize correct predictions, and thus may grossly underperform. For example, the optimal solution algorithmically may be to predict the most common behavior for the entire session, an unappealing solution that may be masked by overall accuracy scores. Finally, measures of accuracy utilized within these studies are different from those traditionally relied upon in social science and thus it is worth additionally assessing reliability using metrics familiar to other psychologists (e.g., Cohen's Kappa; see Bakeman & Quera, 2011).

5.3 | Analysis of behavioral markers

The scale of data possible to collect with sensors is much larger and less bounded than data collected in laboratory settings, meaning that the timescales, operationalization, and impacts of daily activity may not be specified by existing research or theory. Ultimately, identification and operationalization of relevant constructs and their role in shaping developmental trajectories will involve experimentation and, above all, a willingness to become familiar with the data.

5.3.1 | Visualization and iterative processing of data—Data visualizations can provide insight into the structure of phenomena of interest and its variation over time or participants. They can also aid in the generation of hypotheses without the constraints of predetermined analyses that might obscure structure in the data (de Barbara, Johnson, Forster et al., 2013; Gnisci et al., 2008). In particular, iterative cycles between visualizations of raw and semi- processed data can be useful for converging upon variables that can accurately capture phenomena of interest (Fricker, Zhang, & Yu, 2011; Yu, Yurovsky, & Xu, 2012). In later stages, visualizations are key for identifying and refining appropriate quantitative methods to describe or model behaviors of interest. A number of papers detail the process of visualizing high-density mobile sensor data to provide these insights (Polack et al., 2018; Sharmin et al., 2015; Zisook, Hernandez, Goodwin, & Picard, 2013).

5.3.2 | Nested timescales in the analysis of continuous behavior—Leveraging the temporal organization of behavior can be useful for structuring analyses. Behavior is organized across many time- scales. At the timescale of seconds, we can observe contingent gaze shifts and vocalizations, at timescales of minutes and hours, we can observe episodes of play or arguments, and at the timescale of months or years, we can observe the blossoming of a relationship. These nested timescales are relevant when considering how to analyze many hours of data possible to collect with mobile sensors. Parsing continuous datastreams of frame-by-frame activity into meaningful episodes at the timescale of minutes and hours–such as episodes of distress, play or joint activity–is a useful way to structure analyses of high-density data. This can be particularly insightful if these episodes occur repeatedly within a single session as well as within sessions across longitudinal data collection. For example, to study the changing nature of triadic activity across the first year, de Barbara and colleagues identified all instances of "maternal bids" within mother-infant free-play interactions (de Barbara, Johnson, Forster & Deak, 2016). By comparing changes in infants object looking and touching following maternal bids across the longitudinal time period, it was possible to characterize more gradual and continuous shifts in multiple dimensions of activity which contributed to the emergence of classic triadic interactions around the end of the first year (de Barbara, Johnson, & Deak, 2013; de Barbara, Johnson, Forster, & Deak, 2016). In effect, the maternal bid episodes functioned as a naturally occurring spontaneous "trigger" for an event-related analysis. This allowed us to summarize and compare the event-related micro-dynamics of activity across longitudinal time (see also Forster & Rodriguez, 2006, and de Barbara, Johnson, Forster et al., 2013, for a more extensive discussion of this approach).

5.3.3 | Avoiding datamining—The iterative nature of the analysis of high-density data makes it susceptible to the critique of datamining. Insofar as tinkering with data can lead to spurious results, this is a legitimate concern. However, it cannot be a paralyzing fear as iterative analyses are likely necessary at the early stages of computational social science research.

To avoid spurious findings, it is best to select a subset of data which will be used for the sake of gaining insight into the structure of the dataset. Later, this subset can be used to develop and test that any new measures or variables validly and appropriately capture the phenomena of interest. The remainder of the sample should be analyzed only after measures have been finalized within this subset (see also de Barbara, Johnson, Forster et al., 2013). This is a logic analogous to that used in machine learning where a subset of training data is "held out" to test the results of the model. Testing on a subset ensures that the model is not overfit to the particulars of the training dataset. Finally, just as in any scientific paper, accurate reporting of the complete analysis protocol (including all variations tested) is necessary for a fair assessment of the quality and legitimacy of the results.

5.4 | Collaborations and Training

Mobile sensing research requires expertise across domains. Electrical engineers build circuit hardware and the low-level "firmware" that runs on them. Experts in human-computer interaction design the platforms and form factors that allow sensors to seamlessly integrate into daily interactions. Data security and privacy ethicists are necessary to understand and communicate the potential risks of data breaches as it becomes increasingly difficult to guarantee that data are not identifiable. Computer science and data analysts provide powerful algorithms and visualization tools that aid in making sense of the immense amounts of data that can now be collected.

Critically, social scientists provide motivation and insight into phenomena of interest, as well as theoretical and practical knowledge that can carve meaning from multitudes of data, and ultimately construct new theories of development.

For those wanting to become involved in mobile sensing research, the range of necessary expertise may appear a daunting challenge. Given traditional disciplinary boundaries, this will typically be an interdisciplinary effort requiring wide-reaching collaborations. For psychology students wanting to become pursue computational social science research, we provide some training guidelines.

The availability of off-the-shelf sensors is relatively good and thus it is not necessary to gain electrical engineering skills to pursue sensing research. However, for those interested, there are a number of basic hands-on tutorials which provide basic understanding of circuits and the firmware that runs on them, (e.g., the Arduino Starter Kit, available online). Experience with synchronization of sensors and merging data across diverse timescales will need to be learned independently by experimentation, or from working in a developmental lab that collects and annotates video data (see also Bakeman & Quera, 2011; Gottman, 1981).

The basic programming necessary to work with and analyze sensor data can be attained in two to six college-level courses. Introductory programming courses will cover necessary basics of data structure, manipulation, and visualization. This material is also available in online tutorials and user manuals (e.g., at mathworks.com or python.org). Some training in timeseries analyses or machine learning is also beneficial. A timeseries analysis course will cover basics such as interpolation, filtering and time- and frequency- domain analyses used to quantify and describe timeseries. There are excellent resources for timeseries analyses written specifically for those without extensive quantitative or computational training (Gottman, 1981). Nearly every computer science department will have a machine learning course which should be accessible with basic statistics and programming experience. More specialized experience in activity recognition may also be desired, and is available in select engineering or computer science departments, or via online tutorials.

Note that basic formatting, manipulation, and visualization of sensor data can be done in Excel, and commercial software for video annotation and sensor data collection typically include visualization and analysis capabilities. These programs are often restrictive, especially when complex or flexible operationalizations are necessary, and they are typically not programmable, meaning that any work done on an individual session must be repeated over the complete dataset. However, they can be a good start for students uncertain about committing to more extensive computational training. These programs can also be very useful as they allow rapid examination of data and where basic tools are sufficient for analyses.

6 | CONCLUSIONS

Insofar as we can leverage sensor data to access the processes by which daily activity and interaction shape trajectories across the lifespan, mobile sensing tools provide a unique–indeed, transformative–opportunity for developmental science. If successful, these tools can provide insight into new mechanisms of development: how structure and variation in our daily environments and experiences matter for outcomes. With repeated samples at large scale, such data can move us beyond simple causal models of behavior to understanding how complex non-linear dynamics can shape developmental trajectories.

Much work remains to be done before this grand promise of mobile sensors for developmental science can become a reality. Meeting these challenges will require coordinated interdisciplinary efforts with social scientists involved throughout the research process. The involvement of social scientists is critical for guiding not only the questions that are asked, but the next generation of engineering and computer science research: what is sensed and how algorithms discover meaningful structure in data.

Acknowledgments

Funding information

This work was supported by NIMH KOI Award (1K01MH111957-01A1) to Kaya de Barbaro.

REFERENCES

aan het Rot M, Hogenelst K, & Schoevers RA (2012). Mood disorders in everyday life: A systematic review of experience sampling and ecological momentary assessment studies. Clinical Psychology Review, 32(6), 510–523. 10.1016/j.cpr.2012.05.007 [PubMed: 22721999]

Abowd GD, & Mynatt ED (2000). Charting past, present, and future research in ubiquitous computing. ACM Transactions on Computer-Human Interaction (TOCHI), 7(1), 29–58. 10.1145/344949.344988

Adib F, Mao H, Kabelac Z, Katabi D, & Miller RC (2015). Smart homes that monitor breathing and heart rate. Paper presented at the Proceedings of the 33rd annual ACM conference on human factors in computing systems.

Adolph K (2016). Video as data: From transient behavior to tangible recording. APS Observer, 29(3), 23. [PubMed: 28736493]

Arif M, & Kattan A (2015). Physical activities monitoring using wearable acceleration sensors attached to the body. PLoS ONE, 10(7), e0130851 10.1371/journal.pone.0130851 [PubMed: 26203909]

Bae S, Chung T, Ferreira D, Dey AK, & Suffoletto B (2017). Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions. Addictive Behaviors, 83, 42–47. [PubMed: 29217132]

Bakeman R, & Gottman J (1997). Observing interaction: An introduction to sequential analysis. Cambridge, UK: Cambridge Univ Press.

Bakeman R, & Quera V (2011). Sequential analysis and observational methods for the behavioral sciences. Cambridge, UK: Cambridge University Press.

Barrat A, Cattuto C, Tozzi A, Vanhems P, & Voirin N (2014). Measuring contact patterns with wearable sensors: Methods, data characteristics and applications to data-driven simulations of infectious diseases. Clinical Microbiology and Infection, 20(1), 10–16. 10.1111/1469-0691.12472

Basu S (2002). Conversational scene analysis. Cambridge, MA: Massachusetts Institute of Technology Retrieved from http://hdl.handle.net/1721.1/29270

Baumeister RF, Vohs KD, & Funder DC (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? Perspectives on Psychological Science, 2(4), 396–403. [PubMed: 26151975]

Belsky J (1984). The determinants of parenting: A process model. Child Development, 83–96. 10.2307/1129836 [PubMed: 6705636]

Belsky J·, Bakermans-Kranenburg MJ, & Van IJzendoorn MH (2007). For better and for worse differential susceptibility to environmental influences. Current Directions in Psychological Science, 16(6), 300–304. 10.1111/j.1467-8721.2007.00525.x

Ben-Zeev D, Scherer EA, Brian RM, Mistier LA, Campbell AT, & Wang R (2017). Use of multimodal technology to identify digital correlates of violence among inpatients with serious mental illness: A pilot study. Psychiatric Services, 68(10), 1088–1092. 10.1176/appi.ps.201700077 [PubMed: 28669285]

Ben-Zeev D, Brian R, Wang R, Wang W, Campbell AT, Aung MS, … Hauser M (2017). Crosscheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. Psychiatric Rehabilitation Journal, 40(3), 266 10.1037/prj0000243 [PubMed: 28368138]

Berke EM, Choudhury T, Ali S, & Rabbi M (2011). Objective measurement of sociability and activity: Mobile sensing in the community. The Annals of Family Medicine, 9(4), 344–350. 10.1370/afm.1266 [PubMed: 21747106]

Bertz JW, Epstein DH, & Preston KL (2018). Combining ecological momentary assessment with objective, ambulatory measures of behavior and physiology in substance-use research. Addictive Behaviors, 83, 5–17. 10.1016/j.addbeh.2017.ll.027 [PubMed: 29174666]

Black MP, Katsamanis A, Baucom BR, Lee C-C, Lammert AC, Christensen A, … Narayanan SS (2013). Toward automating a human behavioral coding system for married couples' interactions using speech acoustic features. Speech Communication, 55(1), 1–21. 10.1016/j.specom.2011.12.003

Bolger N, & Laurenceau J (2013). Intensive longitudinal methods: an introduction to diary and experience sampling research. 2013 New York, NY: Guilford Press.

Bongard J, & Lipson H (2007). Automated reverse engineering of nonlinear dynamical systems. Proceedings of the National Academy of Sciences, 104(24), 9943–9948. 10.1073/pnas.0609476104

Boonstra TW, Larsen ME, Townsend S, & Christensen H (2017). Validation of a smartphone app to map social networks of proximity. PLoS ONE, 12(12), e0189877 10.1371/journal.pone.0189877 [PubMed: 29261782]

Borjon JI, Schroer SE, Bambach S, Slone LK, Abney DH, Crandall DJ, & Smith LB (2018). A view of their own: Capturing the egocentric view of infants and toddlers with head-mounted cameras. Journal of Visualized Experiments, 140, e58445.

Breiman L (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical Science, 16(3), 199–231. 10.1214/ss/1009213726

Brocklebank LA, Falconer CL, Page AS, Perry R, & Cooper AR (2015). Accelerometer-measured sedentary time and cardiometabolic biomarkers: A systematic review. Preventive Medicine, 76, 92–102. 10.1016/j.ypmed.2015.04.013 [PubMed: 25913420]

Bronfenbrenner U (1977). Toward an experimental ecology of human development. American Psychologist, 32(7), 513 10.1037/0003-066X.32.7.513

Brown NA, Blake AB, & Sherman RA (2017). A snapshot of the life as lived: Wearable cameras in social and personality psychological science. Social Psychological and Personality Science, 8(5), 592–600. 10.1177/1948550617703170

Bulling A, Blanke U, & Schiele B (2014). A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys (CSUR), 46(3), 33 10.1145/2499621

Buss DM (1981). Predicting parent-child interactions from children’s activity level. Developmental Psychology, 17(1), 59 10.1037/0012-1649.17.1.59

Buss DM, Block JH, & Block J (1980). Preschool activity level: Personality correlates and developmental implications. Child Development, 401–408. 10.2307/1129273

Busser H, Ott J, Van Lummel R, Uiterwaal M, & Blank R (1997). Ambulatory monitoring of children’s activity. Medical Engineering and Physics, 19(5), 440–445. 10.1016/S1350-4533(97)00007-6 [PubMed: 9338884]

Bussmann J, & Ebner-Priemer U (2012). Ambulatory assessment of movement behavior: Methodology, measurement, and application. In Handbook of research methods for studying daily life (pp. 235–250).

Byrnes HF, Miller BA, Morrison CN, Wiebe DJ, Woychik M, & Wiehe SE (2017). Association of environmental indicators with teen alcohol use and problem behavior: Teens’ observations vs. objectively-measured indicators. Health & Place, 43, 151–157. 10.1016/j.healthplace.2016.12.004 [PubMed: 28061392]

Cacioppo JT, Uchino BN, & Berntson GG (1994). Individual differences in the autonomic origins of heart rate reactivity: The psychometrics of respiratory sinus arrhythmia and preejection period. Psychophysiology, 31(4), 412–419. 10.1111/j.1469-8986.1994.tb02449.x [PubMed: 10690921]

Caldwell BM, & Bradley RH (1984). Home observation for measurement of the environment. Little Rock, AR: University of Arkansas.

Campos JJ, Anderson DI, Barbu-Roth MA, Hubbard EM, Hertenstein MJ, & Witherington D (2000). Travel broadens the mind. Infancy, 1(2), 149–219. 10.1207/S15327078IN0102_1

Chaix B, Meline J, Duncan S, Merrien C, Karusisi N, Perchoux C, … Kestens Y (2013). GPS tracking in neighborhood and health studies: A step forward for environmental exposure assessment, a step backward for causal inference? Health & Place, 21,46–51. 10.1016/j.healthplace.2013.01.003 [PubMed: 23425661]

Chan L, Swain VD, Kelley C, de Barbara K, Abowd GD, & Wilcox L (2018). Students’ experiences with ecological momentary assessment tools to report on emotional well-being. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2(1), 3 10.1145/3191735

Chatterjee S, Hovsepian K, Sarker H, Saleheen N, al'Absi M, …S. (2016). mCrave: Continuous estimation of craving during smoking cessation. Paper presented at the Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Chin I, Goodwin MS, Vosoughi S, Roy D, & Naigles LR (2017). Dense home-based recordings reveal typical and atypical development of tense/aspect in a child with delayed language development. Journal of Child Language, 1–34.

Christakis DA, Gilkerson J, Richards JA, Zimmerman FJ, Garrison MM, Xu D, … Yapanel U (2009). Audible television and decreased adult words, infant vocalizations, and conversational turns: A population-based study. Archives of Pediatrics & Adolescent Medicine, 163(6), 554–558. 10.1001/archpediatrics.2009.61 [PubMed: 19487612]

Clerkin EM, Hart E, Rehg JM, Yu C, & Smith LB (2017). Real-world visual statistics and infants' first-learned object names. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1711), 20160055 10.1098/rstb.2016.0055

Cook DJ, Duncan G, Sprint G, & Fritz RL (2018). Using smart city technology to make healthcare smarter. Proceedings of the IEEE, 106(4), 708–722. 10.1109/JPROC.2017.2787688 [PubMed: 29628528]

Cordeiro F, Bales E, Cherry E, & Fogarty J (2015). Rethinking the mobile food journal: Exploring opportunities for lightweight photo-based capture. Paper presented at the Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.

Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, & Bolger N (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? Personality and Social Psychology Bulletin, 32(7), 917–929. 10.1177/0146167206287721 [PubMed: 16738025]

Csikczentmihalyi M, & Larson R (1983). The experience sampling method. New directions for naturalistic methods in the behavioral sciences. San Francisco: Jossey-Bass.

de Barbara K, Johnson CM, & Deak GO (2013). Twelve-month "social revolution" emerges from mother-infant sensorimotor coordination: A longitudinal investigation. Human Development, 56(4), 223–248. 10.1159/000351313

de Barbara K, Johnson CM, Forster D, & Deak GO (2013). Methodological considerations for investigating the micro-dynamics of social interaction development. IEEE Transactions on Autonomous Mental Development, 5(3), 258–270. 10.1109/TAMD.2013.2276611

de Barbara K, Johnson CM, Forster D, & Deak GO (2016). Sensorimotor decoupling contributes to triadic attention: A longitudinal investigation of mother-infant-object interactions. Child Development, 87(2), 494–512. 10.1111/cdev.12464 [PubMed: 26613383]

De Choudhury M, Counts S, & Gamon M (2012). Not all moods are created equal! exploring human emotional states in social media. Paper presented at the Sixth international AAAI conference on weblogs and social media.

De Choudhury M, Gamon M, Counts S, SHorvitz E (2013). Predicting Depression via Social Media. ICWSM, 13, 1–10.

De Choudhury M, Kiciman E, Dredze M, Coppersmith G, & Kumar M (2016). Discovering shifts to suicidal ideation from mental health content in social media. Paper presented at the Proceedings of the 2016 CHI conference on human factors in computing systems media.

Deak GO, Krasno AM, Triesch J, Lewis J, & Sepeta L (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. Developmental Science, 17(2), 270–281. 10.1111/desc.12122 [PubMed: 24387193]

Dragon T, Arroyo I, Woolf BP, Burleson W, El Kaliouby R, & Eydgahi H (2008). Viewing student affect and learning through classroom observation and physical sensors. Paper presented at the International Conference on Intelligent Tutoring Systems.

Eagle N, & Pentland AS (2006). Reality mining: Sensing complex social systems. Personal and Ubiquitous Computing, 10(4), 255–268. 10.1007/s00779-005-0046-3

Ebner-Priemer UW, & Trull TJ (2009). Ecological momentary assessment of mood disorders and mood dysregulation. Psychological Assessment, 21(4), 463 10.1037/a0017075 [PubMed: 19947781]

Eronen AJ, Peltonen VT, Tuomi JT, Klapuri AP, Fagerlund S, Sorsa T, …Huopaniemi J (2006). Audio-based context recognition. IEEE Transactions on Audio, Speech, and Language Processing, 14(1), 321–329. 10.1109/TSA.2005.854103

Fausey CM, Jayaraman S, & Smith LB (2016). From faces to hands: Changing visual input in the first two years. Cognition, 152, 101–107. 10.1016/j.cognition.2016.03.005 [PubMed: 27043744]

Forster D (2002). Consort turnovers as distributed cognition in Olive baboons: A distributed approach to mind In Bekoff M, Allen C & Burghardt G, The cognitive animal: Empirical and theoretical perspectives on animal cognition, (pp. 163–171). Boston, MA: MIT Press.

Forster D, & Rodriguez PF (2006). Social complexity and distributed cognition in Olive Baboons (Papio anubis): Adding system dynamics to analysis of interaction data. Aquatic Mammals, 32(4), 528–543. 10.1578/am.32.4.2006.528.

Fricker D, Zhang H, & Yu C (2011). Sequential pattern mining of multimodal data streams in dyadic interactions. Paper presented at the 2011 IEEE International Conference on Development and Learning.

Frijda NH (1986). The emotions. Cambridge, UK: Cambridge University Press.

Froehlich JE, Larson E, Campbell T, Haggerty C, Fogarty J, & Patel SN (2009). HydroSense: Infrastructure-mediated single-point sensing of whole-home water activity. Paper presented at the Proceedings of the 11th international conference on Ubiquitous computing.

Gibson EJ (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. Annual Review of Psychology, 39(1), 1–42. 10.1146/annurev.ps.39.020188.000245

Gill AJ, French RM, Gergle D, & Oberlander J (2008). Identifying emotional characteristics from short blog texts. Paper presented at the Proc. 30th Ann. Conf Cognitive Science Soc., Love BC, McRae K, and Sloutsky VM, eds.

Gnisci A, Bakeman R, & Quera V (2008). Blending qualitative and quantitative analyses in observing interaction: Misunderstandings, applications and proposals. International Journal of Multiple Research Approaches, 2(1), 15–30. 10.5172/mra.455.2.l.15

Goguen JA (1997). Towards a social, ethical theory of information In Bowker G, Gasser L, Star LS, & Turner WA (Eds.), Social science research, technical systems and cooperative work: Beyond the great divide (pp. 27–56). Mahwah, NH: Lawrence Erlbaum.

Goodwin MS, Velicer WF, & Intille SS (2008). Telemetric monitoring in the behavior sciences. BehaviorResearch Methods, 40(1), 328–341. 10.3758/BRM.40.l.328

Gottesman II, & Hanson DR (2005). Human development: Biological and genetic processes. Annual Review of Psychology, 56, 263–286. 10.1146/annurev.psych.56.091103.070208

Gottlieb G (1998). Normally occurring environmental and behavioral influences on gene activity: From central dogma to probabilistic epigenesis. Psychological Review, 105(4), 792 10.1037/0033-295X.105.4.792-802 [PubMed: 9830380]

Gottlieb G (2000). Environmental and behavioral influences on gene activity. Current Directions in Psychological Science, 9(3), 93–97. 10.1111/1467-8721.00068

Gottman J (1981). Time-series analysis: A comprehensive introduction for social scientists. Cambridge, UK: Cambridge University Press Cambridge.

Gottman J, & Roy A (1990). Sequential analysis: A guide for behavioral researchers. Cambridge, UK: Cambridge Univ Press.

Granic I, & Hollenstein T (2003). Dynamic systems methods for models of developmental psychopathology. Development and Psychopathology, 15(3), 641–669. 10.1017/S0954579403000324 [PubMed: 14582935]

Granic I, Hollenstein T, & Lichtwarck-Aschoff A (2016). A survey of dynamic systems methods for developmental psychopathology. In Cicchetti D (Ed.), Hoboken, NJ: Developmental Psychopathology.

Granic I, & Patterson GR (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. Psychological Review, 113(1), 101 10.1037/0033-295X.113.l.101 [PubMed: 16478303]

Gunthert KC, & Wenze SJ (2011). Daily diary methods In Mehl M & Conner T (Eds.), Handbook of research methods for studying daily life. New York, NY: Guilford Press.

Guo R, Li S, He L, Gao W, Qi H, & Owens G (2013). Pervasive and unobtrusive emotion sensing for human mental. Paper presented at the Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare.

Gupta S, Reynolds MS, & Patel SN (2010). ElectriSense: Single-point sensing using EMI for electrical event detection and classification in the home. Paper presented at the Proceedings of the 12th ACM international conference on Ubiquitous computing.

Harari GM, Lane ND, Wang R, Crosier BS, Campbell AT, & Gosling SD (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. Perspectives on Psychological Science, 11(6), 838–854. 10.1177/1745691616650285 [PubMed: 27899727]

Harari GM, Müller SR, Aung MS, & Rentfrow PJ (2017). Smartphone sensing methods for studying behavior in everyday life. Current Opinion in Behavioral Sciences, 18, 83–90. 10.1016/j.cobeha.2017.07.018

Hariharan R, Krumm J, & Horvitz E (2005). Web-enhanced GPS. Paper Presented at the International Symposium on Location-and Context-Awareness.

Healey J (2000). Future possibilities in electronic monitoring of physical activity. Research Quarterly for Exercise and Sport, 71(sup2), 137–145. 10.1080/02701367.2000.11082797 [PubMed: 25680024]

Hodges S, Williams L, Berry E, Izadi S, Srinivasan J, Butler A, … Wood K (2006). SenseCam: A retrospective memory aid. Paper presented at the International Conference on Ubiquitous Computing.

Hovsepian K, al'Absi M, Ertin E, Kamarck S (2015). cStress: Towards a gold standard for continuous stress assessment in the mobile environment. Paper presented at the Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Hutchins E (1995). Cognition in the wild. Cambridge, MA: MIT press.

Imel ZE, Barco JS, Brown HJ, Baucom BR, Baer JS, Kircher JC, & Atkins DC (2014). The association of therapist empathy and synchrony in vocally encoded arousal. Journal of Counseling Psychology, 61(1), 146 10.1037/a0034943 [PubMed: 24274679]

Insel TR (2017). Digital phenotyping: Technology for a new science of behavior. JAMA, 318(13), 1215–1216. 10.1001/jama.2017.11295 [PubMed: 28973224]

Intille S, Haynes C, Maniar D, Ponnada A, & Manjourides J (2016). μEMA: Microinteraction-based ecological momentary assessment (EMA) using a smartwatch. Paper presented at the Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Jayaraman S, Fausey CM, & Smith LB (2015). The faces in infant- perspective scenes change over the first year of life. PLoS ONE, 10(5), e0123780 10.1371/journal.pone.0123780 [PubMed: 26016988]

Jia W, Li Y, Qu R, Baranowski T, Burke LE, Zhang H, …Mao Z-H (2018). Automatic food detection in egocentric images using artificial intelligence technology. Public Health Nutrition, 1–12. 10.1017/S1368980018000538

Juslin PN, & Scherer KR (2005). Vocal expression of affect. The New Handbook of Methods in Nonverbal Behavior Research, 65–135.

Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group. (2008). Continuous glucose monitoring and intensive treatment of type 1 diabetes. New England Journal of Medicine, 359(14), 1464–1476. [PubMed: 18779236]

Kaye JA, Maxwell SA, Mattek N, Hayes TL, Dodge H, Pavel M, … Zitzelberger TA (2011). Intelligent systems for assessing aging changes: Home-based, unobtrusive, and continuous assessment of aging. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66 (suppl_1), i180–i190.

Kim JC, & Clements MA (2015). Formant-based feature extraction for emotion classification from speech. Paper presented at the Telecommunications and Signal Processing (TSP), 2015 38th International Conference on.

Kim KH, Bang SW, & Kim SR (2004). Emotion recognition system using short-term monitoring of physiological signals. Medical and Biological Engineering and Computing, 42(3), 419–427. 10.1007/BF02344719 [PubMed: 15191089]

Kirn-Spoon J, & Grimm KJ (2016). Latent Growth modeling and developmental psychopathology. In Cicchetti D (Ed.), Developmental Psychopathology Hoboken, NJ: Wiley.

Kirschbaum C, Pirke K-M, & Hellhammer DH (1993). The ‘Trier Social Stress Test’-a tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology, 28(1–2), 76–81. 10.1159/000119004 [PubMed: 8255414]

Klasnja P, Consolvo S, Choudhury T, Beckwith R, & Hightower J (2009). Exploring privacy concerns about personal sensing. Paper presented at the International Conference on Pervasive Computing.

Kretch KS, Franchak JM, & Adolph KE (2014). Crawling and walking infants see the world differently. Child Development, 85(4), 1503–1518. 10.1111/cdev.12206 [PubMed: 24341362]

Krull JL, Cheong J, Fritz MS, & MacKinnon DP (2016). Moderation and mediation in interindividual longitudinal analysis. In Cicchetti D (Ed.), Developmental Psychopathology Hoboken, NJ: Wiley.

Kubiak T, & Krog K (2011). Computerized sampling of experience and behavior In Mehl M, & Conner T (Eds.), Handbook of research methods for studying daily life (pp. 124–143). New York, NY: Gilford Press.

Lara OD, & Labrador MA (2013). A survey on human activity recognition using wearable sensors. IEEE Communications Surveys and Tutorials, 15(3), 1192–1209. 10.1109/SURV.2012.110112.00192

Lathia N, Sandstrom GM, Mascolo C, & Rentfrow PJ (2017). Happier people live more active lives: Using smartphones to link happiness and physical activity. PLoS ONE, 12(1), e0160589 10.1371/journal.pone.0160589 [PubMed: 28052069]

Lazer D, Pentland A, Adamic L, Aral S, Barabasi AL, Brewer DM (2009). Life in the network: The coming age of computational social science. Science (New York, N.Y.), 323(5915), 721–723. 10.1126/science.1167742 [PubMed: 19197046]

LeCun Y, Bengio Y, & Hinton G (2015). Deep learning. Nature, 521(7553), 436 10.1038/naturel4539 [PubMed: 26017442]

Lee DK, Cole WG, Golenia L, & Adolph KE (2018). The cost of simplifying complex developmental phenomena: A new perspective on learning to walk. Developmental Science, 21(4), el2615 10.1111/desc.12615

Leventhal T, & Brooks-Gunn J (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. Psychological Bulletin, 126(2), 309–337. 10.1037/0033-2909.126.2.309 [PubMed: 10748645]

Li JJ, & Lansford JE (2018). A smartphone-based ecological momentary assessment of parental behavioral consistency: Associations with parental stress and child ADHD symptoms. Developmental Psychology, 54(6), 1086–1098. 10.1037/dev0000516 [PubMed: 29608072]

Lu H, Frauendorfer D, Rabbi M, Mast MS, Chittaranjan GT, Campbell AT, … Choudhury T (2012). StressSense: Detecting stress in unconstrained acoustic environments using smartphones. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.

Ludwig T (2018). Automated detection of infant crying and fussing for clinical applications. Paper presented at the The European Society for Paediatric Gastroenterology, Hepatology and Nutrition, Geneva, Switzerland.

Maekawa T, Anderson SJ, de Brecht M, & Yamagishi N (2018). The effect of mood state on visual search times for detecting a target in noise: An application of smartphone technology. PLoS ONE, 13(4), e0195865 10.1371/journal.pone.0195865 [PubMed: 29664952]

Matthews M, Abdullah S, Gay G, & Choudhury T (2014). Tracking mental well-being: Balancing rich sensing and patient needs. Computer, 47(4), 36–43.

McConnell MV, Shcherbina A, Pavlovic A, Homburger JR, Goldfeder RL, Waggot D, … Myers J (2017). Feasibility of obtaining measures of lifestyle from a smartphone app: The MyHeart Counts Cardiovascular Health Study. JAMA Cardiology, 2(1), 67–76. 10.1001/jamacardio.2016.4395 [PubMed: 27973671]

Mead MI, Popoola O, Stewart G, Landshoff P, Calleja M, Hayes M, … Dicks J (2013). The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. Atmospheric Environment, 70, 186–203. 10.1016/j.atmosenv.2012.11.060

Mehl MR (2017). The electronically activated recorder (EAR) a method for the naturalistic observation of daily social behavior. Current Directions in Psychological Science, 26(2), 184–190. 10.1177/0963721416680611 [PubMed: 28529411]

Mehl M, & Conner T (2011). Handbook of research methods for studying daily life. New York, NY: Guilford Press.

Miller G (2012). The smartphone psychology manifesto. Perspectives on Psychological Science, 7(3), 221–237. 10.1177/1745691612441215 [PubMed: 26168460]

Montanari A, Nawaz S, Mascolo C, & Sailer K (2017). A Study of Bluetooth Low Energy performance for human proximity detection in the workplace. Paper presented at the Pervasive Computing and Communications (PerCom), 2017 IEEE International Conference on.

Moore E II, Clements MA, Peifer JW, & Weisser L (2008). Critical analysis of the impact of glottal features in the classification of clinical depression in speech. IEEE Transactions on Biomedical Engineering, 55(1), 96–107. 10.1109/TBME.2007.900562 [PubMed: 18232351]

Nam Y, & Park JW (2013). Child activity recognition based on cooperative fusion model of a triaxial accelerometer and a barometric pressure sensor. IEEE Journal of Biomedical and Health Informatics, 17(2), 420–426. 10.1109/JBHI.2012.2235075 [PubMed: 24235114]

NICHD Early Child Care Research Network. (1997). The effects of infant child care on infant-mother attachment security: Results of the NICHD study of early child care NICHD early child care research network. Child Development, 68(5), 860–879. [PubMed: 29106728]

Olguin DO, Gloor PA, & Pentland AS (2009). Capturing individual and group behavior with wearable sensors. Paper presented at the Proceedings of the 2009 aaai spring symposium on human behavior modeling

Olguin DO, Waber BN, Kim T, Mohan A, Ara K, & Pentland A (2009). Sensible organizations: Technology and methodology for automatically measuring organizational. Behavior IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 39(1), 43–55. 10.1109/TSMCB.2008.2006638

Osmani V, Carreras I, Matic A, & Saar P (2014). An analysis of distance estimation to detect proximity in social interactions. Journal of Ambient Intelligence and Humanized Computing, 5(3), 297–306. 10.1007/sl2652-012-0171-6

Oyama S (1985/2000). The ontogeny of information: Developmental systems and evolution. Durham, NC: Duke University Press.

Peeters F, Berkhof J, Delespaul P, Rottenberg J, & Nicolson NA (2006). Diurnal mood variation in major depressive disorder. Emotion, 6(3), 383 10.1037/1528-3542.6.3.383 [PubMed: 16938080]

Pellegrini AD, & Smith PK (1998). Physical activity play: The nature and function of a neglected aspect of play. Child Development, 69(3), 577–598. 10.1111/j.1467-8624.1998.tb06226.x [PubMed: 9680672]

Pennebaker JW, Mehl MR,& Niederhoffer KG(2003). Psychological aspects of natural language use: Our words, our selves. Annual Review of Psychology, 54(1), 547–577. 10.1146/annurev.psych.54.101601.145041

Pentland A (2000). Looking at people: Sensing for ubiquitous and wearable computing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1), 107–119. 10.1109/34.824823

Pirsiavash H, & Ramanan D (2012). Detecting activities of daily living in first-person camera views. Paper presented at the Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on.

Ploetz T, Chen C, Hammerla NY, & Abowd GD (2012). Automatic synchronization of wearable sensors and video-cameras for ground truth annotation - A practical approach. Paper presented at the Wearable Computers (ISWC), 2012 16th International Symposium on.

Ploetz T, Hammerla NY, Rozga A, Reavis A, Call N, & Abowd GD (2012). Automatic assessment of problem behavior in individuals with developmental disabilities. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.

Plomin R, DeFries JC, Knopik VS, & Neiderhiser JM (2016). Top 10 replicated findings from behavioral genetics. Perspectives on Psychological Science, 11(1), 3–23. 10.1177/1745691615617439 [PubMed: 26817721]

Polack PJ Jr, Chen S-T, Kahng M, Barbara KD, Basole R, Sharmin M, & Chau DH (2018). Chronodes: Interactive multifocus exploration of event sequences. ACM Transactions on Interactive Intelligent Systems (TiiS), 8(1), 2. [PubMed: 29515937]

Poliak JP, Adams P, & Gay G (2011). PAM: A photographic affect meter for frequent, in situ measurement of affect. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

Rachuri KK, Musolesi M, Mascolo C, Rentfrow PJ, Longworth C, & Aucinas A (2010). EmotionSense: A mobile phones based adaptive platform for experimental social psychology research. Paper presented at the Proceedings of the 12th ACM international conference on Ubiquitous computing

Rahman MM, Bari R, Ali AA, Sharmin M, Raij A, Hovsepian K, … Kumar S (2014). Are we there yet?: Feasibility of continuous stress assessment via wireless physiological sensors. Paper presented at the Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics.

Rahman T, Adams AT, Ravichandran RV, Zhang M, Patel SN, Kientz JA, & Choudhury T (2015). Dopplesleep: A contactless unobtrusive sleep sensing system using short-range doppler radar. Paper presented at the Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Raikes H, Alexander Pan B, Luze G, Tamis-LeMonda CS, Brooks-Gunn J, Constantine J, … Rodriguez ET (2006). Mother-child bookreading in low-income families: Correlates and outcomes during the first three years of life. Child Development, 77(4), 924–953. 10.1111/j.1467-8624.2006.00911.x [PubMed: 16942498]

Ramírez-Esparza N, García-Sierra A, & Kuhl PK (2014). Look who’s talking: Speech style and social context in language inputto infants are linked to concurrent and future speech development. Developmental Science, 17(6), 880–891. 10.1111/desc.12172 [PubMed: 24702819]

Ramsay G, Bailey J, & Ghai S (2016). Developmental cascades in early social interaction: Vocal biomarkers of risk for autism in the first two years of life. Paper presented at the International Conference on Infant Studies, New Orleans, LA.

Rao H, Kim JC, Clements MA, Rozga A, & Messinger DS (2014). Detection of children’s paralinguistic events in interaction with caregivers. Paper presented at the Fifteenth Annual Conference of the International Speech Communication Association.

Rehg J, Abowd G, Rozga A, Romero M, Clements M, Sclaroff S, … Kim C (2013). Decoding children’s social behavior. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Robillard R, Hermens DF, Naismith SL, White D, Rogers NL, Ip TK, … Smith KL (2015). Ambulatory sleep-wake patterns and variability in young people with emerging mental disorders. Journal of Psychiatry & Neuroscience: JPN, 40(1), 28 10.1503/jpn.130247 [PubMed: 25203899]

Roy BC, Frank MC, DeCamp P, Miller M, & Roy D(2015). Predicting the birth of a spoken word. Proceedings of the National Academy of Sciences, 112(41), 12663–12668. 10.1073/pnas.1419773112

Rude S,Gortner E-M, & Pennebaker J (2004).Language use of depressed and depression-vulnerable college students. Cognition & Emotion, 18(8), 1121–1133. 10.1080/02699930441000030

Sadeh A (2011). The role and validity of actigraphy in sleep medicine: An update. SleepMedicine Reviews, 15(4), 259–267. 10.1016/j.smrv.2010.10.001

Sadeh A, Sharkey M, & Carskadon MA (1994). Activity-based sleep- wake identification: An empirical test of methodological issues. Sleep, 17(3), 201–207. 10.1093/sleep/17.3.201 [PubMed: 7939118]

Saleheen N, Ali AA, Hossain SM, Sarker H, Chatterjee S, Marlin B, … Kumar S (2015). puffMarker: A multi-sensor approach for pinpointing the timing of first lapse in smoking cessation. Paper presented at the Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Sandstrom GM, Lathia N, Mascolo C, & Rentfrow PJ (2016). Opportunities for smartphones in clinical care: The future of mobile mood monitoring. The Journal of Clinical Psychiatry, 77(2), el35 10.4088/JCP.15coml0054

Schlotz W (2011). Ambulatory psychoneuroendocrinology: Assessing salivary cortisol and other hormones in daily life In Mehl M & Conner T (Eds.), Handbook of research methods for studying daily life (pp. 193–209). New York, NY: Guilford.

Schmid Mast M, Gatica-Perez D, Frauendorfer D, Nguyen L, & Choudhury T (2015). Social sensing for psychology: Automated interpersonal behavior assessment. Current Directions in Psychological Science, 24(2), 154–160. 10.1177/0963721414560811

Sharma N, & Gedeon T (2012). Objective measures, sensors and computational techniques for stress recognition and classification: A survey. Computer Methods and Programs in Biomedicine, 108(3), 1287–1301. 10.1016/j.cmpb.2012.07.003 [PubMed: 22921417]

Sharmin M, Raij A, Epstien D, Nahum-Shani I, Beck JG, Vhaduri S, … Kumar S (2015). Visualization of time-series sensor data to inform the design of just-in-time adaptive stress interventions. Paper presented at the Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Sherman SM, Mumford JA, & Schnyer DM (2015). Hippocampal activity mediates the relationship between circadian activity rhythms and memory in older adults. Neuropsychologia, 75, 617–625. 10.1016/j.neuropsychologia.2015.07.020 [PubMed: 26205911]

Shiftman S, Stone AA, & Hufford MR (2008). Ecological momentary assessment. Annual Review of Clinical Psychology, 4,1–32. 10.1146/annurev.clinpsy.3.022806.091415

Silk JS, Steinberg L, & Morris AS (2003). Adolescents' emotion regulation in daily life: Links to depressive symptoms and problem behavior. Child Development, 74(6), 1869–1880. 10.1046/j.1467-8624.2003.00643.x [PubMed: 14669901]

Silk JS, Forbes EE, Whalen DJ, Jakubcak JL, Thompson WK, Ryan ND, … Dahl RE (2011). Daily emotional dynamics in depressed youth: A cell phone ecological momentary assessment study. Journal of Experimental Child Psychology, 110(2), 241–257. 10.1016/j.jecp.2010.10.007 [PubMed: 21112595]

Simons JS, Wills TA, Emery NN, & Marks RM (2015). Quantifying alcohol consumption: Self-report, transdermal assessment, and prediction of dependence symptoms. Addictive Behaviors, 50, 205–212. 10.1016/j.addbeh.2015.06.042 [PubMed: 26160523]

Sitnick SL, Goodlin-Jones BL, & Anders TF (2008). The use of actigraphy to study sleep disorders in preschoolers: Some concerns about detection of nighttime awakenings. Sleep, 31(3), 395 10.1093/sleep/31.3.395 [PubMed: 18363316]

Slatcher RB, & Robles TF (2012). Preschoolers' everyday conflict at home and diurnal cortisol patterns. Health Psychology, 31(6), 834 10.1037/a0026774 [PubMed: 22229929]

Slatcher RB, & Trentacosta CJ (2011). A naturalistic observation study of the links between parental depressive symptoms and preschoolers' behaviors in everyday life. Journal of Family Psychology, 25(3), 444 10.1037/a0023728 [PubMed: 21534668]

Slatcher RB, & Trentacosta CJ (2012). Influences of parent and child negative emotionality on young children's everyday behaviors. Emotion, 12(5), 932 10.1037/a0027148 [PubMed: 22390707]

Smith L, & Gasser M (2005). The development of embodied cognition: Six lessons from babies . Artificial Life, 11(1-2), 13–29. [PubMed: 15811218]

Smith LB, Jayaraman S, Clerkin E, & Yu C (2018). The developing infant creates a curriculum for statistical learning. Trends in Cognitive Sciences. 10.1016/j.tics.2018.02.004

Smith L, & Thelen E (2003). Development as a dynamic system. Trends in Cognitive Sciences, 7(8), 343–348. 10.1016/S1364-6613(03)00156-6 [PubMed: 12907229]

Smith LB, Yu C, Yoshida H, & Fausey CM (2015). Contributions of head-mounted cameras to studying the visual environments of infants and young children. Journal of Cognition and Development, 16(3), 407–419. 10.1080/15248372.2014.933430 [PubMed: 26257584]

Soderstrom M, & Wittebolle K (2013). When do caregivers talk? The influences of activity and time of day on caregiver speech and child vocalizations in two childcare environments. PLoS ONE, 8(11), e80646 10.1371/journal.pone.0080646 [PubMed: 24260443]

Soska KC, Adolph KE, & Johnson SP (2010). Systems in development: Motor skill acquisition facilitates three-dimensional object completion. Developmental Psychology, 46(1), 129 10.1037/a0014618 [PubMed: 20053012]

Sroufe LA (2012). Ritalin gone wrong (p. 29). New York Times.

Sroufe LA, Egeland B, Carlson EA, & Collins WA (2005). The Development of the Person: The Minnesota Study of Risk and Adaptation from Birth to Adulthood. New York, NY: Guilford Press.

Starner T, Mann S, Rhodes B, Levine J, Healey J, Kirsch D, … Pentland A (1997). Augmented reality through wearable computing. Presence: Teleoperators & Virtual Environments, 6(4), 386–398. 10.1162/pres.1997.6.4.386

Stiles J (2008). The fundamentals of brain development: Integrating nature and nurture. Boston, MA: Harvard University Press.

Stirman SW, & Pennebaker JW (2001). Word use in the poetry of suicidal and nonsuicidal poets. Psychosomatic Medicine, 63(4), 517–522. 10.1097/00006842-200107000-00001 [PubMed: 11485104]

Stone AA (2017). Challenges remaining for the field of real-time, self- reported data. Paper presented at the Society for Ambulatory Assessment, Esch-sur-Alzette, Luxembourg.

Stone AA, & Shiftman S (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. Annals of Behavioral Medicine, 24(3), 236–243. 10.1207/S15324796ABM2403_09 [PubMed: 12173681]

Sweetland Edwards H (2017). Alexa takes the stand: Listening devices raise privacy issues. Time Magazine.

Taguchi T, Tachikawa H, Nemoto K, Suzuki M, Nagano T, Tachibana R, … Arai T (2018). Major depressive disorder discrimination using vocal acoustic features. Journal of Affective Disorders, 225, 214–220. 10.1016/j.jad.2017.08.038 [PubMed: 28841483]

Tamis-LeMonda CS, Kuchirko Y, Luo R, Escobar K, & Bornstein MH (2017). Power in methods: Language to infants in structured and naturalistic contexts. Developmental Science, 20(6), el2456 10.1111/desc.12456

Tawari A, Martin S, & Trivedi MM (2014). Continuous head movement estimator for driver assistance: Issues, algorithms, and on-road evaluations. IEEE Transactions on Intelligent Transportation Systems, 15(2), 818–830. 10.1109/TITS.2014.2300870

Taylor P (2000). Distributed agency within intersecting ecological, social, and scientific processes In Cycles of contingency: Developmental systems and evolution (pp. 313–332). Cambridge, MA: MIT Press.

Thelen E, & Smith LB (1994). A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press.

Thomaz E, Essa I, & Abowd GD (2015). A practical approach for recognizing eating moments with wrist-mounted inertial sensing. Paper presented at the Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Thomaz E, Parnami A, Bidwell J, Essa I, & Abowd GD (2013). Technological approaches for addressing privacy concerns when recognizing eating behaviors with wearable cameras. Paper presented at the Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing.

Timmons AC, Baucom BR, Han SC, Perrone L, Chaspari T, Narayanan SS, & Margolin G (2017). New frontiers in ambulatory assessment: Big data methods for capturing couples' emotions, vocalizations, and physiology in daily life. Social Psychological and Personality Science, 8(5), 552–563. 10.1177/1948550617709115

Tobin ET, Kane HS, Saleh DJ, Wildman DE, Breen EC, Secord E, & Slatcher RB (2015). Asthma-related immune responses in youth with asthma: Associations with maternal responsiveness and expressions of positive and negative affect in daily life. Psychosomatic Medicine, 77(8), 892 10.1097/PSY.0000000000000236 [PubMed: 26407226]

Tokson M (2017). The Supreme Court's cell-phone-tracking case has high stakes. New Yorker Magazine.

Twenge JM (2017). Have smartphones destroyed a generation. The Atlantic, 3.

VanDam M, Warlaumont AS, Bergelson E, Cristia A, Soderstrom M, De Palma P, & MacWhinney B (2016). HomeBank: An online repository of daylong child-centered audio recordings. Paper presented at the Seminars in speech and language

Varkey JP, Pompili D, & Walls TA (2012). Human motion recognition using a wireless sensor-based wearable system. Personal and Ubiquitous Computing, 16(7), 897–910. 10.1007/S00779-011-0455-4

Vernon-Feagans L, Willoughby M, & Garrett-Peters P (2016). Predictors of behavioral regulation in kindergarten: Household chaos, parenting, and early executive functions. Developmental Psychology, 52(3), 430 10.1037/dev0000087 [PubMed: 26751500]

Verplaetse C (1996). Inertial proproceptive devices: Self-motion-sensing toys and tools. IBM Systems Journal, 35(3.4), 639–650. 10.1147/sj.353.0639

Wang R, Blackburn G, Desai M, Phelan D, Gillinov L, Houghtaling P, & Gillinov M (2017). Accuracy of wrist-worn heart rate monitors. JAMA Cardiology, 2(1), 104–106. 10.1001/jamacardio.2016.3340 [PubMed: 27732703]

Wang R, Harari G, Hao P, Zhou X, & Campbell AT (2015). SmartGPA: How smartphones can assess and predict academic performance of college students. Paper presented at the Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing.

Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, … Campbell AT (2014). StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones. Paper presented at the Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing.

Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, … Campbell AT (2017). StudentLife: Using smartphones to assess mental health and academic performance of college students In Mobile health (pp. 7–33). Berlin, Germany: Springer.

Wang Y, Hartman M,Aziz NAA, Arora S, Shi L, & Tunison E (2017). A systematic review of the use of LENA technology. American Annals of the Deaf, 162(3), 295–311. 10.1353/aad.2017.0028 [PubMed: 28824021]

Wang Z, Pan X, Miller KF, & Cortina KS (2014). Automatic classification of activities in classroom discourse. Computers & Education, 78, 115–123. 10.1016/j.compedu.2014.05.010

Warlaumont AS, Richards JA, Gilkerson J, & Oiler DK (2014). A social feedback loop for speech development and its reduction in autism. Psychological Science, 25(7), 1314–1324. 10.1177/0956797614531023 [PubMed: 24840717]

Wass S, de Barbara K, & Clackson K (2015). Tonic and phasic co-variation of peripheral arousal indices in infants. Biological Psychology, 111, 26–39. 10.1016/j.biopsycho.2015.08.006 [PubMed: 26316360]

Watson D, Clark LA, & Tellegen A (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54(6), 1063–1070. 10.1037/0022-3514.54.6.1063 [PubMed: 3397865]

Weiser M (1991). The computer for the 21st century. Scientific American, 265(3), 94–105. 10.1038/scientificamerican0991-94

Weisleder A, & Fernald A (2013). Talking to children matters early language experience strengthens processing and builds vocabulary. Psychological Science, 24(11), 2143–2152. 10.1177/0956797613488145 [PubMed: 24022649]

Westeyn TL, Abowd GD, Starner TE, Johnson JM, Presti PW, & Weaver KA (2012). Monitoring children’s developmental progress using augmented toys and activity recognition. Personal and Ubiquitous Computing, 16(2), 169–191. 10.1007/S00779-011-0386-0

Weusthoff S, Baucom BR, & Hahlweg K (2013). Fundamental frequency during couple conflict: An analysis of physiological, behavioral, and sex-linked information encoded in vocal expression. Journal of Family Psychology, 27(2), 212 10.1037/a0031887 [PubMed: 23458696]

Wilhelm FH, & Grossman P (2010). Emotions beyond the laboratory: Theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment. Biological Psychology, 84(3), 552–569. 10.1016/j.biopsycho.2010.01.017 [PubMed: 20132861]

Wilhelm P, Perrez M, & Pawlik K (2012). Conducting research in daily life In Mehl MR & Conner TS (Eds.), Handbook of research methods for studying daily life (pp. 62–86). New York, NY: Guilford Press.

Wyatt D, Choudhury T, & Bilmes J (2007). Conversation detection and speaker segmentation in privacy-sensitive situated speech data. Paper presented at the Eighth Annual Conference of the International Speech Communication Association.

Wyatt D, Choudhury T, Bilmes J, & Kitts JA (2011). Inferring colocation and conversation networks from privacy-sensitive audio with implications for computational social science. ACM Transactions on Intelligent Systems and Technology (TIST), 2(1), 7 10.1145/1889681.1889688

Yao X, Ploetz, T, Johnson M, & deBarbaro K (under review). Automated detection of infant-holding and -carrying behaviors via body-worn motion sensors. Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Yatani K, & Truong KN (2012). BodyScope: A wearable acoustic sensor for activity recognition. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.

Yu C, Yurovsky D, & Xu TL (2012). Visual data mining: An exploratory approach to analyzing temporal patterns of eye movements. Infancy, 17(1), 33–60. 10.1111/j.1532-7078.2011.00095.x

Zhang H, De Choudhury M, & Grudin J (2014). Creepy but inevitable? The evolution of social networking. Paper presented at the Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing.

Zhang X, Pina LR, & Fogarty J (2016). Examining unlock journaling with diaries and reminders for in situ self-report in health and wellness. Paper presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, Santa Clara, California, USA.

Zhang YC, & Rehg JM (2018). Watching the TV watchers. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2(2), 88 10.1145/3214291

Zimmerman FJ, Gilkerson J, Richards JA, Christakis DA, Xu D, Gray S, & Yapanel U (2009). Teaching by listening: The importance of adult-child conversations to language development. Pediatrics, 124(1), 342–349. 10.1542/peds.2008-2267 [PubMed: 19564318]

Zisook M, Hernandez J, Goodwin MS, & Picard RW (2013). Enabling visual exploration of long-term physiological data. Paper presented at the Proceedings of the 2013 IEEE Conference on Visual Analytics Science and Technology.

TABLE 1

List of salient developmental constructs across the lifespan, paired with theoretically-relevant daily activities and sensing solutions. All mention of annotation refers to coding of data by trained human annotators; all mention of “detection” refers to automated coding via trained machine learning algorithms

Developmental construct	Relevant behaviors/exposures	Sensing possibilities	Sample studies
Temperament & Personality	Activity Daily behavior, routines	Amount of motion from accelerometer Annotated from wearable camera; EMA	Buss et al. (1980) Brown et al. (2017)
Child stress and emotional adaptation	Child mood & distress Regulation Context of child distress	Early: Crying or laughing detected from wearable audio Later: EMA Early: Duration to soothe detected via wearable audio; Later: EMA Sporadic annotation (e.g. time locked to tantrums detected from wearable audio)	Rao et al. (2014); Kim and Clements (2015) Gunthert and Wenze (2011); Silk et al. (2011) Ludwig (2018); Silk, Steinberg, and Morris (2003)
Caregiving characteristics	Tone of interactions Physical presence and contact Joint activity Contingent language Parental Sensitivity	Detected from wearable audio Detected from bluetooth or wearable motion sensors Annotated reading or playing; detected mutual gaze from wearable camera or audio Detected via LENA wearable audio algorithms Sporadic annotation (e.g. event-locked to distress detected from wearable audio or within audio “snippets”)	Black et al. (2013); Weusthoff et al. (2013) Olguin et al. (2009); Yao et al. (under review) Rehg et al. (2013); Soderstrom and Wittebolle (2013) Zimmerman et al. (2009) Tobin et al. (2015)
Parental stress	Mood Neurobiological stress reactivity	EMA Ambulatory cortisol, wearable physiology monitors	Gunthert and Wenze (2011) Schlotz (2011); Hovsepian et al. (2015)
Household characteristics	Chaos of home environment Household routines Food availability	Detected via wearable or static audio or camera Regularity of meal-time detected via water or electrical meter Wearable camera	Cook etal. (2018) Froehlich et al. (2009); Gupta et al. (2010) Jia etal. (2018)
Early perceptual development	Visual and auditory experience Gross motor experience Fine motor experience	Wearable cameras and audio recorders Posture and fall detected via wearable motion sensors Wrist-worn motion sensors; toys with embedded motion sensing	Smith et al. (2015) Nam and Park (2013) Varkey, Pompili, and Walls (2012); Verplaetse (1996); Westeyn et al. (2012)
Language learning	Language input and contingency of parental speech	Detected via wearable audio (e.g. via LENA algorithms)	Zimmerman et al. (2009)
Attention & executive function	Daily Cognitive assessments Time spent distracted in attentional focus	Active assessment collected via mobile phone app Detected via wearable motion sensors Cellphone use monitoring application	Maekawa, Anderson, de Brecht, and Yamagishi (2018) Tawari et al. (2014) Wang, Chen et al. (2014)
Peer, teacher, or family interactions (incl. e.g. marital or sibling conflict)	Quantity of interactions Classroom activity Quality of interactions	Contingent speech detected via wearable audio Proximity detected via wearable Bluetooth sensors Detected via static audio recordings	Berke et al. (2011) Olguin et al. (2009) Wang, Pan, Miller, and Cortina (2014) Slatcher and Trentacosta (2012); Tobin etal. (2015)

Developmental construct	Relevant behaviors/exposures	Sensing possibilities	Sample studies
		Conflict or supportiveness annotated from snippets via wearable audio recorder	
Puberty	Hormonal changes	Ambulatory assessment of saliva samples	Schlotz (2011)
Identity formation	Visible advertising or stereotypes	Wearable camera with some automated assistance	Zhang and Rehg (2018)
Academic success	Time spent studying Afterschool activities Classroom engagement	Detected via multimodal cellphone sensor monitoring Geocoded from GPS data Detected via pressure-sensitive seats, motion and physiological sensors	Wang, Chen et al. (2014) Byrnes et al. (2017) Dragon et al. (2008)
Risky activity (sex, drug abuse)	Frequency of content words Alcohol or other substance use Common activity by location	Transcription of snippets from wearable audio Detected via multimodal cellphone sensor monitoring; specialized biomarker sensors Geocoded from GPS data	Pennebaker, Mehl, and Niederhoffer (2003) Bae et al. (2017);Bertz et al. (2018) Byrnes et al. (2017)
Aging	Cognitive and physical decline	Detected via static multimodal platforms	Kaye et al.(2011)
Physical & mental health	Psychiatric symptoms Sleep; physical activity; eating episodes	EMA Detected via motion sensors	aan het Rot, (2012) Bussmann and Ebner-Priemer(2012); Robillard et al. (2015); Thomaz et al. (2015)
Other relevant exposures	Phone use Nutrition Sugar consumption	Cell phone monitoring application Annotated via participant-collected photo logs Continuous glucose monitoring sensors	Harari et al. (2017) Cordeiro et al. (2015) Juvenile Diabetes Research Foundation (2008)