

@BabySteps: Design and Evaluation of a System for using Twitter for Tracking Children’s Developmental Milestones

Hyewon Suh, John R. Porter, Alexis Hiniker, Julie A. Kientz

Human Centered Design and Engineering

University of Washington

{hyewon25, jrporter, alexisr, jkientz}@uw.edu

ABSTRACT

The tracking of developmental milestones in young children is an important public health goal for ensuring early detection and treatment for developmental delay. While numerous paper-based and web-based solutions are available for tracking milestones, many busy parents often forget to enter information on a regular basis. To help address this need, we have developed an interactive system called @BabySteps for allowing parents who use Twitter to track and respond to tweets about developmental milestones using a special hashtag syntax. Parent responses are parsed automatically and written into a central database that can be accessed via the web. We deployed @BabySteps with 14 parents over a 3-week period and found that parents were able to learn how to use the system to track their children’s progress, with some using it to communicate with other parents. The study helped to identify a number of ways to improve the approach, including simplifying the hashtag syntax, allowing for private responses via direct messaging, and improving the social component. We provide a discussion of lessons learned and suggestions for the design of interactive public health systems.

Author Keywords

Children; health; microblogging; Twitter; social media; data capture; memories; parents; public health.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

The tracking of developmental progress in young children is one of the key ways of ensuring the early detection of developmental delay. Typically, the earlier these children are diagnosed, the sooner treatment can begin, which can greatly improve outcomes for these children [7]. One cur-

rent approach for screening for developmental progress is to ask parents a series of questions about their children’s abilities in areas such as gross motor, communication, and problem solving. For example, “does your baby say two syllable sounds, such as ‘da da’ or ‘ga ga’?” or “can your child walk up stairs while holding onto a railing?” Standardized screening questionnaires administered by doctors and public health clinics may consist of many of these questions periodically. One popular screener tool, the Ages and Stages Questionnaire (ASQ), consists of 22 questionnaires with 30 questions each across the child’s first five years [19]. The screener can either be conducted via paper or through a website, and is usually administered at Well Child Visits with the child’s pediatrician. There are some limitations to using milestone questionnaires for screening, such as child variability or the potential for parent anxiety. However, they are currently the most effective method for screening according to the American Academy of Pediatrics [4].

The problem with paper-based or even web-based screener surveys is that it is often difficult to ensure that parents complete each survey, especially since there are more screener questionnaires than scheduled Well Child Visits during the first 5 years of a child’s life. In addition, when parents answer questions in a single sitting, they may not have the opportunity to try each activity with their child if the question asks about an activity that does not always occur. Also, busy parents often just forget to keep track of milestones using both paper and web-based systems [12]. Tools that can help prompt parents to answer questions at more regular intervals can possibly result in more complete and accurate responses and better parent involvement and awareness of their child’s activities. The CHI community has recently begun developing such software (e.g., [11]).

There are limitations with standalone software applications in that they require the parent to remember to open the application to answer questions. Even with proactive notifications or email reminders, parents may still not engage as easily and as often as they might otherwise. In addition, because these tools are intended to be used over a long period (e.g., up to 5 years), they require methods to maintain parent engagement. To help address some of these limitations, one solution is to design and develop tools to help parents answer questions using technology with which they are already familiar and engaged with, such as mobile

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada

Copyright is held by the owner/author(s). Publication rights licensed to ACM. ACM 978-1-4503-2473-1/14/04...\$15.00.

<http://dx.doi.org/10.1145/2556288.2557386>

phones or social media. A recent study [16] suggests that integrating a health intervention into social sites that people already visit can increase adherence. Although parents are often too busy to remember to complete baby books or other manual tracking systems [11], they are already using social networking systems to share information on their children [15]. Inspired by earlier studies, our first approach has been to develop a tool to automatically prompt parents via Twitter to respond to questionnaires about their child, which can be stored in a centralized database.

In our system, @BabySteps, we use a specified syntax of hashtags to allow parents to respond to age-appropriate milestone questionnaires automatically tweeted by accounts based on a child's birth month. A script then uses the Twitter API to read parent responses and write them to a database. The child's developmental progress can be viewed via a companion website we have developed. We conducted a 3-week feasibility study of @BabySteps with 14 parents of children aged 2 months to 2 years. Parents responded to milestone prompts via Twitter and then could access a website to review their child's progress. The study explored how the specialized syntax could be used to create an interactive public health system and encourage parents to continue engagement. The results of the study helped identify areas for improvement and implications for the design of public health applications using Twitter or other social media. The main contributions of this work are the design and development of a novel system for using Twitter to track developmental progress in young children and a study understanding how these types of systems are accepted by users, which leads to design considerations.

RELATED WORK

@BabySteps builds upon previous work on supporting the needs of parents with young children and leveraging technologies and social media for health and safety.

Designing for the Health of Young Children

Designing interactive technologies to support the health and wellbeing of young children and their families has been a growing area of interest in the field of human-computer interaction. A workshop called "Motherhood and HCI" was held at CHI last year, showing an increase in interest in supporting parents and the health issues of their children. D'Alessandro & Dosa [2] describe how information technology can be empowering to parents and children, and thus there is great motivation to use it to support their health.

The motivation and design of @BabySteps was inspired directly by a formative study on designing for tracking developmental progress in young children [12] and the subsequent Baby Steps software application [11]. We extend this work by developing and studying a Twitter-based application that can link into a website component modeled after the software application. Similar applications have tracked other aspects of health of young children. The ENSURE system developed by Park et al. [18] looked at how interactions with a child's pediatrician can be supported through

interactive technologies, such as tracking temperature or medications. Tang et al. [21] developed a mobile application called Estrellita for observing premature infants' daily living activities (such as feeding and diaper changes), prompting parents to do exercises at home, and reminders of upcoming appointments. VisiBabble is a tool for monitoring emerging speech in young children [5], whereas Child's Play uses sensor-enabled toys to monitor children's development [24]. Similar commercial applications, such as Trixie Tracker [22] and LENA [15] also allow for parents to track information about feeding and diapers and tracking language respectively. While these applications all have the goal of tracking information about young children, including their development, @BabySteps is the first application we are aware of to use Twitter in an interactive format.

Interactive Technologies & Social Media

In recent years, public health outreach efforts have benefited greatly from the large-scale communication afforded by SMS messaging. With the rising prevalence of prepaid mobile plans, public health campaigns integrating SMS are able to communicate with populations who have traditionally been difficult to reach as a result of scale, geographic dispersal, limited technology use, or other concerns. For example, beginning in 2003, the telecommunication company MTN began sending out mass quarterly informational SMS messages to all of its subscribers in Nigeria in an effort to promote HIV/AIDS awareness [23]. Similar projects like the outreach campaign from Hoefman & Apunyu in Uganda used two-way communication, interspersing 'quiz' and survey questions amongst informational messages to which recipients could respond [8].

Juxtaposing this sort of 'blanket' approach to SMS outreach regarding a ubiquitous health concern, similar programs have more recently been implemented to target more specific populations. Jordan et al.'s Text4baby system used an opt-in model, wherein expectant mothers signed up for the service via an initial SMS, after which they received three informational SMS messages per week about their pregnancy and related health concerns [10]. By using an opt-in model where users provide relevant personal information during an initial phase of interactivity, it is possible to tailor outgoing content specifically for the recipient. In the case of Text4baby, provision of the user's due date enabled sending of time-appropriate information based on pregnancy trimester. Woolford et al. used a more complex approach in their development of an SMS system to encourage adherence to a weight management program among obese adolescents; a detailed initial survey was used to tailor the set of potential messages a user might receive to optimally appeal to their values, motivations, and goals [25].

Meanwhile, the Internet and social networking platforms have facilitated numerous different efforts related to health and wellness. Numerous online communities and information hubs have been developed to provide direct and indirect counseling and psychotherapeutic aid to those in

need [1]. De Choudhury et al. developed a system which could assess the public tweets of women to predict with some confidence those most likely to undergo significant postpartum emotional or behavior change [3]. While this work demonstrated that health data of this type could be gathered on a large scale via social media, it is limited in the nature of the data available to it. Although psychological analyses of this sort can be based on sentiment gleaned from any user-authored text, more specific health information is unlikely to spontaneously appear in social media at all, let alone in workable form. For that reason, it is necessary for the user to be made aware of not only the nature of the desired content, but also the form.

While the latter might seem a massive obstacle given the freeform nature of social media postings, recent work has shown that Twitter in particular has a great deal of promise. The Tweak the Tweet project has demonstrated on an ongoing basis that, even in dynamic and high-stakes contexts, populations of Twitter users could be made to quickly adopt an efficient standardized syntax for the sharing of specific pieces of information, such as needs and availabilities during crisis events [20].

When considering the landscape of these outreach and coordination systems, a useful way to classify them is according to two distinct dimensions: the engagement level of the end user or users, and the communicative relationship being encouraged or supported. Figure 1 marks the position of some of the aforementioned major systems on these dimensions. Its horizontal axis represents relationship and direction of information flow, ranging from a one-to-many model where a single agent outwardly communicates with a large population, to a many-to-one model in which the constituents of a population channel a large amount of information towards a single agent; the midpoint can be seen as either one-to-one or many-to-many, as these relationships are functionally similar in this context. The vertical axis denotes engagement level and ranges from passive to interactive.

Viewing the existing work through this lens, the relative

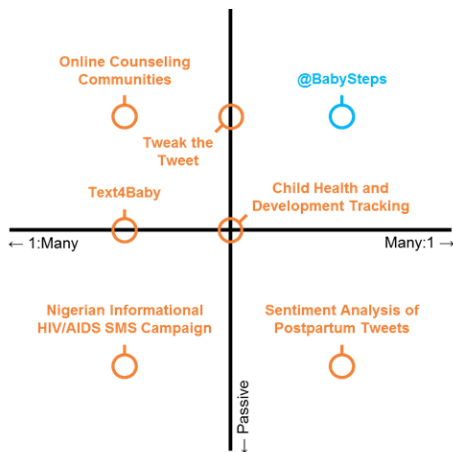


Figure 1: Space of interactive technology for public health

emptiness of the upper right quadrant suggests the untapped potential of a system that seeks to *interactively* relate to users on a *many-to-one* basis – that is, a system that explicitly solicits a population to provide it with specific information. @BabySteps seeks to occupy this region and test its viability.

@BABYSTEPS SYSTEM DESIGN

In this section, we describe the design and implementation details for @BabySteps. The Twitter-based system was designed to be a component of a larger suite of tools for allowing parents to respond to milestone question prompts, which also includes a website and a mobile text messaging system. A centralized database stores all of the parent information and responses. The design of @BabySteps was based on a series of formative studies ([6,9,12]), with additional input from a study on the use of social networking by mothers of young children [16].

Functionality and Usage

To use the @BabySteps system, users first register in our system with their name, email address, and Twitter handle. For each of their children, they provide a name, gender, and date of birth. If a user is uncomfortable with their child’s real first name being used in public communications, they are able to provide a pseudonym that can be used instead. After registration, a user follows the main Baby Steps Twitter account (@BabyStepsUW), which is used to send out announcements and study information to all participants and is manually operated by the research team. They also follow a Twitter account associated with their child’s birth month, and if applicable, grant permission for it to follow them back. So, for example, a user with a child born in August 2011 would follow @BabyStepsAug11.

At regular intervals, this latter account sends out tweets asking age-appropriate milestone questions about the child’s development. Each of these tweets contains a unique closing hashtag that we refer to as a milestone ID (see Figure 2 for an example).



Figure 2: Exemplary milestone question tweet

Users respond to these questions either publicly by posting to their timeline or privately via direct message to the account that originally posted the question. In either case, the user’s response must contain a minimum of two specific elements: 1) the milestone ID hashtag from the prompt they are responding to and 2) their response to the question in the form of a hashtag – either #yes, #sometimes, or #no / #notyet (these two are handled identically). For users who have multiple children registered in our system, they must additionally include the relevant child’s first name or pseudonym as a hashtag. Beyond these required elements, users

are free to include any other free text or additional hashtags in their tweet as well. This allows them to naturally include their response in an unassuming tweet that makes sense to their own followers, should they choose to do so.

If all required pieces are included, such as in Figure 3, their response to the milestone will be logged in our database.

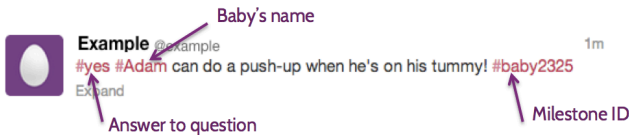


Figure 3: Exemplary response from users

Some milestones have figures that go along with them, such as showing a close up of a child doing a pincer grasp. We uploaded figures to pic.twitter.com and included a link in the prompt for relevant milestones. In addition to specific responses to our prompts, users are also able to create a free-form “memory” about their child by including the hashtag *#babymemory* in a tweet (e.g., *#Billy* visited the zoo today! *#babymemory*). When doing so, the remainder of the text in the tweet will be included as a memory on their child’s Baby Steps timeline on the companion website. Also, because there are only 30 milestones in a 2-3 month period, we had static tips and activities that were tweeted on regular basis, such as advice about vaccinations and ideas for activities that encourage development. Tips and activities used the *#tips* and *#activity* hashtags respectively.

Twitter System Implementation

The Twitter system was developed in Python, and consists of a series of scripts which run on the server that also houses the website. MySQLdb, a Python MySQL connector module, is used to interface with a centralized database shared across different pieces of the Baby Steps suite. Twitter interaction is achieved by making calls to Twitter’s REST API v1.1. @BabySteps uses a collection of automated Twitter accounts, each of which have authorized full read/write access to our registered Python application. These accounts are each named with a given month and year combination, e.g. @BabyStepsAug11, @BabyStepsNov12. Generally speaking, our system can be broken into two general pieces: the “push” component and the “pull” component. Their basic operation is described as follows.

Push Component

Three times per day, during normal daytime hours in our local time zone, this component of our system is executed via a cron job. It iterates through each of the authorized month-specific accounts sequentially. For each account, we identify the integer ID of the last milestone question that was tweeted out. This ID is then incremented, and the next milestone in the sequence is checked in the database. Each milestone has an age range for when it is appropriate to be asked. The age implied by the month and year in the Twitter account’s handle is then compared against this range. If the new milestone is age-appropriate, it is sent out as a new tweet, with the hashtag *#baby[ID]* appended to the end,

where [ID] is the milestone’s integer ID. If all relevant milestones for the age bracket have been exhausted, the system sends out a random age-appropriate activity or tip.

Pull Component

This script executes every half-hour around the clock and also iterates through our accounts sequentially. It begins by identifying the unique Twitter-assigned ID of the last incoming tweet that was received. All public tweets from the users followed by and following this account, as well as all direct messages, created since this last check ID are then collected for processing. From the raw text of each individual tweet, the set of hashtags are extracted. To identify milestone responses, we look for tweets whose hashtag sets contain the minimum set of required elements outlined previously. If all required components are present, the user’s response to the milestone prompt is recorded in our database. Alternately, if the tweet contains the hashtag *#babymemory*, the tag is stripped from the raw text, and the remainder is inserted into the user’s log of baby memories. In either case, the entry is time stamped according to when the user tweet was originally authored.

Companion Website

In addition to the Twitter-based system, Baby Steps also had a companion website where participants could log on to edit milestone responses, view a timeline of their child’s progress and memories, and see an assessment of their child’s progress based on their responses (e.g., whether they are on track developmentally or if they should get in touch with a doctor for additional assessment). The website also had a frequently asked questions page with information on how to use the system and a resources page that had contact information for public health resources to encourage parents to get in touch with their doctor if there is any concern about their child’s development. The eventual goal of the Baby Steps work is to have our state’s public health department conduct follow-up phone calls with parents based on their response to milestone prompts, but this system is not yet in place. Thus, the website served as an intermediary for parents to understand their child’s progress.

DEPLOYMENT STUDY DESIGN

To test the feasibility of the @BabySteps system design, understand the usability of our hashtag syntax, and to gauge the acceptability of using Twitter for sharing a child’s milestone progress, we conducted a real world deployment study with parents of young children. We conducted a qualitative, exploratory study, as Klasnja et al. have argued is the best first step for evaluating novel health systems [13] along with some quantitative data logging. This section includes the details of the study design and the participants we recruited.

Study Procedure

Before deploying @BabySteps, we sent out a pre-study survey asking participants about their demographics, Twitter usage, and parenting practices. With this information, we manually registered our participants in our system.

Stage 1: Two Weeks of @BabySteps Usage on Twitter

Because our Twitter system required use of special syntax in tweet responses, we sent out Twitter instructions to participants before Baby Steps' monthly accounts started tweeting out. On the instruction page, there was information about how to use the three hashtags and information about how to follow the Baby Steps main account and monthly accounts. Also, because we acknowledge that privacy can be of concern to parents, we encouraged our participants to check their account's privacy settings. We also did not limit the form of response – it could be a reply to our tweet, a new tweet, a re-tweet, or a direct message. That way, participants have the ultimate say in what they want to share publicly vs. privately. Then, for two weeks, each monthly account tweeted out developmental milestones, a set of functional skills or tasks that most children can do at a certain age range. Based on their observation, participants responded 'Yes', 'Sometimes', 'Not Yet / No' accordingly. Occasionally, both the main account and monthly accounts sent out reminders for syntax and study-related announcements, as inspired by Tweak the Tweet [20].

Stage 2: One Week of @BabySteps plus Website Usage

After participants used @BabySteps and responded for two weeks, we made the Baby Steps companion website available for parents to view their child's progress based on milestones they answered via Twitter. The goal of this study was not to test the website itself, although participants were welcome to provide feedback on the system's design. For this study, the script only parsed tweets which used the right syntax. Due to this, not all of the participants' Twitter responses were successfully recorded on the website. However, because the progress report may not accurately show child developmental progress until all 30 milestones are answered for a certain age range, we notified participants

about this fact and encouraged them to answer any milestones that they have missed via the website before reviewing the progress report. To reduce any unnecessary parental stress, we explained in recruitment materials that development is variable with a wide range of what is typical as children grow in different areas at different rates. Participants had approximately one week to use our website before we deployed a post-study survey, though they were free to use the website indefinitely after our study completed. We did not open up the site initially since we wanted to study how parents used Twitter for responding to milestone prompts before they had a chance to answer questions via the website. The post-study survey included questions about the frequency of tweets, the ease of learning the hashtag system, any concerns about sharing child developmental information publicly, and open-ended questions about how to improve the design of @BabySteps.

Participants

As the main body of our study occurred in Twitter, we recruited our participants via Twitter. The research team tweeted a short study description and a link to a screener survey to their personal accounts and asked for people to re-tweet. Some followers re-tweeted or to forwarded it to a relevant person, with the majority of participants coming from a re-tweet by a person with nearly 20,000 followers (who also enrolled in our study as P5). The screener survey asked if they were parents of a baby between 2 months and 2 years, the number of children in the household, their experience with Twitter, and optional demographic information. We received 72 responses to the screener survey. Of those that responded to the screener, we sent invitations to approximately 40 participants to recruit a large, diverse set. Of the invitations sent, a total of 19 parents ended up consenting, completing the pre-study, and enrolling in the

Table 1: Participant details

ID	Age	Gender	# of children	Years Using Twitter	# of Twitter accounts	# Followed	% Personal acquaintance	# of Followers	% Personal acquaintance	Protected tweets	Twitter Frequency	Baby's Initial Age(s) in months	Baby's Gender(s)
P1	25 - 30	F	1	> 3 years	1	329	30-50%	226	< 10%	No	Several per day	21	F
P2	25 - 30	F	2	> 3 years	1	772	< 10%	329	< 10%	No	Several per day	4, 24	F, F
P3	25 - 30	M	1	> 3 years	3+	402	50-70%	602	30-50%	No	Several per day	12	M
P4	25 - 30	F	1	> 3 years	2	5	> 70%	5	> 70%	Yes	Once per 2 weeks	13	F
P5	25 - 30	F	1	> 3 years	1	8,169	< 10%	19,736	< 10%	No	Several per day	13	M
P6	31 - 40	F	2	1 - 3 years	2	175	< 10%	235	< 10%	No	Several per day	8	F
P7	31 - 40	F	1	> 3 years	2	577	< 10%	1598	< 10%	No	Several per day	22	M
P8	31 - 40	F	2	1 - 3 years	1	277	< 10%	214	< 10%	No	Several per day	16	F
P9	31 - 40	F	5	> 3 years	1	385	30-50%	269	30-50%	No	Several per day	5, 5	F, M
P10	31 - 40	F	1	1 - 3 years	1	150	< 10%	80	< 10%	No	Several per day	16	M
P11	31 - 40	F	2	1 - 3 years	1	305	10-30%	68	> 70%	Yes	Once per day	13	M
P12	31 - 40	F	1	> 3 years	1	795	< 10%	364	< 10%	Yes	Several per day	10	F
P13	31 - 40	F	2	> 3 years	3+	~900	10-30%	~4,000	10-30%	No	Several per day	22	M
P14	31 - 40	F	2	1 - 3 years	2	~100	10-30%	50	50-70%	No	Several per day	20	F

study. During the course of the study, one decided to withdraw because of concerns over sharing tweets about her child publicly on Twitter, and two people never sent tweets to @BabySteps and thus dropped out. Two participants replied to via Twitter to milestone posts, but did not respond to our post-study survey (P1 and P7). Overall, we have @BabySteps usage data from 14 participants and 12 responses from the post-study survey (see Table 1).

The participants were a mix of demographics from all over the United States, though most were younger as they had young children. Since we recruited via Twitter, most of our participants had good experience with it. Six participants had multiple Twitter accounts and nine participants have used Twitter more than 3 years. All participants currently record information in varying ways about their child, and many often bring notes to doctors. However, lack of time and forgetting to record were the most common difficulties they face in keeping records regularly, which makes them ideal candidates for @BabySteps. Two participants (P2 and P9) used @BabySteps with their two children, with P9 having fraternal twins. All participants consented via the web and received a \$20 Amazon.com gift card as a token of appreciation. See Table 1 for participant details.

RESULTS

We here report findings from the @BabySteps deployment study on how participants used our system, learned the pre-defined syntax, reacted to privacy issues, interacted with one another, used the website in complement to Twitter, and their overall reaction to the system.

Overall Tweeting Activity

For this study, Baby Steps set up twelve accounts including the main Baby Steps account. Over the three weeks, each account associated with a baby’s birth month tweeted out age-appropriated milestone questions as well as useful parenting tips (#tip) and activities (#activity) that parents can try with their children. A total of 670 tweets were sent (5 from the main account) averaging approximately 60.5 tweets among eleven monthly accounts (average 2.88 tweets per day). The composition of those 60.5 tweets was 42.1 milestone questions, 8.1 tips, 8.4 activities, and 1.9 study-related announcements (e.g. reminders to use pre-defined hashtags, functionalities, etc.). Depending upon the age range, there could be slight difference in the number of milestone questions, tips, and activities. In response to our tweets, 14 participants made 319 milestone responses (22.3 per participant, see Table 2).

We also wanted to know what participants thought about the frequency and timing of tweets made by Baby Steps. In this study, an average of 2.88 tweets were sent out at random times per day, rather than tweeting at set intervals, to be more noticeable by participants checking Twitter at any time of the day. Eight participants thought this frequency was just right, while four thought it was too often. However, when it came to tweeting each tweets at different times, we received mixed opinions.

P6: “(...) whenever I checked twitter I usually would see at least one tweet from baby steps, which I really liked.”

P12: [in response to how to improve system] “Sending out the tweets at designated times every day (or even repeating them) so I would know to expect them.”

Our system was designed to tweet parenting tips and activities after all milestone questions for an age group had been exhausted. One participant expressed that she would prefer if the system sends out them in parallel.

P5: “The ideal frequency would be a balance of the milestones with the tips. Some days were just tips. Others were just milestones. It seemed unbalanced, hard to predict, irregular.”

From our pre-study survey, we found out that all participants had their own ways of recording child development and are currently keeping records of it. The most common difficulties they face are lack of time and forgetting to record. As Twitter has a “Favorite” feature, some participants used it to save some of our tweets for later.

P5: “Because I checked the Twitter when [my baby] was asleep or not with me, and it was a milestone activity I hadn’t tried yet with him. I wanted to wait until I could try it, and sometimes favorited the milestone so I’d remember to.”

Table 2: Overall usage by the 14 participants in our study. The sparklines to the right are meant to give an overview of activity over 3 weeks. P5 started the study later than others.

ID	Total Tweets	Avg. per Day	Tweet Frequency by Time
P1	6	0.3	
P2	9	0.4	
P3	31	1.3	
P4	28	1.2	
P5	22	1.0	
P6	10	0.4	
P7	25	1.1	
P8	14	0.6	
P9	26	1.1	
P10	47	2.0	
P11	26	1.1	
P12	23	1.0	
P13	12	0.5	
P14	33	1.4	
Avg.	22.3	1	

It was more common to favorite #tip or #activities rather than milestone questions, but overall favoriting was low.

Using Twitter Syntax

For our system to correctly recognize and store user responses, we required three hashtags (in any order) to be included in each milestone response: 1) answer to the milestone question, 2) baby's name or pseudonym, and 3) milestone ID. Hashtags are very common on Twitter, however, they are generally created by users organically. We thought it might be new for participants to be required to include specific hashtags for a certain purpose and might need some training. Thus, we sent out instructions (via a website link) and occasional reminder tweets about the syntax. As a result, among 319 milestone responses, 198 (62%) were correctly formatted and 121 (38%) were missing some parts.

There were a few cases (8 tweets) where the response had different answers (e.g. maybe, rarely don't know, not sure, etc.) instead of the given answer options (yes/sometimes/no/notyet) but the majority of errors came from missing the # sign before answers to the question or a missing milestone ID. For example:

- @BabyStepsAug12 #baby2464 yes. Can turn pages himself [missing # before 'yes']
- @BabyStepsDec12 #yes so cute! [missing milestone ID]

Because of reminder tweets, most of the participants learned the correct syntax over time, but some participants expressed difficulties following the rules.

P2: "It [Syntax] was good but a little complicated - would have been nice to not have to do a hashtag for all responses, as that doesn't seem like a typical way to use hashtags. As in, you would never normally do a hashtag in front of 'yes' or 'no'. I had to keep going back to the email with instructions to remember what format you wanted the responses in."

The most preferred method of responding to milestone questions was via the Twitter mobile phone or tablet application (54% of all responses), which made it more difficult to follow the syntax. Approximately 28% of the posts were via the Twitter website, and another 11% were either via a special app like TweetDeck or TweetBot for iOS and an additional 7% were of unknown origin (due to being direct messages). Four participants reported they used only the mobile phone application, whereas there was only one participant who solely used the Twitter website.

P11: "It was difficult to remember to use the correct syntax, period, but probably a bit harder on the phone."

P9: "Using it from mobile, I had to remember the correct hashtag (ex. #baby1234) because it was not visible when I was typing. This would not be the case with Twitter on the web, but I most often use my phone to tweet."

There was one interesting case of the participant who was responding to milestone tweets for her two 5-month old twins. She responded to the milestone prompt for both children within the same tweet (@BabyStepsMar13 #yes #ba-

by2325 #babyname1 #babyname2). Our system could not handle this syntax, so this was not parsed correctly. She eventually split up the tweets into separate ones for each.

Tweeting in Public

Twitter, unless the user explicitly chooses to go private, is placed in the public domain and whatever the user tweets will be broadcasted to the world at large.

Privacy

As the main body of the study occurs on Twitter, privacy issues naturally came to the surface. Some participants expressed that they were completely fine with sharing child developmental progress on Twitter.

P13: "I don't mind publicly responding. I already do a lot of public responses, and tweets about my children. Response did not affect my preference."

P12: "My account is private but I don't mind responding publicly. My child is meeting milestones as expected. Even if she weren't I wouldn't mind responding publicly."

However, not everyone agrees on that. In fact, one of the participants withdrew from the study because she did not feel comfortable in sending tweets with her baby's developmental information, even though her timeline is protected. She expressed that although her account is protected and private, she was concerned about tweeting negative answers to milestone questions. Also, P5, who responded via Direct Message (DM), expressed a similar privacy concern before consenting. After the study, she added her reasons for responding via DM rather than public tweets and frustration she had gone through because of her choice.

P5: "I prefer to respond privately, because a public message is a) incomprehensible to anyone viewing my public timeline and b) runs the risk of giving peering strangers more information about my baby than I care to give."

P10: "I didn't use DM b/c I thought it was too complicated on my phone. (...) I responded publicly b/c it was easier on my phone. But I didn't put a lot of identifying info. Just yes, no, etc."

Too Much Baby Talk?

In addition to and/or aside from privacy and security concerns, some parents have expressed their conscientiousness in posting too many baby stories on Twitter, which was a concern raised in Morris's study of mothers and social networking use [16]. There was also a fear of overshare, which might lead to unfollowing [14]. People with this kind of concern may also choose to answer via Direct Message.

P4: "I feel like most people don't want to see me tweeting about my child's development -- I try not to inundate people with pictures on fb (just instagram), so I would feel the same on twitter."

P6: "I did not know DM was an option but I would have preferred it. I did not want to tweet publicly multiple times in a row about my baby's development. It seemed a little spammy."

Social Interactions

Because the study was relatively short-term and the number of participants was small, opportunities for social interac-

tion were limited. However, as the nature of Twitter is as a social networking service, we were interested whether participants ever interacted with each other. Among the eleven monthly accounts we had, three had multiple participants – two with two participants and one with three participants. We received several testimonies of interaction, from one following the other to a brief conversation between the two.

P5: *"I saw who else was following the @babystepsjul12 and followed those who seemed interesting."*

P10: *"I conversed with a woman on twitter who was also following the same BabySteps birth month account. We talked about enjoying the study."*

In addition, P14 showed her desire to interact with parents with a baby born in the same month. Future studies will include more participants to study this social dynamic at a larger scale.

P14: *"I think I might have been the only person in my group so it wasn't as interactive as I would have hoped."*

Twitter and Website Integration

After responding to milestone questions via Twitter for two weeks, participants were introduced to the Baby Steps website. They could view their child's developmental progress based on their responses via Twitter and were able to answer additional milestone questions, add photos, share a timeline of their child's progress, get useful resource information, etc. Although testing the Baby Steps website was not the main goal of this study, because users had a week to explore, we received quite amount of feedback that it served well as a companion of @BabySteps and they complement each other.

P6: *"I love it! So well structured. And I love that I can use it to put in milestones that I missed on twitter."*

Despite it not being a popular feature from @BabySteps, keeping sentimental records of a child with #babymemory and checking it on the website received positive feedback.

P12: *"I liked that I could record #babymemory bits to check back on later."*

However, because the system only parsed and recorded the tweets with correct syntax for this study, not all of participant responses were recorded on the website, which caused frustration to some participants.

P4: *"It didn't actually pull in the responses that I had posted to twitter -- it only recorded three of the responses, and I had responded to all the questions"*

P11: *"None of my responses are there because I messed up the syntax."*

Overall Feedback

In the post-study survey, we asked participants about their general perceptions and suggestions for improvement or new features. Most participants had a positive reaction to @BabySteps and the website. Nine out of twelve participants who completed survey said that they want to continue to use both @BabySteps and website after the end of the

study. They also said that they would like to recommend our system to friends. For the most part, many favored that the system identified informed tasks and developments that baby should be doing.

P4: *"It was really nice to learn about how her development is potentially matching up with where she should be."*

P8: *"It encouraged me to try some things with my daughter (like letting her feed herself with a spoon) which I hadn't thought to do yet."*

Also, some participants appreciated how @BabySteps made them more aware of their child's development and grateful for their child meeting development progress on time.

P6: *"I loved that it reminded me about the less obvious milestones she was achieving."*

P3: *"I think some of the development that happened we just took for granted... so getting to keep track of it showed that our kid was developing at a pace."*

Ratings in a 5-point Likert scale (5 being the positive answer) revealed that participants think our system can be helpful to both first time parents ($M = 4.25$, $SD = 0.87$) and parents with multiple children ($M = 4.17$, $SD = 1.03$).

P9: *"I can see how the tips and activities could be useful, especially for a 1st time mom."*

P13: *"Used it for my second child, and enjoyed re-learning about various milestones I had experience with my older child >3 years ago."*

It turned out that we had a pediatrician among our participants. She revealed it in the post-study survey. Her motivation for participating was because she recalled that remembering the developmental milestones was difficult during her training. She gave great comments about our system.

P12: *"Full disclosure - I'm a pediatrician... You have created a fantastic tool for tracking and for guidance for parents on how their child is doing developmentally. I would love to be able to use this in my own practice. (...) This is an objective way to track milestones. I love it and I hope I get to continue using it! I've been using a blog that is open to only family to track milestones and post pictures - this is even better for milestone tracking. Thank you so very much!"*

DISCUSSION

Overall, we believe the deployment study showed that the @BabySteps system was successful in achieving its initial goals. In the end, all participants were able to learn the syntax and successfully respond to milestone messages. For most participants, the system kept them engaged over the course of the study, and we received positive feedback on the overall use and idea of the system.

Despite the initial success of our system design in our first trial of it, we believe there are a number of actions we can take to improve the usability and overall acceptability of @BabySteps. For example, our original design had enforced each participant to include 2-3 three different hashtags within their responses to milestone prompts. This included the response (#yes, #sometimes, or #no/#notyet),

the baby milestone ID (e.g., #baby2343), and if there was more than one child registered in the system, the child's name or pseudonym (e.g., #sally or #daughter). However, because participants sometimes forgot the hashtag on the response, it would be easy to modify the system to allow for answers to not require the hashtag. In addition, we could allow participants to omit the baby milestone hashtag if they reply directly to the message, since the Twitter API allows us to identify which tweets are in reply to another message. This would alleviate some of the difficulty of having to copy the hashtag, especially when using a mobile device. Finally, we could simplify the baby code slightly. We chose 4 digit codes initially, thinking they would be more distinct, but because there are only about 660 unique milestones across the entire 5 years of a child's life, we could number the milestones between #baby1 and #baby660 to shorten what the participant has to remember. These changes may prevent more user error and allow for flexibility, though it may still be necessary to be strict in the event of multiple children (like the case with P9, who tried to combine a tweet for her two twins).

In addition to simplifying the hashtag system, we believe there are additional improvements we can make to the system itself that could help participants in the beginning to learn the syntax and to also account for human error, typos, etc. The Tweak the Tweet application [20] had similar issues with having participants tweet using the right hashtags as well, which is why we had reminder tweets about the syntax. We believe adding to our script the capability of sending automated replies in the event that the user sends an improperly formatted tweet would help with correcting errors sooner. We believe these should be via direct message (DM). For example, if the user tweeted "@BabyStepsNov12 #yes, my daughter can kick a ball," we could send a DM back that says "Thanks! However, we cannot process your request due to: missing child name hashtag (#sally), missing milestone id hashtag (#baby234)."

In addition, although we did not encounter this in the short duration of our study, it is possible that a person could learn about @BabySteps via someone they follow and wish to start using the system, which we hope will happen to help broaden the reach of our system's design. However, if they are not registered with our system, their replies cannot be processed. Thus, we could send an automated message to anyone who sends us a message inviting them to register on the Baby Steps companion website.

Using Twitter for a public health application is also not without its challenges, which would be useful for others interested in working in this space. First, the 140-character limit and the need to allow for additional characters for replies was somewhat challenging. The system we built and the milestone content was inspired by the Ages and Stages Questionnaire [19]. However, due to the 140-character limit, we could not use the exact wording of the milestone questions, which in some cases were as long as 300-400

characters. Because the ASQ is a validated tool, this likely would break validation and may not be as reliable as the original. There is a similar issue with other public health campaigns that have used SMS messaging and have run into 160 character limits, and thus there are public health organizations that are working on shortened, validated messages for this format. This could be one solution to this issue, but may take some time and cooperation from the original sources of the content. Second, because Twitter remains in the public domain, privacy can be a critical issue. We cannot expect all of our users to have a strong level of media literacy. Therefore, it is important for us not only to inform them about the risk in advance, but also to give them options that they can choose to control privacy. Our future work is exploring alternatives to Twitter to track milestones, so for those who are not comfortable with using this method, they have several other options.

There is also a question on how frequent postings should be to avoid overload, missing milestones, and also keep the parent engaged without forgetting the syntax. There is currently not much guidance on how often to send messages, and we believe it is likely dependent on the health domain (e.g., it is possible parents may be willing to receive more milestone messages than they would reminders about quitting smoking). Although most of the participants in our study believed that the three milestones per day frequency was just right, there were a few who believed it was too many, or they would like to see milestones balanced more with the tips and activity suggestions. We believe additional studies are necessary to determine this correct balance.

Finally, one last issue is the need for having the companion website. It would be ideal if the system could reside entirely on Twitter, but for developmental milestone tracking, it is not enough to just collect responses from parents. We also needed to prompt them to communicate that information with their doctor and follow up as necessary. Although the eventual use case for Baby Steps is to have our state's public health system monitor parent responses in the database and directly follow up with participants based on the results, this system is not yet in place. Thus, we had to have the website where participants can log in, see all of their responses, print out the results to bring with them to the doctor, and find local resources on who to contact if they are concerned about their child's development. There may be other health applications where this may not be needed, since the tweet itself could direct them to resources.

CONCLUSIONS & FUTURE WORK

We presented the design and implementation details of @BabySteps, a system for using Twitter to create an interactive public health system for recording children's developmental milestone progress. We conducted a 3-week deployment study to understand the feasibility of such a system. Our system was successful in teaching participants to use the proper syntax to reply to tweets, we had engagement with participants over the course of the study, and we

received a number of positive comments about the system's design. Our study revealed several issues in how we can improve these types of systems. This research has two primary contributions: 1) the design and development of a novel system for using Twitter as an interactive public health platform and 2) the results of the deployment study demonstrating its feasibility. The findings from our study have implications for designing such systems and can inform others of lessons learned in using Twitter for this purpose.

We believe there are a number of opportunities for future work. This system and study has just scratched the surface for what is possible with using Twitter as an interactive public health platform that can shift the flow of information from many to a single source. We plan to conduct a study experimenting with the frequency of Tweets to determine the right balance to maintain engagement but avoid fatigue. We will also conduct a longer-term study of @BabySteps to determine whether this approach can maintain engagement over the long term and improve self-efficacy and patient activation, as well as understand how parents prefer to respond to and share milestone information when compared to just a website or other mediums, such as text messaging.

ACKNOWLEDGMENTS

We wish to thank Chun-Ku Lin, Sean Munson, Kiley Sobel, Li-Tien Ou, Robert Racadio, Mira Shah, Yuna Shim, Yi-Chen Sung, and Xiaotong Yang for their assistance with this research. This research was funded by National Science Foundation grant #0952623. The University of Washington's Institutional Review Board approved all research.

REFERENCES

- Barak, A., Hen, L., Boniel-Nissim, M., & Shapira, N. A. (2008). A comprehensive review and a meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. *JTHS*, 26(2-4), 109-160.
- D'Alessandro, D. M., & Dosa, N. P. (2001). Empowering children and families with information technology. *JAMA Pediatrics*, 155(10), 1131.
- De Choudhury, M., Counts, S., & Horvitz, E. (2013). Predicting postpartum changes in emotion and behavior via social media. *CHI 2013*, pp. 3267-3276.
- Duby, J. C., et al. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*, 118(1), 405-420.
- Fell, H., Cress, C., MacAuslan, J., & Ferrier, L. (2003). visiBabble for reinforcement of early vocalization. *ASSETS 2003*, pp. 161-168.
- Foucault, B.E. Designing Technology for Growing Families. *Technology@Intel Magazine*. 2005.
- Guralnick, M. J., & Bricker, D. (1987). *The effectiveness of early intervention for children with cognitive and general developmental delays*. Academic Press.
- Hoefman, B., & Apunyu, B. (2010). Using SMS for HIV/AIDS education and to expand the use of HIV testing and counseling services at the AIDS Information Centre (AIC) Uganda. *M4D 2010*.
- Jeong, H. Y., et al. (2008). Opportunities to support parents in managing their children's health. *CHI 2008 Extended Abstracts*, pp. 3225-3230.
- Jordan, E. T., et al. (2011). Text4Baby. *Nursing for Women's Health*, 15(3), 206-212.
- Kientz, J. A., Arriaga, R. I., & Abowd, G. D. (2009). Baby steps: evaluation of a system to support record-keeping for parents of young children. *CHI 2009*, pp. 1713-1722.
- Kientz, J. A., et al. (2007). Grow and know: understanding record-keeping needs for tracking the development of young children. *CHI 2007*, pp. 1351-1360.
- Klasnja, P., Consolvo, S., & Pratt, W. (2011). How to evaluate technologies for health behavior change in HCI research. *CHI 2011*, pp. 3063-3072.
- Kwak, H., Chun, H., & Moon, S. (2011). Fragile online relationship: a first look at unfollow dynamics in twitter. *CHI 2011*, pp. 1091-1100.
- LENA. Retrieved Sept. 18, 2013 from <http://www.lenababy.com>
- Morris, M.R. (2014) Social Networking Site Use by Mothers of Young Children. *CSCW 2014*.
- Munson, S. A., et al. (2010). Happier Together: Integrating a Wellness Application Into a Social Network Site, *Persuasive 2010*.
- Park, S., Jeong, H., & Zimmerman, J. (2008). ENSURE: Support for Parents in Managing their Children's Health. *Design and Emotion*.
- Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires®*. Baltimore, Maryland, 257-182.
- Starbird, K., & Palen, L. (2011). Voluntweeters: Self-organizing by digital volunteers in times of crisis. *CHI 2011*, pp. 1071-1080.
- Tang, K. P., et al. (2012). Designing a mobile health tool for preterm infant wellness. *Pervasive Health 2012*, pp. 232-235.
- Trixie Tracker Software. Retrieved Sept. 18, 2013 from <http://www.trixietracker.com>
- Victoria Udoh. Nigeria: MTN to Broadcast Anti-Aids Messages. (September 2003). Retrieved Sept. 18, 2013 from <http://allafrica.com/stories/200309260162.html>
- Westeyn, T. L., et al. (2012). Monitoring children's developmental progress using augmented toys and activity recognition. *PUC*, 16(2), 169-191.
- Woolford, S. J., et al. (2010). Tailored mobile phone text messages as an adjunct to obesity treatment for adolescents. *JTT*, 16(8), 458-461.