6.

From “So Cool” to “I’m Bored”

Longitudinal Trends in Activity Monitors and Gaming

Danielle Hagood (University of California, Davis) & Cynthia Carter Ching (University of California, Davis)

Abstract

To integrate gaming and physical activity among youth (ages 13 and 14), activity monitors were used to track 42 participants’ physical activity throughout the day and in-turn integrate activity information into a virtual game world. In this analysis of Fitbit data logs, random-effects growth curve analyses were used to model the general activity trend. A two-phase model is introduced that explores how activity changes before versus after the game. This quantitative analysis of activity trends is interpreted using participant interviews. The paper concludes by making the case that game and physical data analytics necessitate complimentary qualitative analyses.

Introduction & Research Questions

A decade ago, games research diverged into two strands. In the first, scholars like James Gee focused on the positive affordances of videogames for learning and engagement (e.g., Gee, 2003; Shaffer, 2006; Squire & Jenkins, 2003). In the second, health and human development scientists explored the possible negative effects of videogames (e.g., Tremblay & Willms, 2003; Vandewater, Shim, & Caplovitz, 2004). Today this divide remains largely intact, with researchers from these different approaches evaluating different questions: How do games affect kids minds? How do games affect kids health? In this research study, we reunite these research perspectives. Instead of asking how games can be more active, like exergames for Wii Fit, or focusing on the games that “sneakily” introduce health information, the current project designed an engaging exploration game called *Terra* that incorporates students’ daily physical activity — collected through wearable physical activity monitors — into a game.

This work fills a current gap in research and game design, as almost all exergames encourage players to move, but do not require players to consider how, when, or why they exercise and move actively. In order to integrate gaming and physical activity, activity monitors were used to track participants’ physical activity throughout the day and in-turn integrate activity information into a virtual game world. In the sense that our design allows users to keep track of physical activity, this study is not unique. We are drawing on the growing approach known as “Quantified Self” in which individuals track intensive personal data (Lee, 2013). Research on the “Quantified Self” promotes the use of wearable devices to track myriad health metrics that provide the user with a slew of data that can be employed in daily decision-making and long-term planning (Swan, 2009).

However, the Quantified Self approach alone is not necessarily enough to change behavior, at least among children (Swan 2006). In contrast, games and video games can be a robust mechanism for transformative individual change (Bogost, 2011; McGonigal, 2011). This program integrates the activity data stream format Quantified Self into the transformative landscape of game play.

Evaluating the efficacy of this combination is an urgent concern. Since 1980s, childhood obesity has doubled from 7% to 18%, while the adolescent obesity rate has tripled, growing from 5% to 18% (Ogden, et al, 2012). This issue creates far-reaching consequences in light of the predictive relationship between obesity and dangerous metabolic and lifestyle-driven diseases later in life. Such diseases currently account for nearly 70% of deaths in the United States (Kung et al., 2008; Reilly & Kelly, 2011). This study will not directly evaluate interventions on childhood health. But, by exploring how something youth already do — play games — can better promote what youth should do more — move and get active — we offer insight into one avenue for change.

In the analyses below we explore general trends in the activity patterns among and between participants. This analysis builds on established growth curve modeling techniques for longitudinal data (Singer & Willet, 2003). These analyses were informed by analyses of interviews and focus groups with participants and learning ecology analyses (Stewart, Hagood, & Carter Ching, in press). We address the following research questions about general use trends and the impact of introducing a physical activity monitor game:

1. What is the general trend of activity among participants?
2. How does activity rate (i.e. steps taken) change after the introduction of the accompanying game?

Research Design

Sample

This study included a sample of 42 middle school students. Participants comprised middle school students at an average sized public school located in a small Northern California town near a large research university. The participants were equally split between two periods of an educational technologies elective course taught by the same instructor (Period 1 = 22, Period 2 = 20).

The individuals' age, gender, and ethnicity are reported here. Participants' were ages 13 and 14 (Mage = 13.7, SD = .55). Both classes enrolled mostly male students; thus participants include mostly boys, with boys making up 83% (N = 35) and girls 17% (N = 7). The self-reported ethnicity of the participants was 48% Caucasian, 17% Latino/a, 12% Asian, 5% Native American, and 2% African American, with 11% of the students not reporting their ethnicity. Note that 2% of this sample is equal to one participant.

Data Collection

In this project participants utilized activity monitors that tracked the number of steps taken and activity intensity (the latter was deemed unreliable). As the dashed boxes in Figure 1 show, participants wore the activity monitors for 30 days before the introduction of *Terra*, a tile-turning, planetary exploration game. The participants then continued wearing activity monitors for another 60 days while playing *Terra* several times each week. During the period of data collection students participated in the game development process, giving feedback and suggestions to the game designers. Play took place as an in-class activity, students were encouraged but not obligated to play. All activity and game data were collected passively as log data. Other personal information about the participants was collected through surveys and interviews.

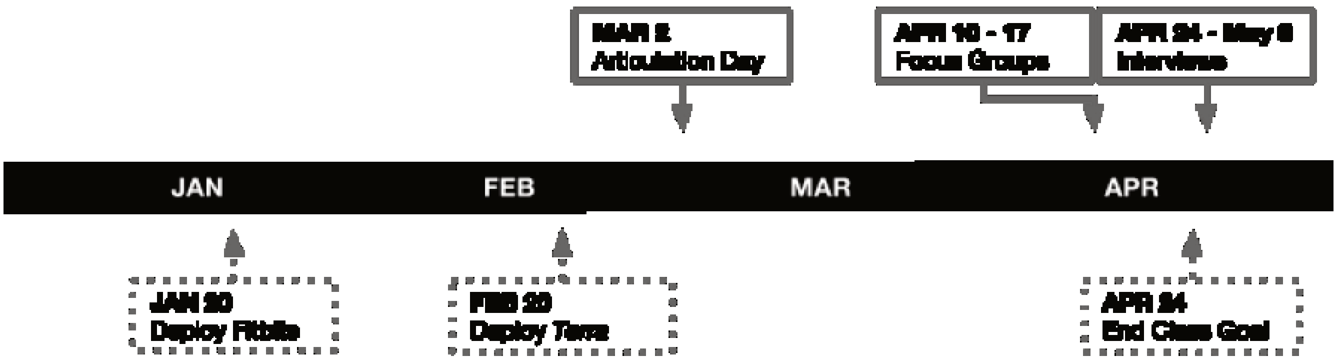

Figure 1. Timeline of Fitbit use, game introduction, and duration.

Figure 2 shows *Terra*, the online browser-based game we designed. Players of *Terra* are space explorers who have landed on a desolate planet. They set up individual domed bases, with the goal of completely “terraforming” the planet so that more of their people can come live there. *Terra* downloads information from the Fitbit online database each time a player logs in. The game displays an “Energy” window that details how many game moves are possible each game day based on in-game metrics and steps. For example, for each 1,000 steps a player has taken the previous “real-world” day, they get one extra move in the game when they log on. The timescale of the game is compressed so that players get seven “*Terra*-days” for each real-world day, while playing a week’s worth of time in the game at each daily login activities like exploring terrain, planting and harvesting crops, building their base, or caring for creatures. As the game progresses, the landscape of the world that players create becomes an aggregate visual representation of their synced activity over the variable time frame of the game campaign, with each player’s landscape reflecting not only strategic in-game decisions but also the extent of their daily fitness.

Figure 2. Screenshot of player's developed landscape in Terra.

Measures

Quantitative data were captured as passive measures through activity monitors and game logs. Since these are not psychological measures, we do not report psychometric properties like reliability and validity. A survey including several scales was used to collect additional participant data, especially demographic data. Below we introduce the variables of interest.

The primary outcome measure of interest is activity, measured the number of steps an individual takes. Steps were measured at an hourly level, then aggregated to the week level. Each weekly step observations represents the average steps a participant took each day that week, taking into account how frequently they used the device. The average number of daily steps (across weeks) is 6415 (SD = 3738).

Data Analysis Plan

Following established model building practice (Singer & Willet, 2003), we first fit several unconditional models. This process requires iteratively fitting and evaluating each model to identify the model that best describes game play over time (see Table 1). Model 1, the unconditional means model, allows us to partition the variance within and between individuals. Model 2, the unconditional growth model, provides a base rate of change over time. We also explored non-linear rates of change. Model 3 is like the unconditional growth model (Model 2), but adds quadratic time effects. This model explores how the introduction of the game mid-study could impact the number of steps taken.

Based on Models 1–3, and on interview reports, we next explored if two-phase models better described the patterns of activity throughout the study. Model 4A provides the two-phase growth model with linear rates of change. While Model 4B fits the two-phase growth model with quadratic rates of change. In both Modal 4A and Model 4B, separate coefficients were calculated for before the game was introduced (“BG”) and after the game play started (“AG”).

Model	Equation
Model 1	$average\ steps_{ij} = \gamma_{00} + \zeta_{0i} + \varepsilon_{ij}$
Model 2	$average\ steps_{ij} = \gamma_{00} + \gamma_{10}WEEK_{ij} + (\zeta_{0i} + \zeta_{1i}WEEK_{ij} + \varepsilon_{ij})$
Model 3	$average\ steps_{ij} = \gamma_{00} + \gamma_{10}WEEK_{ij} + \gamma_{20}WEEK_{ij}^2 + (\zeta_{0i} + \zeta_{1i}WEEK_{ij} + \zeta_{2i}WEEK_{ij}^2 + \varepsilon_{ij})$
Model 4A	$average\ steps_{ij} = \gamma_{00} + \gamma_{10}BG.WEEK_{ij} + \gamma_{20}AG.WEEK_{ij} + (\zeta_{0i} + \zeta_{1i}BG.WEEK_{ij} + \zeta_{2i}AG.WEEK_{ij} + \varepsilon_{ij})$
Model 4B	$average\ steps_{ij} = \gamma_{00} + \gamma_{10}BG.WEEK_{ij} + \gamma_{20}BG.WEEK_{ij}^2 + \gamma_{30}AG.WEEK_{ij} + \gamma_{40}AG.WEEK_{ij}^2 + (\zeta_{0i} + \zeta_{1i}BG.WEEK_{ij} + \zeta_{2i}BG.WEEK_{ij}^2 + \zeta_{3i}AG.WEEK_{ij} + \zeta_{4i}AG.WEEK_{ij}^2 + \varepsilon_{ij})$

Table 1. Taxonomy of models followed to identify best fit.

Missing Data

The outcome variable of interest, steps, has missing values. Based on interview data we can conceptualize missingness for the steps, when 0 is recorded for steps, as made up of two parts. First a zero might be recorded when the participants chose not to use their fitbits. The second is when the fitbit deleted records because it was not syncing frequently enough. Based on our interviews we believe the latter case is missing completely at random. However, the former, missing step data based on disuse is likely related to engagement. At best, some other predictors in our model may predict this type of randomness. At worst, these data are not missing at random.

To improve this situation, we created a use frequency variable which creates a binary variable noting when daily steps as is zero and non-zero. This variable also includes missing data, but they are missing due to technical errors. We also evaluated multiple imputation techniques for dealing with missingness in other predictor variables, however, more than 5% of the survey data that would be used for multiple imputation was also missing. Since the amount of missing data exceeds the is a recommended threshold for imputing values (Graham, 2009), we did not complete multiple imputation. Later in the discussion we return to the concept of missingness, and explore how missing data in game and physical data analytics necessitates companion qualitative data.

Results & Interpretation

The first research question explores the general trend of activity among participants. To address this question we compare the models to determine best fit. It is important to note that Model 1 outlines the unconditional means model, which describes the initial variation among participants. This model is useful in that it partitions the variation among and between individuals, using the intraclass correlation coefficient (ICC). The ICC in this case is .65, meaning that 65% of the total variation in average daily steps is attributable to differences among participants. Model 1 also sets a baseline AIC and BIC our goodness-of-fit indices to compare among the Models (lower AIC and BIC are preferred).

Next, as shown in Table 2, it is unclear if Model 2 or Model 3 best represents the general trend in activity. Based on descriptive data visualization we suspected a non-linear model would fit best. However, neither Model 2 (linear change) nor Model 3 (quadratic change) offered meaningfully lower AIC or BIC. Thus, we moved to evaluating a two-phase model. In the two-phase model *BG* coefficients indicate coefficients for weeks occurring before the game was introduced, while *AG* coefficients indicate second phase weeks occurring after participants started playing the game. In both Model 4A and Model 4B the intercept and *AG* linear change are significant. However, the addition of quadratic effects reduces (i.e. improves) the AIC in Model 4B, but not BIC. Thus, either of the Model 4s could arguably be interpret as best fitting our data. One note is that Model 4B does not converge when including a random-effect for the after-game quadratic effect (AG Week^2), thus this term is not included in the model

So, regarding the first hypothesis—what is the general trend of activity among participants—a two-phase model separating before- and after-game effects is better than a single phase model. But, among two-phase models, modeling time as linear or quadratic offers similar results. Figure 3 shows the average trend line in the two-phase model.

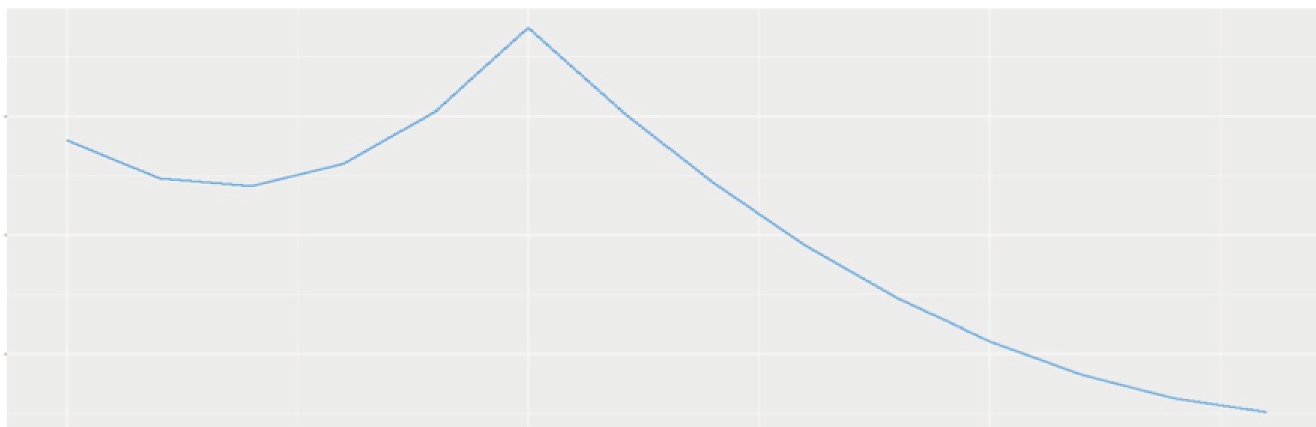

Figure 3. Two-phase model: average steps over time.

The two-phase models (Modes l4A and 4B) address the second research question—how activity relates to the introduction of the game. Since a two-phase model better describes the trends than a one-phase model, this indicates that something is different before- and after-game. Considering Model 4B, we see that the average number of daily steps taken in the first week of the study is 6787. In the weeks of the before-game phase, individuals took less steps each week. The decrease in steps became less sever in each subsequent week. Then, after the game is introduced, participants took 6082 steps daily and this average declined by 1131 daily steps each subsequent week. So overall, we found that each phase of the study began with interest and then declined. But the weekly decline was smaller each subsequent week. Figure 3 shows the predicted trend of average daily steps by week.

		Model 1	Model 2	Model 3	Model 4A	Model 4A
One Phase Models						
Fixed Effects						
Intercept	γ_{00}	8028**	7228**	6556**		
		(368)	(490)	(509)		
Week (Linear)	γ_{10}		-141*	238		
			(63)	(154)		
Week (Quadratic)	γ_{20}			-33*		
				(12)		
Variance						
Within-person	σ_ϵ^2	4.62E6	6.83E6	6.18E6		
In initial status	σ_0^2	8.42E6	7.52E6	6.99E6		
Linear Slope	σ_1^2		9.79E4	4.33E5		
Quadratic Slope	σ_2^2			2.09E3		
Covariance	σ_{01}			-6.96E5		
Covariance	σ_{02}			9.67E3		
Covariance	σ_{12}			-2.62E4		
Two Phase Models						
Fixed Effects						
Intercept	γ_{00}				6586**	6787**
					(467)	(5484)
BG.Week (Linear)	γ_{10}				117	-426
					(119)	(375)
BG.Week (Quadratic)	γ_{20}					12
						(74)
AG.Week (Linear)	γ_{30}				-403**	-705*
					(101)	(223)
AG.Week (Quadratic)	γ_{40}					37
						(27)
Variance						
Within-person	σ_ϵ^2				5.90E6	5.47E6
In initial status	σ_0^2				6.55E6	9.09E6
BG Week(Linear)	σ_1^2				2.95E5	1.26E6
BG Week (Quadratic)	σ_2^2				2.10E5	7.01E4
AG Week(Linear)	σ_3^2					2.95E5
Covariance	σ_{01}				-5.70E5	-2.53E6
Covariance	σ_{02}				-5.52E5	-1.07E6
Covariance	σ_{03}				-3.24E4	4.14E5
Covariance	σ_{12}					4.99E5
Covariance	σ_{13}					-2.76E5
Covariance	σ_{23}					-1.05E5
Goodness-of-Fit						
AIC		8028	7994	7983	7972	7968
BIC		8040	8018	8023	8013	8033

** $p < .001$; * $p < .05$

Table 2. Hierarchical growth curve results.

Discussion

Answering the Research Questions

The answer to the first research question: What is the general trend of activity among participants? Is that there are two distinguishable phases in activity rates, one before the game is introduced and one after the game is introduced (see Models 4A, 4B and Figure 3). Regarding the second research question: How does the activity rate (i.e. steps taken) change after the introduction of the game Terra? As Figure 3 shows, before the game the youth record less steps on their fitbits over time. However, there is a resurgence of step activity during the week when the game is introduced, but this peak return quickly to the pre-game level of activity and subsequently decreases more.

One major concern for these analyses are that the data are not necessarily missing at random. This was discussed previously, but it should be noted that future iterations of this project should keep this issue in mind in the design process. Especially because the missing data due to technical problems adds measurement error to “use rate.” Future models including should include the use-rate variable to account for missing at random data.

The activity level measured by steps in this study varies greatly between individuals (see Figure 5). This observation is confirmed by the ICC value that of the total variation in average daily steps, 65% is attributable to differences among participants. Future work with these data should explore between-participant variables to explain this difference. Further extensions of the present study will evaluate within-participant and time-varying predictors that better predict the number of steps taken. Currently, we have only evaluated uncontrolled models, but a controlled effects model could answer future research questions like: What predicts the rate of change before the game is introduced versus after the game is introduced?

The Argument for Companion Qualitative Analyses

Toward the end of the study we also conducted individual interviews and focus groups with a subset of students (see Figure 1 for timeline); approximately 50% of students participated in both classes. In these conversations we asked students to recollect and describe their experiences wearing the Fitbits and playing *Terra* throughout the duration of the project: at the beginning (when they first received the devices), when we introduced *Terra*, and what researchers conducting the conversations referred to as “now” (i.e., in the fourth month, when focus groups and interviews took place). Students described diachronic reflections, comparing experiences across time between “then” and “now,” which we can then examine relative to the quantitative model depicted in Figure 3.

What emerged from that analysis (which we do not have the space to delve into fully, beyond general themes), is that while many students describe an initial surge of interest followed by a decrease in game engagement and device wear, both of which correspond to the spike and subsequent decrease in steps and patterns of “missingness” in the latter portions of the quantitative model, there were some fairly different categories of reasons students described for their disengagement in particular. In the following excerpt Gabriel suggests that anxiety around having to be responsible for the device resulted in his decreased interest:

> Sara: At first, how often were you wearing it?
>
> Gabriel: At first I was wearing it all the time.
>
> Sara: And now how often would you say you’re wearing it?
>
> Gabriel: I don’t really wear it at all anymore.
>
> Sara: So what changed for you?
>
> Gabriel: Well, I would be always like losing it, or I’m thinking it’s lost and then it’s not, and I didn’t want to break it…(pause). I have it somewhere.

Students also described a decrease in social motivation to wear the Fitbit at school, wherein at first it was a symbol of “specialness” that generated peer interest but ultimately became mundane and unnoticed.

Other students described frustration with occasional bugs in the game that prevented syncing their devices, or frustration with inaccuracies in how the device measured movement, or frustrations with poor alignment between their chosen physical activity (e.g., swimming) and the form factor of the device. Still others became bored with various aspects of the game over time.

What these different themes show is that “missing data” or “decreased engagement” are not all the same, and it is dangerous to ascribe the same meanings to similar data patterns across individuals. Studies of pedometer interventions in health research tend to explain missing data as “non-compliance” or “failure of fidelity,” but few deeply examine participant meanings around their devices and their data. Our ongoing goal is to find ways of combining quantitative and qualitative data to provide a better picture of what we can see in this study of physical activity monitor gaming.

Acknowledgments

This research is funded by grants from the National Science Foundation IIS-1451446 and IIS-1217317.

References

Bogost, I. (2011). *How to do things with videogames* (Vol. 38). U of Minnesota Press.

Gee, J. P. (2003). *What video games have to teach us about literacy and learning.* New York: Palgrave Macmillan.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual review of psychology*, 60, 549-576.

Kung, H. C., Hoyert, D. L., Xu, J., & Murphy, S. L. (2008). Deaths: final data for 2005. *Natl Vital Stat Rep*, 56(10), 1-120.

Lee, V. R. (2013). The Quantified Self (QS) Movement and some emerging opportunities for the educational technology field. *ITLS faculty publications*. Paper 480.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world.* Penguin.

Ogden, C. L., Carroll, M. D., Kit. B. K., & Flegal, K. M. (2012). Prevalence of obesity and trends in body mass index among US children and adolescents, 1999-2010. *JAMA, 307*(5), 483-490. doi:10.1001/jama.2012.40.

Reilly, J. J. & Kelly, J. (2011). Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review. *International Journal of Obesity, 35,* 891-898. doi:10.1038/ijo.2010.222

Shaffer, D. W. (2006). *How computer games help children learn.* New York: Macmillan.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* Oxford university press.

Stewart, M., Hagood, D., & Carter Ching, C. (in press). Virtual games in real-world communities: Environments that constrain and enable physical activity in games for health. *International Journal of Game Based Learning.*

Squire, K., & Jenkins, H. (2003). Harnessing the power of games in education. *Insight, 3*(1), 5-33.

Swan, M. (2009). Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified selftracking. *International Journal of Environ Res Public Health, 6*(2), 492-525.

Tremblay, M. S., & Williams, J. D. (2003). Is the Canadian childhood obesity epidemic related to physical inactivity? *International Journal of Obesity, 27*(9), 1100-1105.

Vandewater, E. A., Shim, M. S., & Caplovitz, A. G. (2004). Linking obesity and activity level with children's television and video game use. *Journal of adolescence, 27*(1), 71-85.