11-16-2016

Towards Computational Human Behavior Modeling for Just-in-Time Adaptive Interventions

Tylar Murray
University of South Florida, tylarmurray@mail.usf.edu

Towards Computational Human Behavior Modeling for Just-in-Time Adaptive Interventions

by

Tylar Wayne Cole Murray

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Electrical Engineering
College of Engineering
University of South Florida

Co-Major Professor: Andrew Raij, Ph.D.
Co-Major Professor: Wilfrido Moreno, Ph.D.
Eric Hekler, Ph.D.
Daniel E. Rivera, Ph.D.
Donna Spruijt-Metz, Ph.D.
Alex Savachkin, Ph.D.
Pooja Patnaik Bovard, Ph.D.

Date of Approval:
November 14, 2016

Keywords: JiTAI, CHBM, systems, HCI, avatars

ACKNOWLEDGEMENTS

This dissertation was made possible thanks to the support of my friends, family, fellow students, advisors, and professors. The support and guidance of Dr. Andrew Raij throughout this work has been invaluable in my studies as a graduate student, and I cannot thank him enough for the opportunities he has provided. Secondly I would like to thank Dr. Wilfrido Moreno for his mentorship throughout my work as a TA and as a doctoral candidate. Drs. Eric Hekler, Daniel E. Rivera, Donna Spruijt-Metz, Pooja Patnaik Bovard and Alex Savachkin also deserve thanks for their help overseeing my dissertation and my research. Lastly, thanks is owed to the University of South Florida Electrical Engineering Department for supporting me as a TA in my early years as a graduate student, and to The Charles Stark Draper Laboratory for supporting me as a research fellow.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The advent of powerful wearable devices and smartphones has enabled a new generation of "in-the-wild" user studies, adaptive behavioral intervention strategies, and context measurement. Though numerous proof-of-concept studies continue to push the limitations of what a behavioral scientist can do with these technologies, there remains a major methodological roadblock separating behavioral theory and application. Avatar-user interaction theory, for example, is not well defined in its formulation, and thus guidelines for intervention designers depend on heuristic methods and designer intuition. Computational modeling has been slow to move into behavioral science in general, but a growing population of behavioral scientists recognize this shortcoming and are eager to apply new technology to their work. In order to help close this disciplinary rift between systems engineers and behavioral scientists, human-computer interaction principles must be applied to make the seemingly inaccessible "magic" of modeling and simulation techniques accessible to behavioral scientists. Thus, this dissertation presents formative work to help bring engineering methodology to human behavior modeling and simulation.

Using theories of avatar-user interaction theory, physical activity regulation, and "information overload" as applications to drive toolkit design, usability considerations and interface needed to connect behavioral scientists with dynamical systems modeling are explored. A number of challenges unique to the modeling of human behavior and quirks of extant modeling efforts in behavioral science mean that existing modeling tools do not satisfy the needs of the community, and a novel design to address these shortcomings is presented.

Exploration of the fundamental design questions which arise from application of engineering principles to this unique problem will produce quality publications in software engineering, HCI, and behavioral science. Furthermore, both the “behaviorSim” toolkit and the innovative inclusion of modeling and simulation represent significant contributions to the development and application of human behavioral theory.

CHAPTER 1: INTRODUCTION

1.1 Behavioral Choices in Healthcare

Leading research indicates that poor personal health decisions are the leading cause of death [1, 2]. Health-care costs attributable to obesity alone are projected to double every decade, engulfing an estimated 16-18% of total US health-care costs by 2030 [3]. Behavioral interventions have been shown effective at initiating a change in health decisions related to obesity [4, 5] and smoking cessation [6, 7]. Despite significant advances in the theory and practice of behavioral science, humans continue to make poor behavioral choices on a daily basis, and the reasons for those choices remain an open research challenge. The consequences of these daily choices are often insignificant in the moment, but over time build up to larger individual and societal problems.

Habitual inactivity, poor diet, and smoking are likely to lead to a variety of health problems (e.g., obesity, diabetes, heart disease, cancer, chronic pain, depression, etc.), lower quality of life and shortened lifespan [8, 9, 10, 11]. Similar challenges exist outside the realm of personal health. Academic success is a function of attending class and completing assigned tasks, among other daily behaviors [12]. In personal finance, poor day-to-day purchasing decisions can add up to large financial debts [13].

From stress disorders, sedentary behaviors (sitting at a computer all day) [14], poor eating choices (i.e. choosing french fries over salad) [15] and addictive substances, modern society is plagued by chronic illnesses avoidable through behavior change.

1.2 Mobile Health (mHealth) and Behavior Change

Health behavior change methodologies are rapidly evolving thanks to recent advances in mobile health technologies. The recent emergence of mobile and wearable devices as a platform for biomedical data collection, processing, and display has enabled a new generation of "in-the-wild" user studies, adaptive behavioral intervention strategies, and context measurement. The ubiquitous nature of these wearable, pocket technologies offers unprecedented opportunities for appropriate and timely biobehavioral feedback anytime and anywhere. Proof-of-concept mobile health (mHealth) systems have changed health behaviors and outcomes with varying levels of success [16, 17, 18].

Pedometers alone have been shown to increase physical activity by providing step-count feedback [19, 20], however the staying power of these changes is largely unknown. Step-count data can also be used to set adaptive goals which best motivate positive changes in participant behavior [21].

SMS text messaging has been shown to be effective at motivating behavior change in many domains including diabetes management, smoking cessation, and increasing of physical activity. Effective text messages typically incorporate regular reminders [20, 22], support messages [23, 24, 25], and feedback [26, 27] to individuals as well as for collection of data [28].

The confluence of pervasive sensing, machine learning, network access, and computation is facilitating new approaches to data collection and adaptive interventions. Systems can detect behaviors and psychological states such as stress [29, 30], physical activity [31, 32], social interaction [33], and smoking [34], automatically and often in real-time. These data streams provide new opportunities for mobile behavioral interventions that help users make better in the moment behavioral choices related to health [35,36], productivity [37, 38, 39], personal finance [40], and environmental stewardship [41].

Researchers theorize that an intervention which can tailor based on the user and context may be an elegant solution to empower self-management of unhealthy behaviors like substance abuse, overeating, and sedentary behavior [42]. These persuasive technologies aim to utilize contextual information (i.e. data collected from the participant's surroundings and history) to deliver personalized interventions at the optimal moment in time. One emerging class of persuasive technologies which aim to leverage real-time behavioral data is the “Just-In-Time Adaptive Intervention” (or JiTAI) which describes an intervention that adapts to an individual's changing needs and circumstances to deliver tailored support at the time when it is most needed [43]. These interventions use data that characterizes the context and individual history of the participant to adapt the intervention and present a maximally potent action at the optimal time. Imagine, for example, an anti-stress application which knows not to interrupt work meetings, but also knows when to play a favorite song to help relieve stress on the drive home. Or consider a smoking cessation application that knows precisely when and where craving is most likely before the desire to smoke is noticeable (on a work break, for example), and prompts the user to play a distracting game until the vulnerable circumstances have passed. Real-time monitoring of data to identify vulnerability to poor behavioral decisions or receptivity to intervention at any given moment is possible [44], and proof-of-concept applications have demonstrated the ability to adapt interventions to users [45, 46] and context [47, 48]. Although the potential applications of JiTAIs are numerous, there remain significant challenges to be overcome by the research community before the potential of JiTAIs can be unlocked.

1.3 New Theories Needed to Support Emerging Behavioral mHealth

A major limitation on the path to digitally-enhanced self-control is our limited understanding of why and how people make unhealthy choices in spite of goals. Current methods for conceptualizing the system driving human behavior take a piece-wise, descriptive approach, examining a phenomenon in detail, but often overlooking how the model fits into the bigger picture. These methods are sufficient for analysis of traits which do not change much over days or weeks, but data collection and intervention delivery timing is now available to the microsecond for physiological data, behavioral features at the minute-level, and psychological constructs (via EMA [49]). JiTAIs can be tailored and delivered through automated messaging systems, smartphone applications but “a major gap exists between the technological capacity to deliver JITAIs and existing health behavior models.” [42]

Extant behavioral theories focus on nomothetic and static insights that do not offer the granularity and specificity to support the full potential of JiTAIs [50]. The extreme level of detail required to allow a JiTAI application to select from the myriad of intervention options, intervention timings, and tailoring features based on the growing set of contextual information available (including intervention history) is not offered by any modern behavioral theories. Such an application requires a detailed quantification of the relationships between contextual inflows and the selection of intervention options. Furthermore, these relationships may be unique to each participant, and may need to be personalized.

A common approach to the problem of inte prevention adaptation, tailoring, or timing-optimization is to use a set of if-then-style decision rules which define the behavior of the application. For instance, consider the following simple rule: if the user has been sedentary recently then deliver the intervention. This intervention is JiT, but it is not adaptive. The rule could be modified to use location context to adapt the intervention, perhaps delivering a different kind of intervention at work and at home and not intervening at all in the car. Expressing this increase in complexity can become quite wordy, but the behavior is fairly straightforward to express in pseudocode:

```
IF has_been_sedentary
    IF home
        Intervention1
```

```
ELSE IF work
     Intervention2
```

If-then and if-then-else logical structures like these are common across many programming languages and are an effective means for codifying the behavior of systems with relatively few conditional statements, but by using current methods the complexity of the behavioral model underlying a JiTAI application grows exponentially as the complexity of the intervention design increases. This is because with each additional contextual consideration or intervention tailoring option made available, each cross-condition must be considered. In our initial example we started with two contextual states (sedentary, non-sedentary) and two intervention options (intervention, no intervention), leading to a single if-else statement which expresses intervention output at the two possible states. After adding an additional contextual element with three states (work, home, car) we now must express behavior at six possible user states in our decision rule structure. Consider now an application which takes into account 10 distinct locations as well as 3 levels of physical activity, 3 levels of eating behavior health, and 3 levels of sleep quality measured over the past hour, day, and week; such an application must describe behavior across 810 possible user states. Decision trees allow for more concise expression of if-then-style application behavior, but the complexity of system behavior required for realistic application of JiTAIs is not feasibly expressed in the form of decision rules.

A more robust method of codifying application behavior is to develop a mathematical model of the decision process. Machine learning techniques can be used to develop a data structure that can apply controls to the system or predict system outcomes. Such a model might, for example, learn the correlations between contextual variables and the intervention which best optimizes behavior. The model could then be used to determine which intervention should be delivered given the user’s current context. Unfortunately, training a machine learning model from data alone requires a large amount of data to learn from. Multiple data points in each dimension of intervention tailoring and contextual input would be needed. This means that as the number of ways to tailor an intervention increases, the data becomes increasingly sparse. The problem of context-intervention training data is compounded by the fact that behavioral responses to interventions are extremely varied and difficult to predict. Due to the extreme complexity of the human system, behavioral datasets will be plagued by unaccounted confounds and unexplained behavioral responses. Even further complicating this problem is the notion that there may not be a single model which works for all users; users may differ so greatly from one another that data may not hold predictive value across users.

An alternative approach to the use of machine learning to encode the decision process is to build a model of the system based on *a priori* assumptions about the model structure which can then be used to optimize the delivery of interventions. Using a model of the human system to optimize intervention delivery may also have the benefit of helping to inform intervention designers based on the underlying theory of the model. This control systems approach is not yet popular in the behavior science community, but methods of model-based control for intervention optimization have been proposed for treatment of fibromyalgia [51], tobacco addiction [52], childhood anxiety [53], gestational weight gain [54]. In order to close the gap between systems modeling and behavioral science new behavioral theories, new terminologies, and new experimental methods need to be developed.

1.4 Contributions by Chapter

To further motivate the need for better modeling in the development of JiTAIs, the next chapter of this dissertation presents a myriad of avatar-based intervention options available. This example application domain demonstrates the complexity of designing an adaptive mHealth intervention, even without the additional complications of intervention timing and multiple streams of contextual information. The use of avatars specifically highlights an under-explored portion of behavioral theory with many nuanced and poorly understood intervention tailoring options, making this study of avatar-user interaction an effective means of showcasing the shortcomings of extant intervention design methodology.

To delve deeper into the issue of avatar-based intervention design chapter three presents the design of a glanceable mAvatar and the results of a preliminary study to explore its effects on youth. In this chapter, the study opens more questions about the implicit human system model. The study shows no statistically significant difference between interactions, but participant responses are overwhelmingly positive and seem to support our theories of user-avatar interaction. The chapter ends with a call for better methods of modeling user state and analyzing mHealth data.

In an attempt to better analyze the effects of mHealth “interventions” like the mAvatar, chapter four introduces methods for visualization-based analysis of in-the-wild behavioral data. Through exploration of three physical activity datasets (including the mAvatar), our methods reveal effects which go unnoticed by traditional statistical analysis. These findings further hint at the need for more robust modeling of the human system in that existing models do not account for the observed dynamical behaviors we see in the data.

The fifth chapter of this dissertation presents a vision of applied behavioral modeling through formalization of computational human behavior models. In chapter five terminology and concepts to bridge the gap between systems modeling and behavior science are presented.

In closing, the sixth chapter steps back again, outlining a broad range of preliminary data from a series of studies which investigate the emerging role of various software as tools to enable human system modeling for behavioral intervention design applications.

CHAPTER 2: USER-AVATAR INTERACTION THEORY[1]

User-avatar interaction theory as presented in this chapter serves as an example behavioral theory which may underlie an mHealth JiTAI application. The merits of using avatars specifically are discussed, and the details of using an avatar as an intervention or as an interface to communicate personal health data are covered.

2.1 Why Use Avatars

The use of avatars as an interface is valuable in that avatars are a visualization primitive which can encode a great deal of information simultaneously. Furthermore, avatars are uniquely useful in that they leverage our innate abilities to interpret the human form. The 'bandwidth' of traditional visualization strategies is being strained by the ever-growing influx of data, and yet emerging 'affective computing' [55] methods call for even more highly tailorable interfaces. Avatars are uniquely suited to fill the role of influencing behavior due to their use of the human-like form as a communication medium. Humans constantly communicate using their bodies by changing their appearance and behavior, and understanding the meaning behind these changes (i.e. social cognition and perception) is typically hard-wired into our thought processes [56]. The bandwidth of this interaction is immense when contrasted with current data visualizations; humans have evolved to interact with other humans (and we do it very well), whereas graph interpretation must be learned and can only span a few dimensions before

[1] This chapter has been adapted from an article published and presented at the International Conference of Design, User Experience, and Usability. Murray, T., Hardy, D., Spruijt-Metz, D., Hekler, E., & Raij, A. (2013, July). Avatar interfaces for biobehavioral feedback. In International Conference of Design, User Experience, and Usability (pp. 424-434). Springer Berlin Heidelberg. Permission to reproduce here is included in Appendix A.

becoming overwhelming. Thus, manipulating the form of human-like avatars has the potential to be a powerful, effective, and easy-to-understand communication format.

In addition to the theoretical support for avatar interfaces, there is also significant empirical evidence that human-like avatars do influence behavior. Previous research indicates that there are at least two mechanisms whereby digital self-representations can influence individuals: the Proteus Effect and operant conditioning.

2.2 User-Avatar Interaction Effects

The Proteus Effect occurs when an individual conforms to implicit cues from a self-like avatar. Several studies on the Proteus Effect in non-mobile contexts indicate that manipulating an avatar's appearance and behavior affects a user's behavior in the real world. For example, seeing one's avatar running on a treadmill can encourage physical activity [57]; using an elderly avatar improves attitudes towards the elderly and increases saving for retirement [58, 59]; using an avatar to saw virtual trees encourages less paper use [60]; and manipulating an avatar's gaze can make the avatar more persuasive [61, 62]. In these cases, the Proteus Effect demonstrates how an avatar can exert an influence over users' perception of themselves and over their behavior. Although the precise psychological mechanism for this influence requires more investigation, one plausible theory is that users see their avatar as a model for their own behavior [63]. Alternatively, the avatar's influence could be explained by a perceived relationship between the user and his/her avatar (i.e., a shared identity [64] or an empathetic bond [65]).

Operant conditioning can influence behaviors by having an avatar function as a visual representation of success or failure. Even when avatars do not take an explicitly human form, they appear to influence behavior via this mechanism. For example, previous work has explored the use of an avatar as an operant conditioning agent and feedback mechanism for promoting physical activity. In this chapter, the physical activity of an individual is mapped to the actions and mood of an anthropomorphized virtual bird avatar [66, 67]. As physical activity increases, the bird becomes happier and more playful, flies faster, and sings more songs. Pilot work suggests that this avatar can promote increased physical activity among individuals [67]. Moving one step further from ’user-likeness’, Fish’n’Steps translates daily steps into the growth and happiness of a virtual fish [68]. Even more abstract from the concept of ’avatar’, UbiFit displays a garden on the background wallpaper of a phone. The garden is similar to an avatar which displays a user’s history, providing feedback on the user’s physical activity when glancing at the phone [69].

These examples, though spanning varying degrees of ’avatar-ness’, still serve in some sense as virtual representations of the self. Behavior change applications nearer the abstract edge of the user-likeness spectrum allow for more creative designs, but sacrifice benefits of innate interpretation. The distinction between avatar and non-avatar systems is not well defined currently, however future research will likely reveal that the display must meet some (personalized) criterion of realism, interactivity, self-presence, customization, or abstraction to be considered self-like enough to utilize the Proteus or similar effects.

This evidence, when combined with conceptual knowledge of human-avatar interaction, suggests that the use of avatar-like interfaces may create behavior change through motivation, rather than purely informative visualization methods. Thus, avatars may be a powerful new technological medium for providing core methods for behavior change based on behavioral science (i.e., goal-setting, self-monitoring, modeling, and positive reinforcement).

2.3 The Language of BioBehavioral Feedback

Before we are able to identify guidelines for the use of mobile avatars in biobehavioral feedback, we must first have an abstract model of information flow and interaction in any biobehavioral feedback system (with or without avatars). Figure 1 is a pictorial representation of the components of a generic biobehavioral feedback system and the information flow within it. It is important to note here that we use the term ’feedback’ in a loose sense in which it represents any output to the user based on user input which may affect future user behaviors. Starting from the top-left of figure 1, a description of the user’s current behavior (input) is provided via self-report or sensor. This description of in-the-moment user behavior is passed to a feedback algorithm, along with any relevant historical information. Some examples of historical information which may be taken into account are the previous day’s user behavior, feedback given to the user previously, or data on the impact of a particular form of feedback on the user. With in-the-moment and historical information, the algorithm then generates the feedback. Output is then observed which may or may not immediately convey the feedback. As a demonstration of this model, consider a typical time-series feedback visualization which displays level of physical activity inferred from accelerometers. The input in this scenario is the accelerometer data. The feedback algorithm includes the method of inferring physical activity, as well as the mapping of activity level to a 2D plot of timestamps and activity. The graph of past physical activity (created from the mapping, user settings, and/or input parameters for graph creation) makes up the virtual world, and the user navigates the world through a pan/zoom window, which determines the output.

2.4 Adding Avatars to the BioBehavioral Feedback Model

An avatar-based implementation of the model presented differs from a more traditional visualization system only in the feedback algorithm and the output to the user. Design of the feedback algorithm to map input to output is a complex task, which cannot be properly explored without better knowledge of the avatar outputs available.

Guidelines for the outputs of conventional data visualization are well established [70]; here we aim to identify and organize the wide variety of outputs available to an avatar display and move towards the identification of similar guidelines. Just as the use of item location, color, and size can be used to convey information in a chart or graph, we propose that characteristics of the avatar display can be altered to convey information. However, the critical difference between innate avatar interpretation and learned graph reading suggests that the most useful encoding attributes of an avatar are based in the psychology of avatar perception.

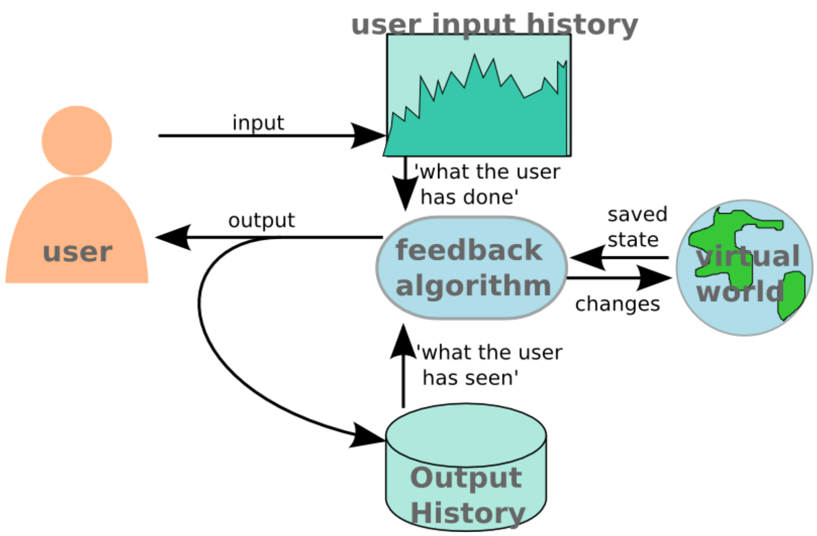

Figure 1: Information flow diagram for biobehavioral feedback algorithms.

2.4.1 Encoding Attributes in the Avatar's Physical World

Encoding attributes available in an avatar display are more numerous than those available in other visualizations due to the extremely vast amount of information humans can gain from interaction with another human-like entity. Many of these attributes, however, may have subtle or implicit influence, and impact can differ significantly from person-to-person. Here we present a generalized hierarchy to describe all conceivable changes which can be made to the physical world of an avatar. A consideration of this hierarchy can help a designer find the proper encoding attribute(s) to ensure that the effect on targeted behavior is maximized while reducing other, undesired user perceptions of avatar trait changes.

Changes to the avatar primitive in the most obvious form modify the avatar itself in some tangible way. These are attributes which fall under the physical branch. Much like existing visualization strategies, an avatar’s size, location, shape, color, etc. can be used to convey information, though in the case of an avatar these encodings often have built-in meaning to a user. For instance, inversely relating the level of daily physical activity into the width of the avatar (so he/she appears to grow thinner with exercise) is intuitive, but encoding the same value proportionally seems to send the wrong message to users, since he/she would appear to grow less fit with additional physical activity.

In addition to the encoding attributes available in an avatar’s appearance, avatars provide an additional ability to convey information via a change in their behavior. Attributes under the behavioral branch can be as simple as a change of behavior ’class’ for pre-scripted avatars (e.g. from a physically active behavior to a more sedentary behavior [71]) or may involve character attributes that should be reflected in avatar behavior. For example: a case in which an avatar demonstrates increases in strength by an ability to lift heavier objects is more than just a change in avatar behavior (lifting objects); it is a change in avatar traits (strength). Another set of behavioral attributes available to designers are the ’behavioral biometrics’ - i.e., the personal characteristics of behavior such as gait, voice timbre, and typing rhythm [72].

In addition to manipulation of the avatar primitive, algorithms may manipulate the virtual environment in which the avatar resides in order to affect user perception of the avatar. These attributes fall under the environmental branch. Changes to the environment can be cosmetic or more complex, and in many cases can have profound impact on the avatar display. For instance: avatar location and surroundings can be manipulated to go along with a behavioral avatar change (e.g., the avatar takes a trip to the beach to encourage the user to relax). Environmental changes can play an even larger role for avatars used in games; changes in the virtual environment can be used as gameplay elements. Location and virtual object removal/addition/manipulation can be used as indicators of progress or accomplishment. Similarly, aspects of the environment may be manipulated to behave differently towards the avatar (e.g., a computer-controlled agent becoming friendlier to one's avatar as a social reward for desired behavior).

The hierarchy represented by Figure 2 demonstrates the wide variety of encoding attributes available to visualization designers organized by the categories outlined. This is not intended to be an exhaustive list of possible encoding attributes, but encompasses many possibilities in an organized fashion, so that we may have a language to discuss avatar display changes just as we would discuss changes in shape, color, location, etc. of traditional data visualization. Though all possible attributes cannot possibly be included, we believe that all possible encoding attributes logically fall within the first-level categories presented (physical, behavior, and environment). Some further subdivision is shown, and attributes themselves can in some cases be further broken down (i.e. size subdivided into size of individual body parts).

Each encoding attribute can also be divided into two primary types: 1) literals - these changes have a noticeable, immediate effect on the avatar and are constantly observable. Examples include height, body shape, facial expression, current behavior, and current avatar location. 2) traits - these changes typically have a more subtle effect on the avatar; they are numerical values which describe a certain intangible property of the avatar or virtual environment. Examples include avatar proficiency at a task, behavioral biometric characteristics, and virtual character interaction characteristics. Avatar traits are a common theme in modern games, where a user may achieve a new 'level' or acquire a new 'power up' which will modify their gameplay. Traits typically will trigger a change in the value of a literal, but this change may not be apparent until a certain action is performed or as time passes. Strength, for example, is only observable when performing a strength-dependent behavior, and may display in multiple ways (e.g. speed of lifting, reduced apparent strain of lifting, increased lift height). These primary types of attributes can be found at any place in the formulated hierarchy; more examples of 'traits' and 'literals' can be identified by color as blue and red, respectively, in Figure 2.

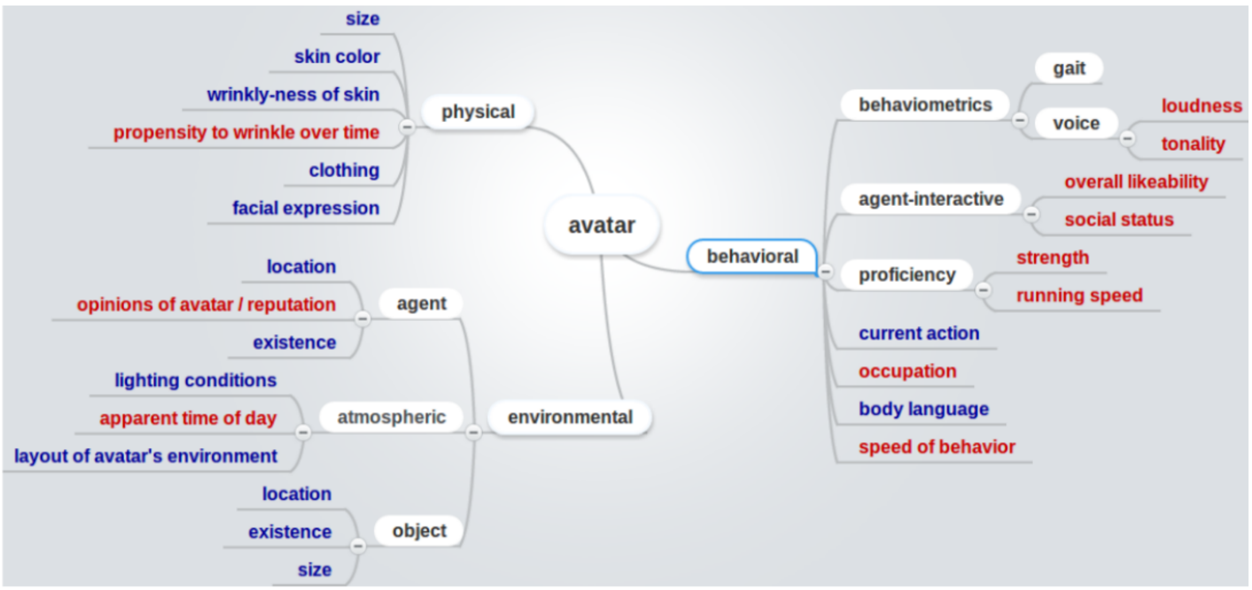

Figure 2: A hierarchal organization of potential avatar encoding attributes.

In many cases, multiple physical attributes of the avatar could be changed to express a single change. This is the case for dramatic changes in avatar identity, such as changing an human-like avatar to a plant-like creature as a reward for adopting environmentally friendly behavior or transforming the avatar's head into a greasy cheeseburger to encourage changes in diet. It must also be considered that many of these principles can perhaps apply for individual body parts (e.g., eye color, hair length, etc.). Given the practically unlimited options available to an intervention designer, it becomes important to rely on heuristic knowledge of behavioral theory and extant interventions to guide design choices. One way to avoid becoming overwhelmed with possibility is to consider a higher level of abstraction based on psychological interpretation of the human form.

2.4.2 Encoding in the Psychological

Changes in avatar appearance can be simple changes to physical or behavioral literals (e.g. change in avatar height or change in running speed), but these types of changes do not differ in principle from more traditional data visualization unless they can be interpreted without explanation. That is, a non-intuitive encoding strategy such as using avatar height to encode sleep quality is, in principle, a bar graph with human-shaped bars. However, when using an avatar primitive, simple encodings will almost always have a complex psychological effect on the user. For instance, encoding a user’s caloric intake in the overall size of an avatar could have the unintended consequence of making the user view the avatar as more attractive as he/she grows taller. This complication arises because the mapping from the user’s interpretation of the avatar to the physical or behavioral space of the avatar is not well defined; indeed, a simple change in the physical space almost always creates complex changes in the user’s perception of the avatar.

Avatar displays designed to leverage the psychology of avatar interpretation should instead aim to adjust the user’s perception of a specific, high-level trait of the avatar which is relevant to the targeted behavior change. For instance, one could aim to change the perceived abilities of the avatar by making it appear frail, weak, or elderly. By attempting to encode values in high-level interpretation rather physical traits, we can utilize heuristic knowledge of human-form interpretation in our intervention design. Figure 3 provides a minimal demonstration of selected ’low-level’ physical attributes and ’high-level’ psychological encoding attributes.

2.5 A Guide to Application of Avatar Interface

The work of Yee et al. [73] provides some example of avatar visualization design from the psychological perspective. Experiment designers wished to modify a psychological construct (the attractiveness of the user) and did so by using height as a proxy based on existing research. To further deepen the effect, other modifications could have been made to the avatar in order to modify the perceived attractiveness. For instance, adjusting the facial features [74] could have also been used. Care must be taken not to assume that multiple changes combine linearly, however. In general, adjustment of multiple encoding attributes could cause an entirely different effect than the original two.

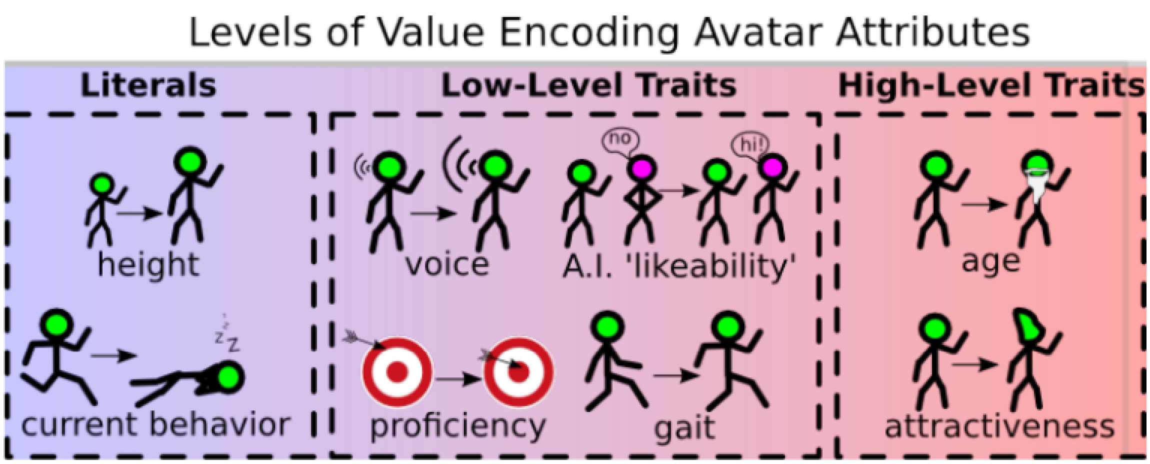

Figure 3: Levels of encoding attributes used to modify user perception.

Below we describe suggested stages of development for designing an avatar interface. Like many design processes, traversing the stages is iterative in nature; progression through the stages often reveals a need to return to a previous stage to further refine the design. A constant re-checking of past decisions is crucial to creating an avatar interface that is coherent across all dimensions of the system. We present some novel examples as well as some from literature at each suggested stage in Table 1.

1) Identify Behavior Change Plan - In this step one must identify our general plan for behavior modification. The targeted behavior must be explicitly defined and a theoretical basis for motivating a change must be found. In this chapter we argue that motivation for change is generated with avatars through both the Proteus effect and operant conditioning, but other psychological theories could be applied here as well. Existing literature provides additional guidance on the use of behavioral theories within an HCI context [75], as well as explanation and tools for defining, understanding, and describing behavior change [76, 77].

2) Identify Target Trait in User's Perception - Once we have an overall plan for motivating change, we must identify precisely what part of the user's perception of the avatar we aim to use as the encoding attribute. The resulting 'high-level' trait(s) should come from behavior change literature or designer intuition and not from the hierarchy of low-level traits and literals.

3) Map the Target to Avatar's Physical Encoding Attributes - Once the targeted high-level encoding attribute is found, the desired effect should be mapped to physical changes in the avatar such as those laid out in Figure 2. At times an easily manipulable physical attribute can act as a proxy for conveying a more psychosocial concept (e.g.: perceived attractiveness could be changed by manipulating height or facial symmetry), but in many cases there may not be significant literature on perception of the targeted attribute. Sometimes this mapping is so intuitive that researchers may (rightfully) not see it worthy of investigation; for example: perceived age can certainly be conveyed via wrinkling of the skin and whiteness of hair. However, it is important to explicitly consider this process of assumption to ensure that the targeted trait is conveyed most effectively. Returning to Figure 3, this process moves us leftwards away from the high-level user-perception space. In fact, encoding attributes must be

reduced leftwards completely to the avatar's physical space in order to implement encoding attributes, which, in turn, forces the user to do a great deal of interpretation. Here an understanding of the target audience is extremely important, since cultural or personal differences can greatly change user interpretation. Just as in human-human interaction, subtleties such as clothing, posture, and body language carry a great deal of information to the user - even if designers do not intend them to. In this way, all physical attributes are constantly interpreted, so implementations should be carefully checked for potential misinterpretation. Ultimately, some confounds and unintended effects are inevitable, but at this stage we minimize potential harm through careful consideration and iterative testing of many possible mappings.

Table 1: Example applications at each stage of interface development. Three examples given based on existing research, and two hypothetical examples.

	Outline Behavior Change Plan	Identify targeted trait(s)	map to Avatar's Physical Encoding Attributes
Time to Eat [21]	to improve eating habits subjects are shown a virtual pet with mood determined by the healthiness of breakfast	mood of the pet is the targeted trait	mood of the pet is conveyed on a scale between happy and sad through images with varying facial expression, posture, and text.
MILES [17]	The happiness of a bird avatar is determined by user level of physical activity.	percieved level of happiness of the bird is the targeted avatar trait	the mood of the bird is shown via increased flight speed, and increased song singing.
Yee et al. [24]	hypothesis: increase in avatar attractiveness changes subject's social behavior	Percieved avatar attractiveness is targeted trait.	Avatar height used as a proxy for 'attractiveness'
stressReduce	Desire: reduce overall stress. Motivate behavior change via operant conditioning; users see avatars recieve rewards and experience happiness when avatar is relaxed, unhappiness when avatar is stressed.	levels of stress' and 'levels of happiness' need to be conveyed	level of stress' conveyed via frizzing of avatar's hair, hands placed on temples, bags under eyes, amount of steam coming out of ears. 'level of happiness' via smile achieved through manipulation of mouth and eyes.
ActivitySuggest	to increase user physical activity (PA) via the proteus effect user is presented with PA avatar to increase desire to be PA. Avatar mirrors user level of PA as measure from accelerometers most of the time, but occasionally 'suggests' a higher level by simply displaying it.	Aim is to modify user's percieved level of avatar physical activity	scale of percieved level of PA created using activies of varying intensity. Running, bicycling, swimming are 'high' PA level; walking, playing catch are 'medium' PA level; sleeping and studying are 'low' PA level.

2.6 Open Questions and Concerns

The dangers of unintended consequences via misinterpretation may become more serious as we develop systems with more powerful behavior change methods and as we first explore these uncharted methods for providing feedback via avatars. The problem of misinterpretation becomes an even greater concern for the described systems since the method of interpretation for these outputs to the user is no longer something which is taught, but can be entirely dependent on the user’s perception of the avatar. For example, a user who may suffer from distorted bodily self-perception may interpret the body shape of the avatar much differently from the norm.

One of the largest challenges remaining for an implementation of human-in-the-loop feedback with avatars is that the method of mapping inputs to outputs (the algorithm itself) may need to vary from application to application and from user to user. Due to the large search space, identifying the best mapping from input to output may require significant iterative design and personalization along with advanced analytic methods such as the use of control systems engineering and dynamical systems modeling [78].

In conclusion, the design process for avatar interfaces is given some foundation through the use of described methodology, but much more exploration is needed to address the questions posed throughout this chapter. Through additional implementations guided by behavior change theory, it is our expectation that avatars will prove an extremely powerful tool for behavior change science. However, additional research into the modeling and analysis of data for systems which adapt to users’ needs and deliver complex interventions in-the-wild are prerequisite to a deeper understanding of user-avatar interaction.

CHAPTER 3: GLANCEABLE M-AVATAR

Chapter two explored the use of avatars as a behavior-change interface; to expand on this front, chapter three presents data from a self-avatar-based, glanceable intervention targeting physical activity behaviors in adolescents (aged 11-14). Details of the trial study (n=13) are presented followed by results which highlight some of the challenges facing extant JiTAI study design and analysis methodology.

Self-avatars (or just avatars, for brevity) can take many forms. They can be as simple as a picture of a user on a social network [79] or a far more complex, animated character in a virtual world whose actions can be controlled by the user [80]. Avatars serve as facilitators of social interaction in virtual worlds by providing bodies for users to manipulate to express themselves and communicate with others (not unlike using one's body to communicate nonverbally in the real world) [81, 82]. Avatars are prominent in video games, and exercise games (or Exergames). Examples include Wii Sports Boxing and Microsoft Kinect Adventures: Reflex Ridge, where the user's real-time movements are tracked and transformed to similar movements by the avatar.

Another emerging user interface for behavior change is the ”glanceable” [83], ”ambient” [84], ”always-on” [85], or ”peripheral” [86] displays delivered via mobile phones to improve health behavior choices. The ubiquitous nature of these mHealth wallpapers allow individuals to remain constantly in-tune with their physical activity goals and health information, which is speculated to cause activation of behavioral goals and improve self-regulation of planned behavior [87].

Yee, Bailenson, and others have shown that the behavior of a participant in a virtual world can be influenced by their avatar's physical characteristics in both laboratory settings [81] as well as in 'real-life' online interactions [73]. It has also been shown that changing the behavior of a virtual representation of one's self can be used to positively affect opinions on health and physical activity [57]. However, these avatar effects have not been demonstrated in a mobile context.

There is a call in the research community for evaluation of the potency of this effect outside of immersive virtual environments [57]; this trial study explores the outer edge of the domain in which avatars may have an effect - glanceable visualizations in which the user is very loosely tied to their avatar. More specifically, this chapter examines the theoretical fidelity of a system designed to test the "doppelganger effect" applied to overall physical activity within a mobile context. The doppelganger effect, much like the aforementioned proteus effect, is a way in which a user's self-avatar can alter the user's behavior. The doppelganger effect is observed when a user is motivated to copy the actions of a self-like avatar. For instance, a running avatar might inspire the user to be more physically active. In contrast with the proteus effect, the doppelganger effect applies motivation through a difference in avatar action and user action, whereas the proteus effect influences self-perception through a difference between user and avatar appearance.

Our mAvatar application enables testing of the doppelganger effect through observation of changes in user physical activity in response to a pervasive mobile display which shows rudimentary user doppelgangers performing actions of varying physical activity levels. To ensure an effective test of the concept in a mobile context, there are a variety of important design issues to consider such as 1) how and to what degree can a user customize the avatar, 2) the determination of an appropriate delivery mechanism to provide the intervention, 3) the conscious and unconscious connection between the avatar and the users, and 4) the perceived and unperceived influence of the avatar on a person's behavior. All of these points are essential for establishing acceptable theoretical fidelity to support a proper test of the concept within a system.

3.1 Methods

Participants for the study were recruited using a flyer targeting parents posted on a university campus and distributed via various university mailing lists. Approximately 40 responded, and approximately 20 scheduled to learn more about the study. None were excluded from the study. To reduce subject-side bias on behavior, participants were told that we were interested in using avatars to influence behavior, but were not told that we focused on physical activity specifically.

Participants carried the phone and a FitBit electronic pedometer for at least 8 days while they went about their everyday lives. Throughout the observation period the smartphone displayed a glanceable, avatar on its background wallpaper. The avatar was personalized by superimposing a photo of the participant's face onto the cartoon avatar’s head. Each day, the avatar adopted one of two types of behaviors: either physically “active” (e.g., walking, playing basketball) or “sedentary” (e.g., watching television on the couch, using the computer). Participants were not told how the avatar chose behaviors.

Physical activity measurements are continuously captured using a validated, smartphone-based passive physical activity monitor (mMonitor [88]) as well as a Fitbit One pedometer [89]. The phone-based activity monitor labeled every minute as one of sedentary, light, moderate, and vigorous physical activity. The fitbit provided step-counts at a frequency of one per minute.

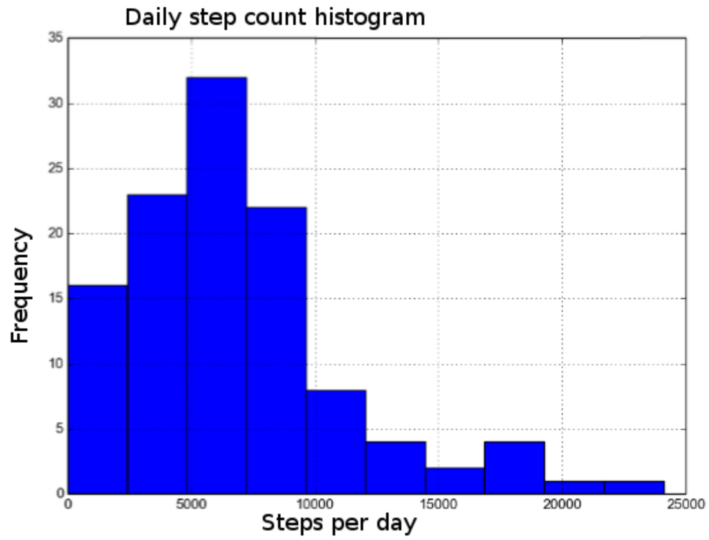

Figure 4: Histogram of day step count total.

Measurements of avatar influence were taken in the form of phone view logs. The amount of time the avatar is displayed to the phone user was recorded by logging visibility change events from the android operating system. These logs were tested to be a very reliable measure of when the avatar is and is not visible to the user.

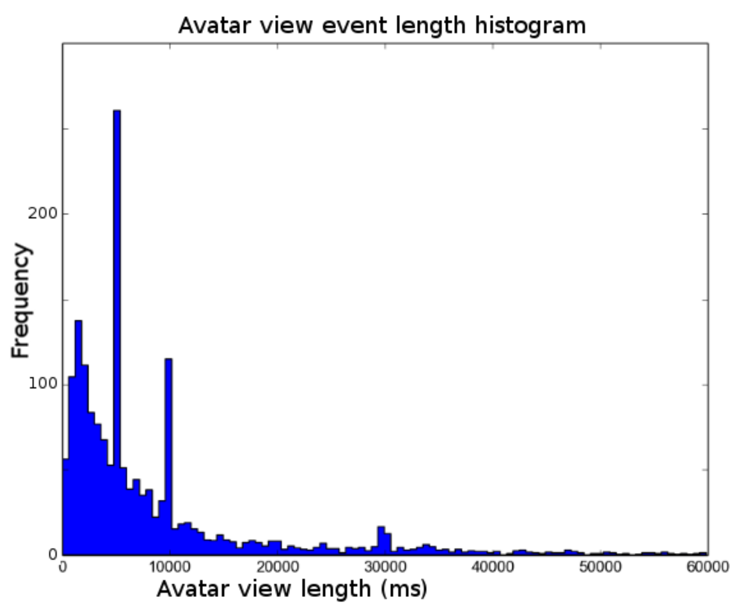

Figure 5: Histogram of avatar view lengths.

3.2 Data Processing and Analysis

3.2.1 Avatar-Intervention Dosage Score

Analogous to the dosage of a medication, a measure of the amount of avatar intervention delivered is introduced as “intervention dosage”. In order to quantify dosage of avatar intervention, a “Avatar Intervention dosage score” is introduced. The score is computed using the amount of time the avatar is viewed. Viewing a physically active avatar results in a positive score, and viewing of a sedentary avatar results in a negative score.

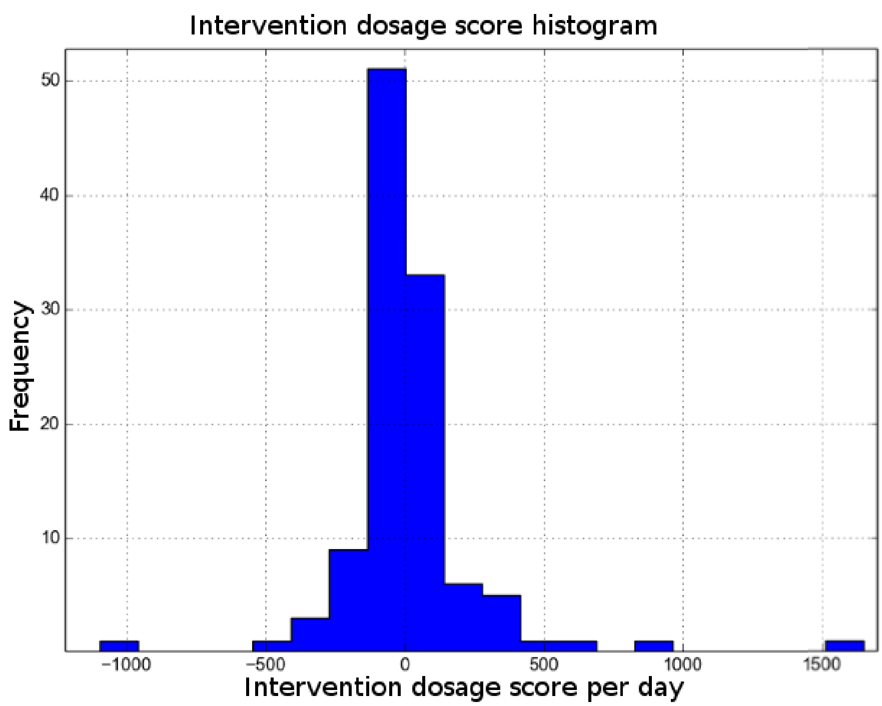

Figure 6: Seconds of avatar intervention dosage per day.

Additional user-experience data was collected in the form of a demographic survey prior to the monitoring period, a user-avatar relationship survey immediately following the monitoring period, and a user-experience interview to conclude participation. Minute-level step counts for

Fitbit were downloaded via Fitabase [90], and GPS location was collected using a free gps logging application called GPSLogger.

3.2.2 Data Processing

3.2.2.1 Removing Outlier View Times

Outliers were identified and removed in measurement of the amount of time that the avatar is displayed to the user. Because the software measures only the amount of time the avatar is visible on the screen, cases where the screen is left on (such as while charging), appear as unreasonably long view times which can dramatically skew analysis. These events (defined as view times longer than 60 seconds) were removed from the data and replaced with short view times (5 seconds) at the start and end of the anomalous view time.

3.2.2.2 Accounting for Every-Other-Day Events

During a concluding interview, participants were asked about their weekly schedules, specifically focusing on physically active events which might take place Tuesday-Thursday or MWF. Sports team practices and PE class schedules were asked about specifically. These kinds of events are of particular interest because they may skew the every-other-day, within-participant study design. An analysis of participants with noted every-other-day physical activity schedules is needed to show that this potential confound is not a cause of any observed effect.

3.3 Results

Because the measurements in this experiment take place “in the wild”, many potential confounds must be considered. The use of within-subjects comparison means that many of these confounds should cancel out, but some additional data analysis techniques were used in an attempt to guide future studies of user-avatar theory.

3.3.1 Macro-scale ‘Intervention’ Effect on the Raw Data

In theory, the doppelganger effect should motivate participants to mirror the behavior of the avatar. Thus participants should be more physically active on days the avatar is physically active and more sedentary on the days the avatar is sedentary. The most straightforward test of this assertion is to use a paired t-test on each participant’s sedentary and active day averages. Due to various technical issues, incomplete data from two participants had to be excluded from this analysis.

As can be seen in Figure 7, on average participants were more active on avatar-active rather than avatar-sedentary days. A paired t-test performed on the average step count from avatar-active versus avatar-sedentary days does not show statistical support for the hypothesis with a p value of 0.35. The relatively small sample size of this pilot study along with the high variability of the in-the-wild data collected made this weak effect undetectable through standard analysis. Future comparison of effect sizes using differently styled avatars can be designed to explore the Doppelganger effect or other avatar-user interaction theories. The effect of a more realistic, interactive, or customized avatar can use this study as a baseline in order to characterize these moderating variables.

Figure 10 (top) shows each day as a point with aforementioned avatar-dosage score on the x-axis and step-count on the y-axis. A positively-sloped correlation confirms the hypothesis that step-count should increase with active-avatar exposure and decrease with sedentary-avatar exposure. In order to explore a possible subgrouping of participants figure 10 (bottom) shows linear correlation attempts for all participants.

3.3.2 Micro-scale Intervention Effect

These data indicate that the physical activity of an avatar wallpaper may influence the physical activity of its user at the day level, but the effect was too small to detect in this study and it is still unclear how quickly this effect may begin and fade. Through analysis of participant physical activity after each “avatar view event”, we can begin to get a better picture of the latency and delay involved.

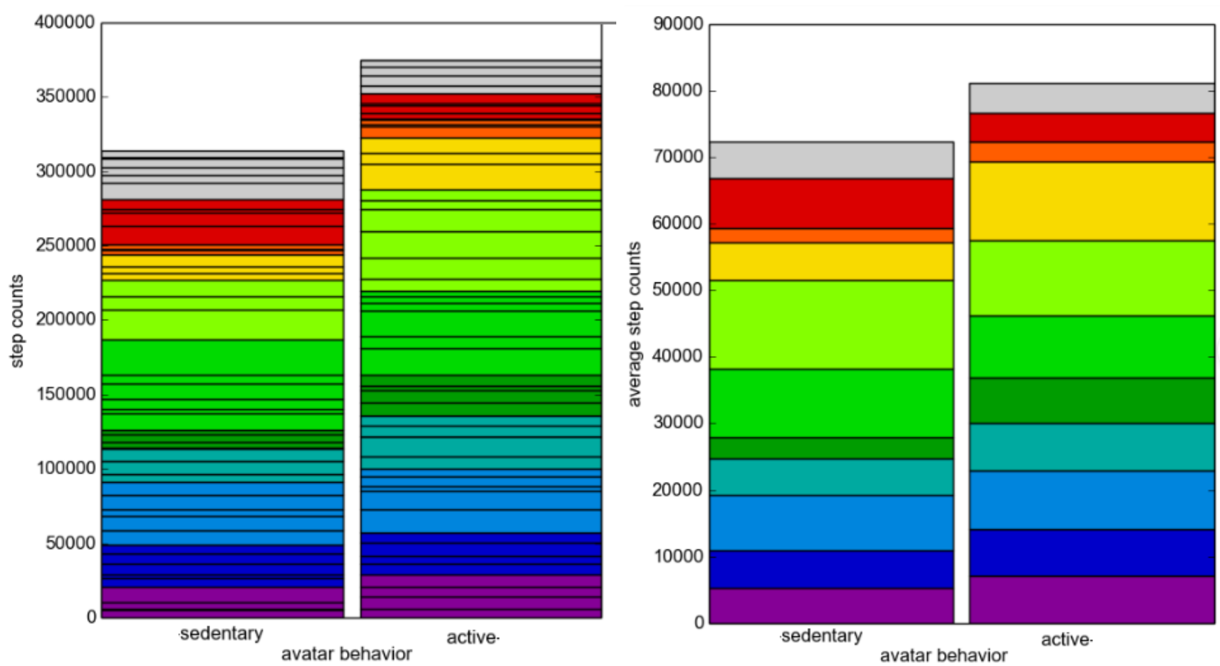

Figure 7: Stacked bar charts of active vs sedentary day step counts. Right shows average for each participant, while left shows each individual day. Each color represents a participant.

3.3.2.1 Defining an Avatar View Event

Given a list of avatar display times, we must identify a point in time at which the avatar view event occurs. There may be some minimum amount of view time required for an avatar to affect its user, but in this analysis we consider any time period >0.01 second to be sufficient.

An “avatar view event” is said to occur when the avatar is displayed on the screen where no other view event has occurred in the prior 2 minutes. This minimum time between view events is referred to as the “recovery period”. If another view event exists in the 2 minutes before, then this avatar viewing becomes an extension of the previous view event. Thus, only one avatar view event can occur within a 120 seconds interval, regardless of the number of times the screen is toggled on and off. This definition is given under the assumption that delays and latencies of the avatar's effect on physical activity are no smaller than 2 minutes.

Additionally, times when the avatar is displayed for more than 1 minute continuously are considered to be unrealistic and have been removed as explained in "Removing Outlier View Times". This is based on the assumption that users will not find the avatar interesting enough to warrant a continuous gaze of more than 1 minute.

In order to reduce variation in the signal following the "view event", the moment in time the "avatar view event" is said to occur is at the last instant that the avatar is displayed within the particular event, whenever the participant has stopped using the phone. The "duration" of the event is the time period prior to this during which the avatar is being viewed.

The placement of the "view event" is demonstrated using the sample data in figure 8.

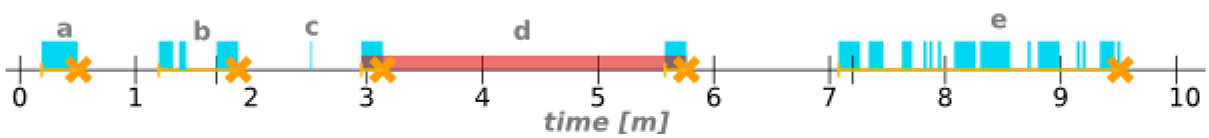

Figure 8: Demonstration of the placement of "view events". Events marked with x given binary avatar-view data time series in scenarios a, b, c, d, and e.

1) (a) "glance" - wallpaper is viewed between 1 second and 60 seconds in length. This represents cases where the phone has been glanced at once and then put away. Use-cases include: checking the time, looking at the phone only for the avatar, and checking for notifications.

2) (b) "usage" - wallpaper viewed multiple times with insufficient time between them, thus they are lumped into one event. This event represents a typical phone usage session during

which the user may change between apps or screens, seeing the avatar background briefly during the transition.

3) (c) “short fault” - wallpaper is viewed for less than minimum view time, no event.

4) (d) “long fault” - unrealistically long view time (red) is replaced with shorter view times at beginning and end.

5) (e) “long usage” - nearly continuous usage of the phone over an extended period of time is observed as many intermittent avatar views; a view event is placed at the end of the usage.

One additional caveat of note: the end of a view event may occur at any point in time, but fitbit step counts begin only at the start of each minute. View event endings are thus “snapped”(rounded) to the nearest minute, meaning that there is up to 30 seconds of variation in the alignment of view events.

3.3.2.2 The Dynamics of Post-Avatar-View Step-Count

In an attempt to visualize the dynamics of step count following an avatar view event, figure 11 shows all participants’ step counts following all 772 active-avatar (red) and 784 sedentary-avatar (blue) view events. Also shown is the average step count in the minutes following the event. In this view, active-avatar and sedentary-avatar effects appear very similar.

3.3.3 Subgroup Analysis

In order to better explore the differences between participants with positively and negatively linear sloped correlations between avatar dosage score and step-count (as separated in figure 10 bottom-left and bottom-right), the difference between average steps following active and sedentary view events was integrated over the 180 minutes following the event to provide a simple score which may indicate the degree to which the avatar’s physical activity inspired participant physical activity. These scores were then compared to several survey metrics which we believed may have some impact on the potency or direction of avatar influence. A summary of these results is shown in figure 9. Since no strong correlation are apparent, we conclude that these metrics are insufficient to explain the between-subjects differences in the data. Through inspection of participant interviews, a potential moderating variable was identified mid-study. Because participants were not told how the avatar would choose behaviors, participants were open to interpret the avatars' behavioral choices in two ways: 1) the avatar mirrors their own behavior, or 2) the avatar is suggesting behaviors for the participant. Some connection between participant behavior and the behavior of the avatar was expected by nearly all participants, but reported expectations were split between these two interpretations. This potential moderator further complicates analysis because each interpretation is expected to influence physical activity in opposite ways. If the participant believes the avatar is mirroring their own behavior, then avatar-sedentary behavior may cause increased physical activity due to the participants' heightened awareness of their own behavior. In this way the avatar could act as a simple biofeedback mechanism. In contrast, the second interpretation predicts participant physical activity behavior to correlate positively with that of the avatar as originally hypothesized.

A) Avatar Susceptibility vs post-event steps

B) Perceived avatar influence vs post-event steps

C) Perceived avatar control vs post-event steps

Figure 9: Average difference (active-sedentary view event) in participant steps. 180m following avatar view event vs potentially moderating variables measured by survey (A) avatar susceptibility score, B) perceived avatar influence, C) perceived control over avatar behavior)

In conclusion, the mAvatar trial study created an interesting dataset with unique data analysis challenges. Although support for the doppelganger effect in a mobile context could not be shown using standard analysis techniques, a more in-depth look at the minutes following intervention delivery may lead to a better understanding of the avatar-effects this study set out to explore.

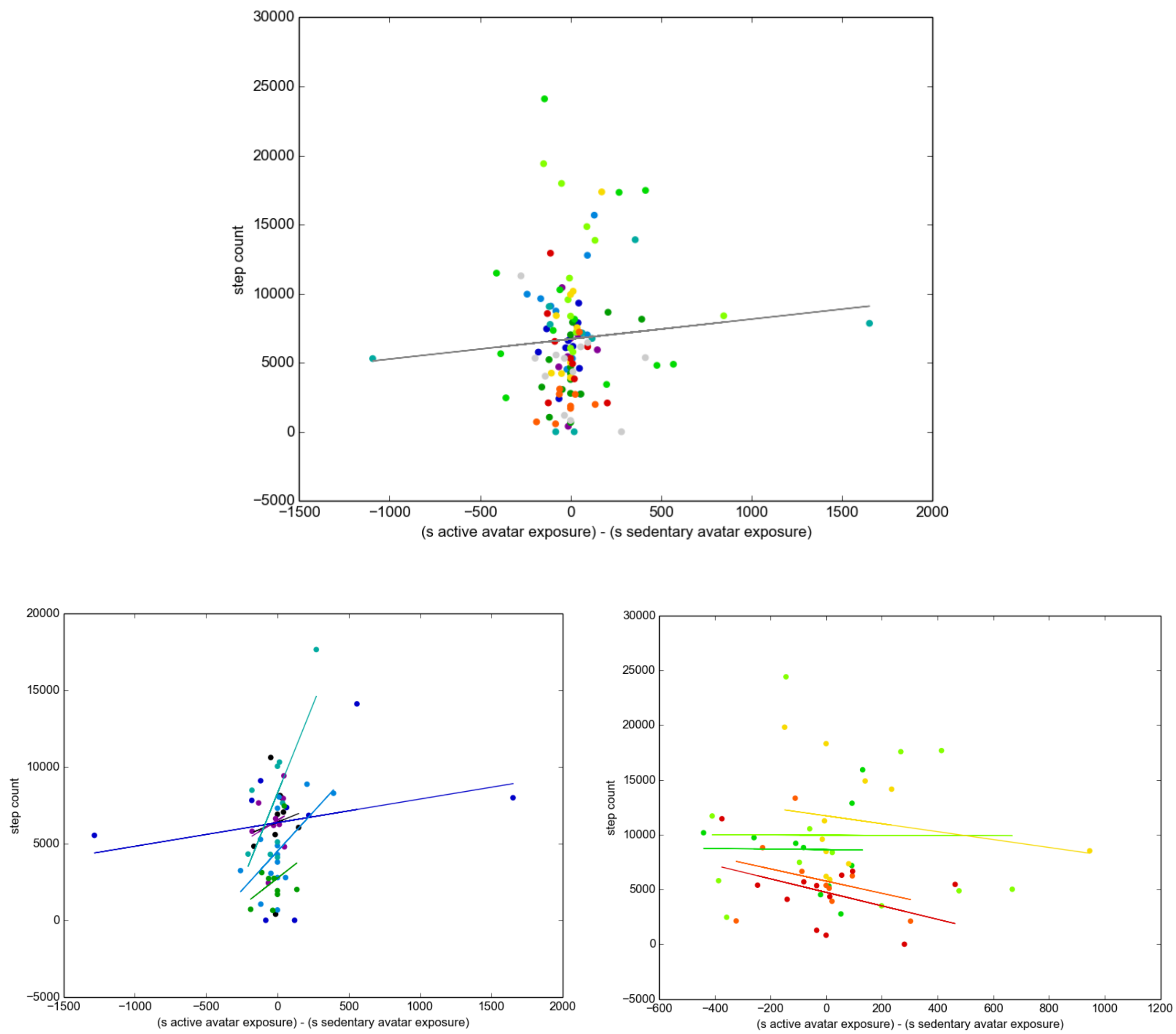

Figure 10: Scatterplot of all participants daily step counts vs avatar exposure score. Top shows correlation across all data. Bottom shows correlations for each participant, split into positively-sloped (left) and negatively-sloped (right).

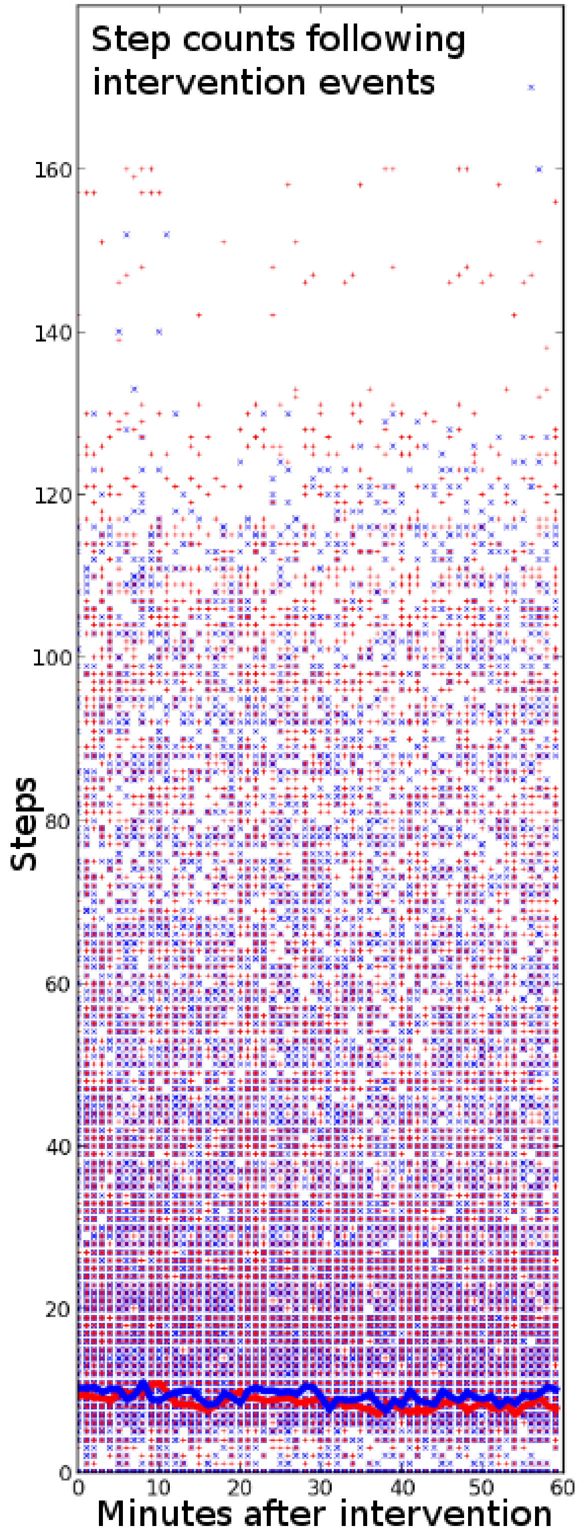

Figure 11: Comparison of step-counts in 60 min following avatar view events. Active shown as red + and sedentary as blue x. Individual values marked as well as average lines shown.

CHAPTER 4: INTERVENTION-VIZ

Though the potential applications of JiTAIs are numerous, there remain significant challenges to be overcome by the research community before the potential of JiTAIs can be unlocked. As shown in the previous chapter, methods for evaluating the efficacy of a JiTAI are not yet established and methods for utilizing computational models of human behavior are even less developed [91]. Existing behavior models appear inadequate to inform state-of-the-art intervention development [50]. Conventional methods of analysis do not offer the level of detail needed to explore the implicit dynamics of JiTAIs, and behavioral theorists need methods and tools to help understand the dynamics of behavioral responses to a stimulus. Applicable methods of intervention analysis and data visualization have been slow to reach behavioral researchers, dramatically limiting their ability to develop of state-of-the-art behavioral theories to address these shortcomings. Without addressing these open questions, models cannot be used to effectively predict or explain behavior in practice, the dynamical aspects of human behavior will remain ignored, and applications will remain artificially limited by the unnecessary complexity of decision rules which are used to implicitly codify models of behavior in existing proof-of-concept systems.

In this chapter methods for analysis of the minute-level dynamical response to a behavioral intervention are outlined. The impulse response of a physical activity intervention is explored and data visualizations which provide insight into the dynamics of health-related events are demonstrated. We evaluate the visualizations from a JiTAI developer's perspective using three datasets, focusing on what research questions are addressed by each visualization, where there is uncertainty in the meaning of the visualizations, and the strengths and weaknesses of each approach. These methods, when combined with a computational modeling approach to understanding human behavior may enable behavioral scientists to formulate more accurate and more application-ready models, leading to more effective behavioral interventions.

4.1 Related Work

Much work exists on both behavioral intervention analysis and event-based time series visualization. However, little existing work addresses the dynamics of a numerical variable's response to a behavioral intervention event.

Event-based analysis is an important topic for business applications as well as in the health domain. Many of the methods applied to analyze consumer behavior can be applied to the health domain. "Lifelines" [92] allow for the exploration of health events in a series for one individual, and new research in event sequence analysis [93], including analysis of event patterns [94, 95, 96] and the relation of multiple symptoms [97], helps researchers examine outcomes on a "macro-scale" across many participants by aggregating records into a single view. Similarly, the problem of identifying patterns at multiple time scales has been partially addressed through clustering of time series [98], and methods for exploring the "paths" traversed by many individuals between many event types and statistical analyses to highlight relationships between events has recently been established [99]. These methods provide useful abstraction at the population level and allow researchers to explore correlations and state transitions of population subgroups, but these methods are most effective for discrete-state measures. For measures with many states or continuous variables (such as many behavioral measures in JiTAI applications), it becomes more difficult to provide statistical support for a particular state-transition hypothesis. The methods presented in this article fill this gap in the literature and focus specifically on visualization of the dynamical behavior of continuous variables surrounding a particular event or aggregation of events.

4.2 Example Application: Physical Activity

As an example application to demonstrate the strengths of the proposed visual analytics two empirical datasets will be used, each with a minute-level metric of physical activity and intervention events delivered throughout a period of several days. In both studies interventions were delivered with the intent of increasing participants' physical activity, and responses to interventions varied between participants and delivery contexts. In addition to these data, a control dataset with known intervention responses is included for comparison.

These datasets provide a good test bed for application of the methods presented here. The interventions in these datasets are all expected to affect the level of the target behavior, but the dynamics of the response may differ greatly. The differences in the chosen datasets serve to highlight the strengths and weaknesses of methodologies outlined. The n-of-one control dataset with a strong intervention acts a baseline with predetermined response characteristics which should be easily identified by our analysis. The KNOWME data represents a JiTAI with a study-wide effect and multiple behavioral measures. Lastly, the mAvatar study data shows less prominent effects study wide, but has potentially interesting subgroups for exploration. Additionally the mAvatar data is unique in that it contains two interventions targeting the same theory, but influencing in opposing directions. A summary of each dataset is shown in Table 2.

Table 2: Summary of datasets analyzed.

Data Set	n	Length (days)	Intervention	measures
control	1	14	N/A	Step count
KNOWME	10	3	SMS Message	HR, Accelerometry
mAvatar	11	8+	Glanceable avatar display	Step count

4.2.1 Control Dataset

The control dataset is the result of manual recording of one participant undergoing an imaginary, very potent intervention. The participant remained sedentary for an interval ranging from 5 to 120 minutes. Then the participant was physically active for a period of no less than 5 minutes. Physical activity was recorded as a step count using a Fitbit One pedometer at a frequency of 1Hz. Thus, the mock intervention is delivered with 100% efficacy and should therefore be easily identified in the data.

4.2.2 KNOWME Study

In this study ten teenagers (mean age 16.3 +/- 1.7 years) were asked to carry a smartphone and wear an accelerometer and a heart rate monitor for 3 days. Physical activity was measured continuously and was monitored in real time using the KNOWME system [100]. When a participant had been continuously sedentary for two hours, a personalized SMS text message was sent to their phone. Each text message is manually crafted to prompt the participant to be more physically active. The text message prompt is expected to cause an increase in PA within minutes to hours after the intervention. This physical activity increase should be detectable in both the accelerometer data as well as the heart-rate data, with the heart-rate data lagging only very slightly behind accelerometer.

4.2.3 mAvatar Study

An alternating treatment design is used to examine participant behavior over a period of 8+ days in order to test the effect size of an avatar-based live wallpaper deployed on Android phones [71]. Participants (n=11) aged 11-14 were exposed to a simple, animated cartoon avatar on their mobile device showing alternating levels of PA. Each day the avatar would either be active (playing basketball, running, bicycling) or sedentary (watching TV, on a computer, or playing video games). Fitbit One electronic pedometers were used to estimate participant levels of physical activity via step count. Depending on the participant's interpretation of the avatar display, one of two effects are expected:

1) The participant believes the avatar is reflecting their own behavior, increasing their awareness of sedentary behavior, causing the avatar to act as a biofeedback mechanism, and boosting their PA.

2) the participant believes the avatar is suggesting how they should behave, possibly inducing the Doppelganger Effect [101], and raising their physical activity to better match the avatar.

For participants in subgroup one a negative correlation between avatar and participant PA is expected. Conversely, for condition two a positive correlation between avatar and participant PA is expected. The dynamics of these two effects are uncertain, but it is hypothesized that effect one has a comparatively shorter delay and decay than effect two, which may be more cumulative in nature.

4.3 Methods

4.3.1 Highlighting Event Dynamics

Existing "macro-scale" methods can determine if an intervention has a significant influence over our target behavior, but they do not give much insight into how the event has an effect over time. In order to explore the dynamical response of an intervention, the shape of the input signal must be defined. In most cases an intervention can be represented as an impulse signal. Using this representation the impulse response can be calculated as the cross-correlation between the intervention signal and the behavioral measure. Figure 12 shows the result of cross-correlation between the intervention input and the heart rate signal across all participants in the KNOWME study.

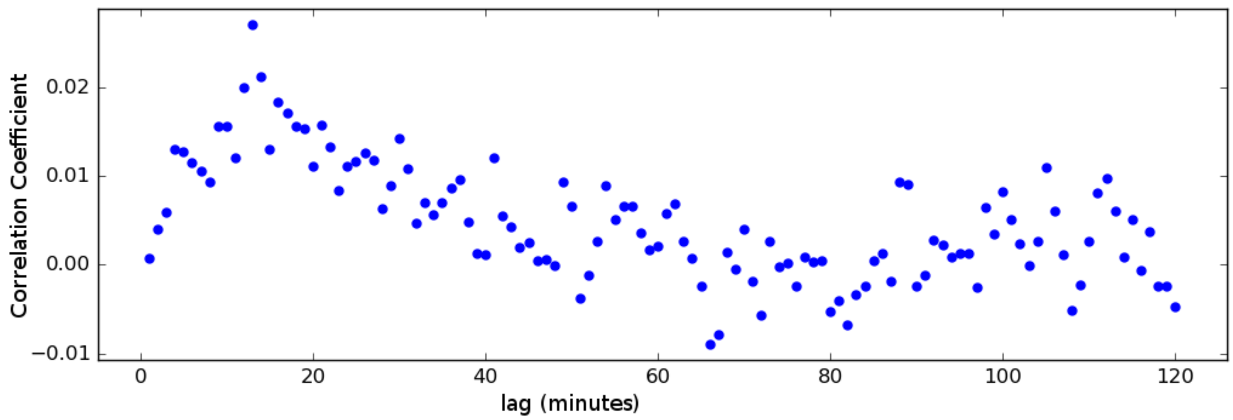

Figure 12: Cross-correlation function showing study-wide heart rate response. Intervention from the KNOWME dataset would be at 0 minutes of lag.

The dynamics surrounding a particular event can also be shown using a raw time series. The instance or span of the event is marked on the time-axis and the value of the behavioral measure (physical activity in this case) is encoded in the height at each point in time.

Figure 13 shows the case where an event instantaneously causes permanent change in the target behavior, but in the many cases the intervention will have a temporary effect on the target behavior and will have some delay before setting in.

These intervention response dynamics shown in figure 14 are critically important for JiTAI developers. Each participant's record can be inspected individually, and events of interest can be marked. Since this examination is taking place over many series, it is prudent to utilize sparklines [103] or horizongraphs [104] to allow for examination of many series simultaneously.

4.3.2 Event-time Alignment

Plotting individual events one-by-one allows a researcher to explore the idiographic details of that particular event, but in order to draw out generalizations across groups of events (be it by participant, context, or another selector) events must be plotted relative to the time of the event, rather than the start of the study. By time-shifting the data view so that each intervention event falls at t=0 in a time-series, we can view many events on a common time frame.

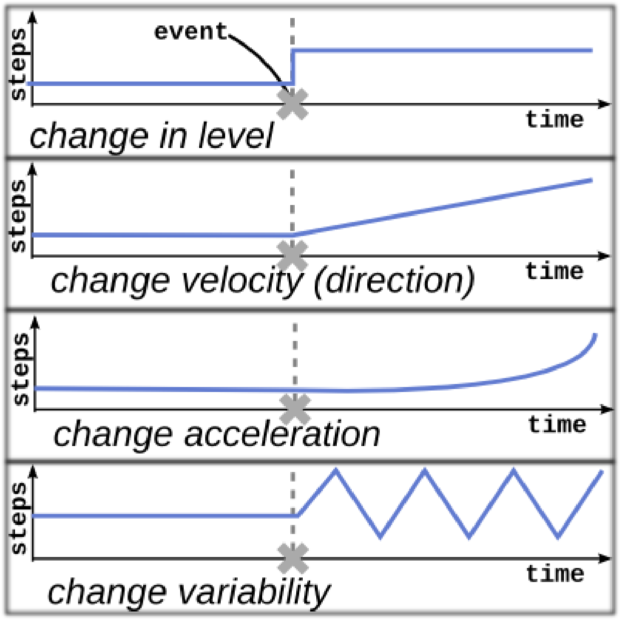

Figure 13: Theoretical responses to intervention.

Adapted from Glass, Willson, & Gottman [26].

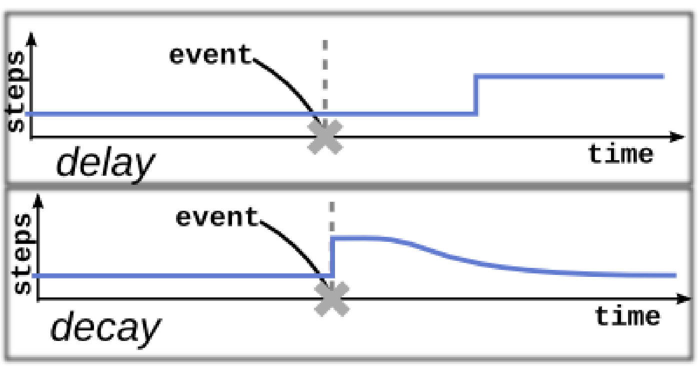

Figure 14: Level-change dynamic effects.

Adapted from Glass, Willson & Gottman [102].

Figures 13 and 14 give us sense of what an intervention should look like, but in reality individual variations in context completely mask the often small effect of an intervention (see figure 15). To a researcher looking at the plot of individual event responses in figure 15, it might seem that only the intervention plotted in purple was an effective intervention, acting with a delay of approximately 30m, and decaying rapidly 120m after the event. However, the control dataset includes interventions that were 100% effective by design, acting with minimal delay and beginning decay at 5m. Since the data has been time-shifted to place the time of event at t=0, an average across all series will reveal nomothetic trends across all events. When looking at all events individually, it is difficult to spot any pattern in the series. When averaging across all event responses, however, a response is evident, and the purple series is exposed as an outlier rather than the only instance of successful intervention.

This approach can be taken for all events in one participant's time series to characterize that participant, or can be applied across participants to characterize a more generalized response to the intervention. In fact, a subset of groups can even be selected and analyzed in order to enable advanced subgroup analysis.

4.3.3 Gauge Effect Size

It is difficult to judge if a sudden increase of, for example, 10 steps/min is statistically relevant for a given participant in a particular context. To help address this, we include an additional y-axis showing the mean and standard deviation of the series to give an increased sense of the significance of this effect relative to data which may be out of frame. In addition to the nearly immediate response in figure 16, a longer-lasting effect reaching out to approximately 180m after the event seems to be boosting step count, though the all-events view in figure 15 as well as the stacked-events display reveals that there are two outlier events which may be the sole cause.

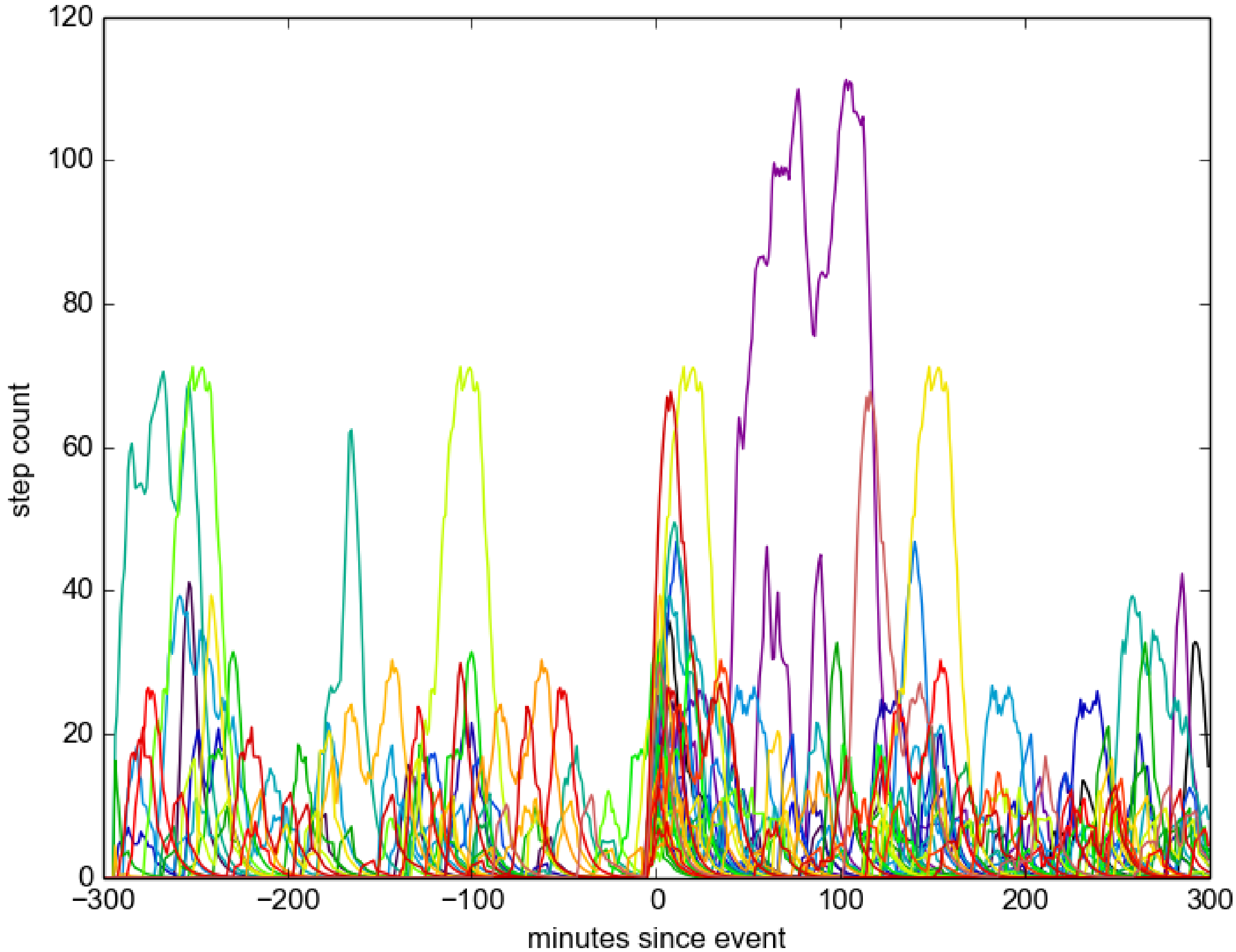

Figure 15: Aligned event responses surrounding the control intervention.

4.3.4 Stacking

To address the shortcomings of using averages, we show all individual events stacked on a single graph. This aggregation method yields the same shape, and the y-axis can be easily normalized to match our average series by dividing by the number of events. While still evening out random contextual influences, this visual also provides indication that the average result is not due to one outlier event, enables easy spotting of missing data or faulty sensors, and gives some indication of the number of events considered. For an n-of-one dataset such as the control dataset, events can be graphed with a unique color.

In figure 16, event colors are chosen based on the order in which they were observed. This encoding scheme may in some cases reveal habituation to an intervention if the later colors show decreasing effect magnitude. Color mapping of events can also be used to visually group events based on time of day, location of the event, or participant.

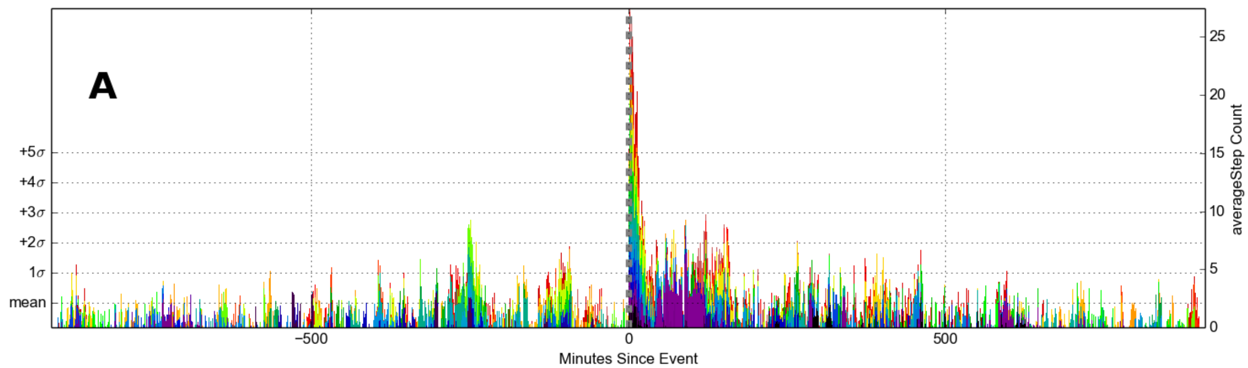

Figure 16: Aggregated step counts surrounding control intervention events. Shows aggregate event response dynamics and individual variations across events.

For a plot of many participants, encoding participant in color allows the visual to display both event-level and event-group-level detail in addition to the overarching response. Figure 17 shows the difference between a plot of various average response lines and the stacked area plot of figure 16 using the KNOWME dataset. The thin lines in figure 17 represent the response of each participant to the event averaged across all events for that participant. The thick gray line shows the average across all participants' average series. The stacked bars in figure 18 are colored by participant ID, and each bar represents one unique event - stacked in order of event incidence. This allows researchers to search for both participant outliers within the set as well as event outliers within each participant. For instance, it is clear that the participant shown in purple responded to the intervention, but we can also see that this effect is largely the result of a single event within the participant's series. This reveals that intervention was effective on average, while also showing that there exists some variable within participants moderating the efficacy of the intervention.

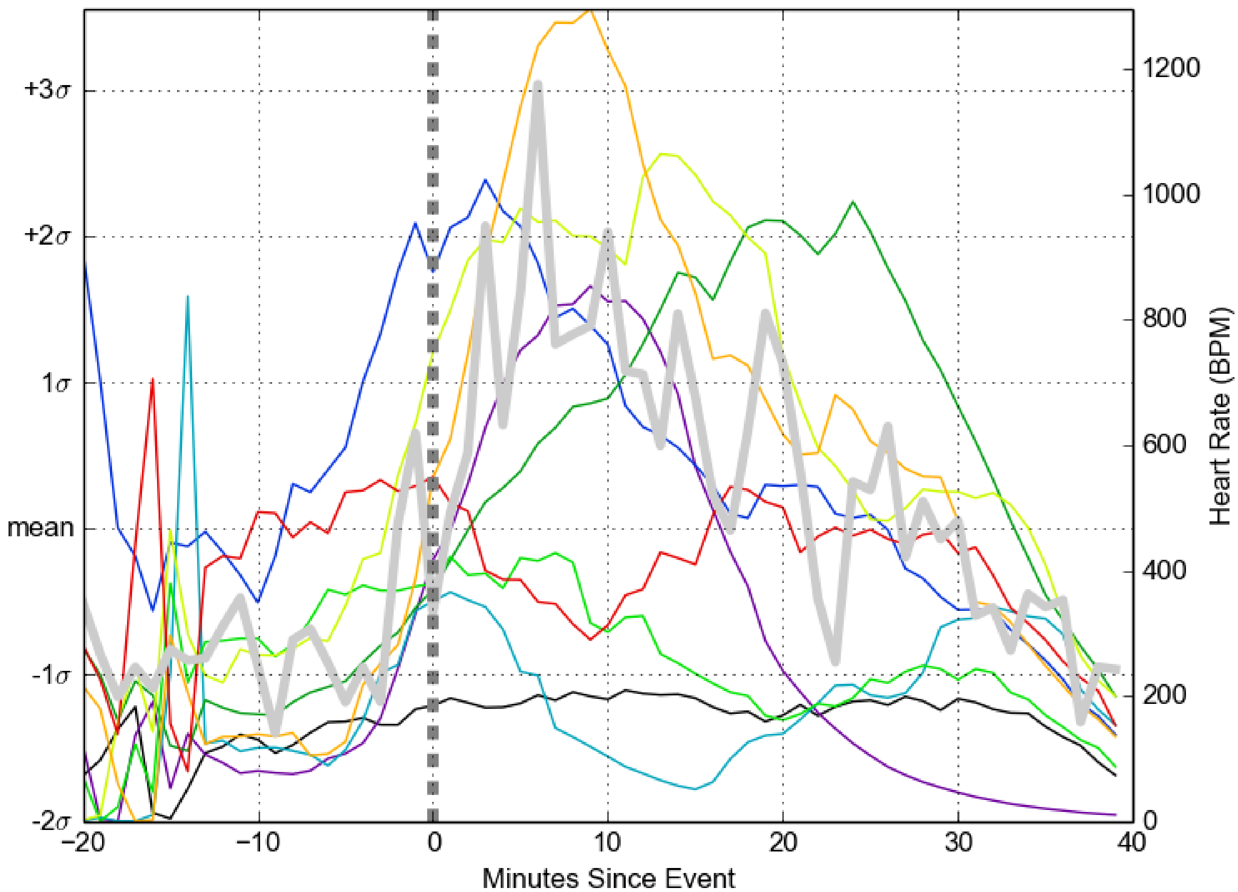

Figure 17: Average heart rate for each participant surrounding an intervention event. From the KNOWME dataset (smoothed over 15m rolling window).

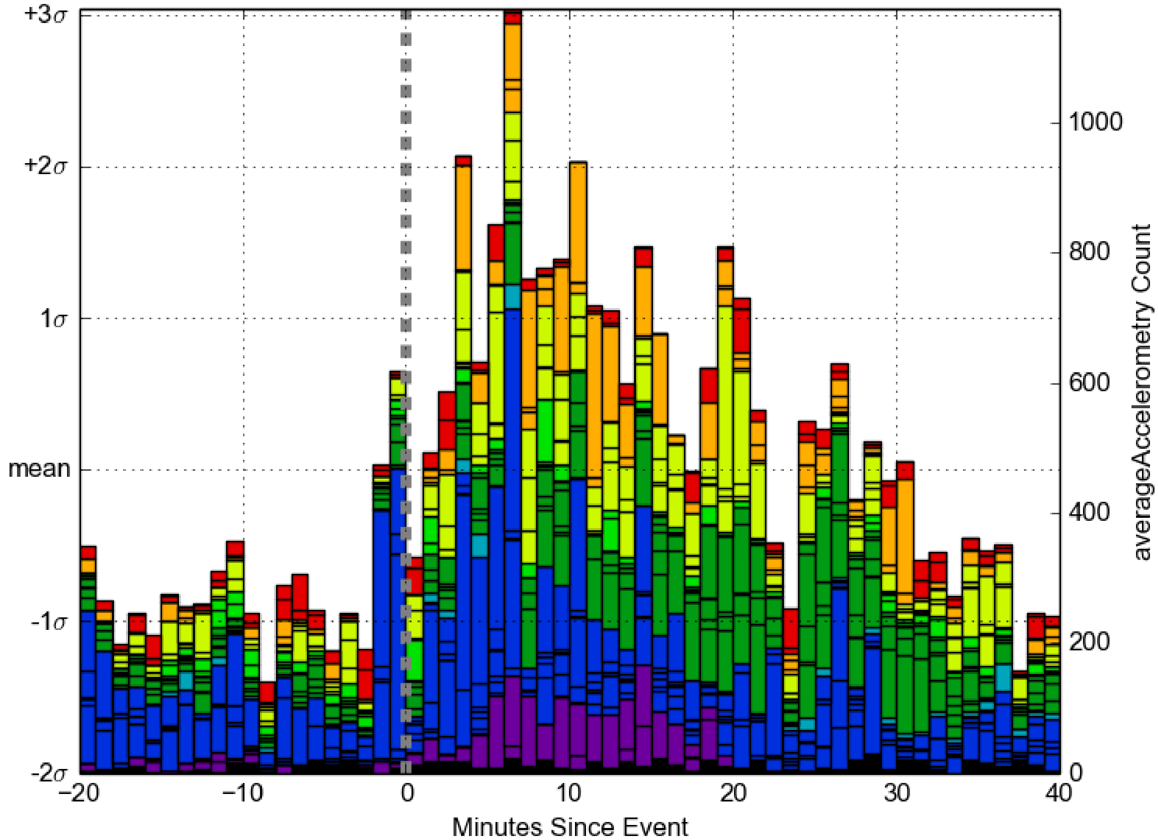

Figure 18: Stacked bar chart showing accelerometry for each participant. Using intervention events from the KNOWME dataset.

Figures 17 and 18 show an increase in physical activity following the delivery of a physical-activity-suggesting sms message. Though the behavioral measure differs from that used in the control dataset and figure 15, a comparison of the y-values in terms of standard deviation also reveals that this effect is less extreme than what we observe in the control intervention. The deviation from the mean as measured relative to the standard deviation gives a sense of how unlikely the signal is to be a random artifact, but detailed methods for evaluating the statistical likelihood of observing a particular shape are not covered here. For additional comparison to the control data, also consider the stackplot shown in figure 19. Though the highlighted windows are relatively small (to highlight the intervention response), much wider context around the event can be plotted, such as that shown in figure 16.

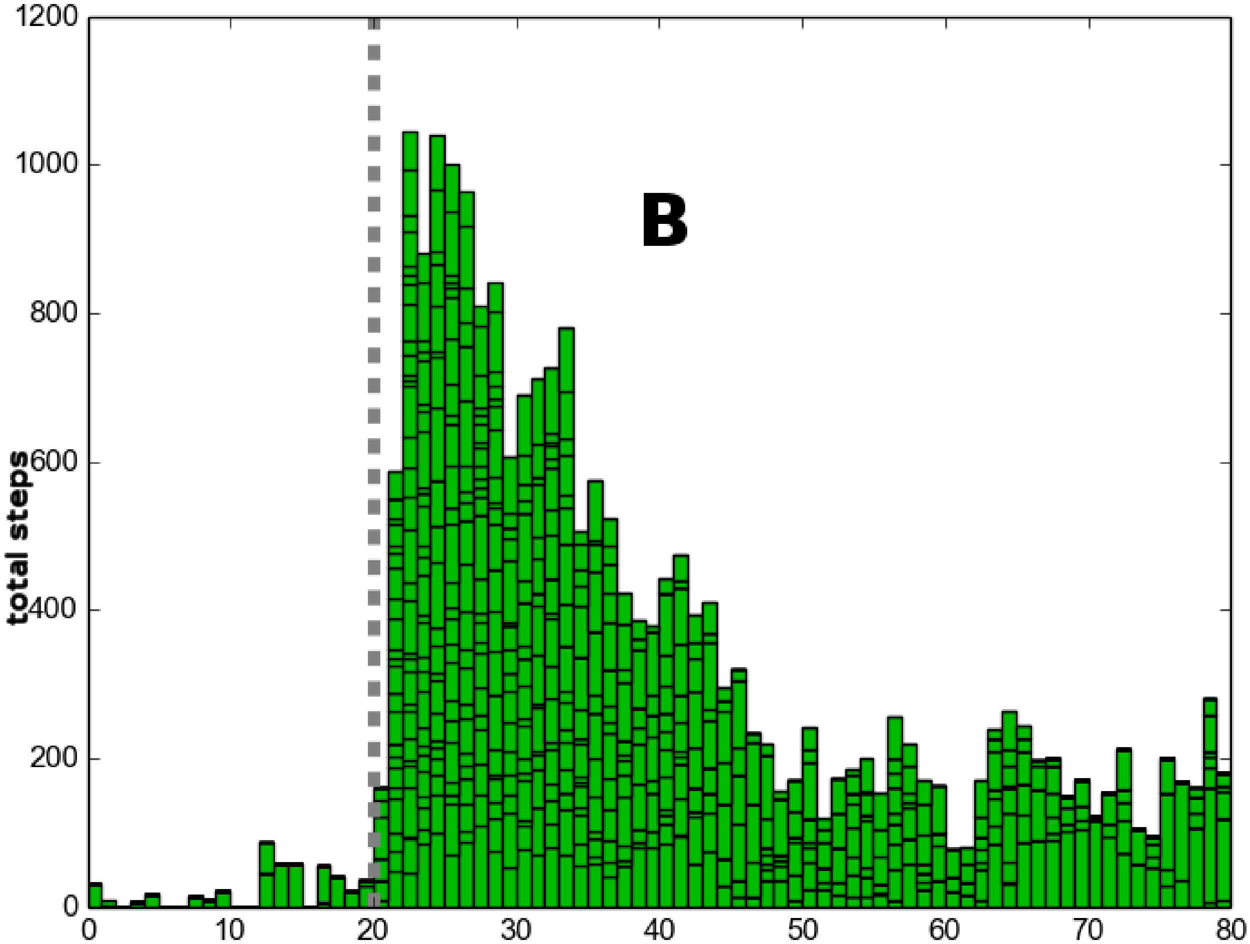

Figure 19: Aggregation of step counts showing dramatic response to the control intervention.

This same analysis is applied to figure 20, but with another variable in the KNOWME dataset, heart rate. Both the accelerometry counts and heart rate signals should act as proxies of physical activity. Note however, the different dynamics of each variable's response. Accelerometry counts are more directly tied to behavior - which can be erratic and non-linear, thus the dynamics observed are more volatile, while heart rate acts as smoothed function of accelerometry, responding less quickly and decaying more slowly than accelerometry data.

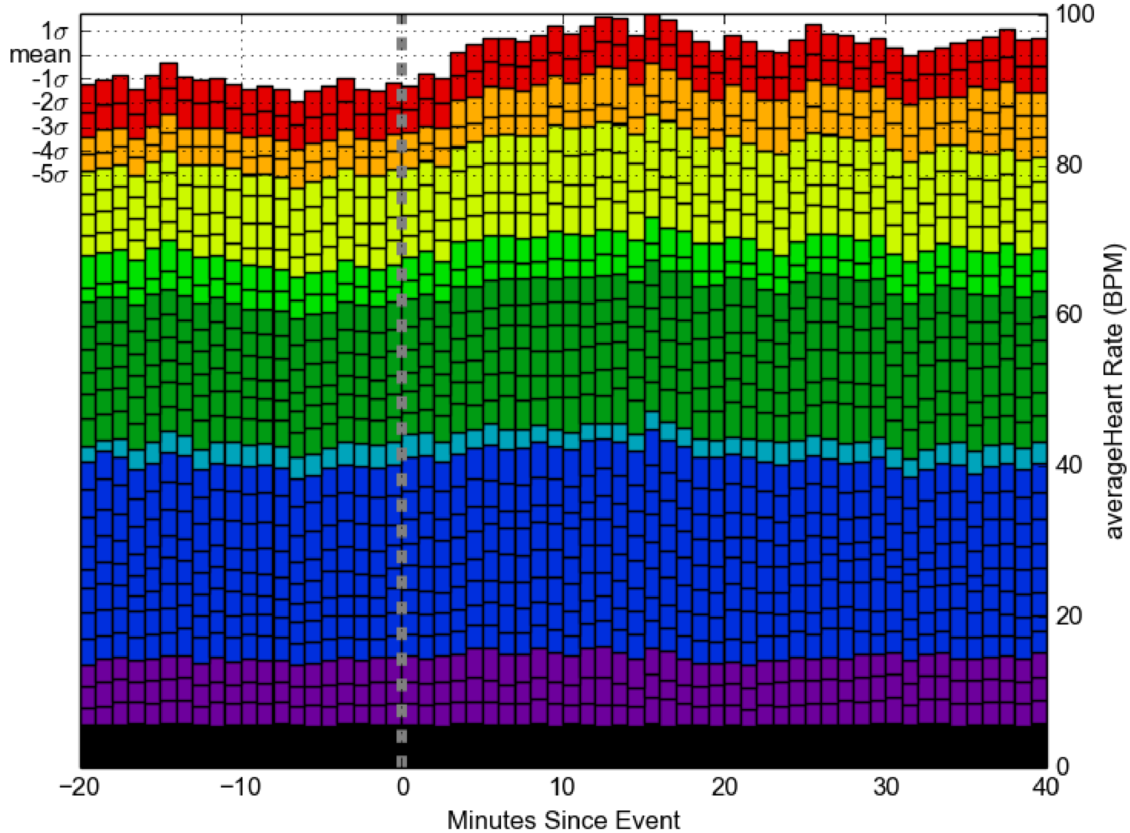

Figure 20: Heart rate data aggregated across KNOWME participants.

Shows a mild response to an SMS intervention.

The line graph allows for characterization of unique individuals, but the stackplot better highlights the overall effect and also shows the number of events considered.

4.3.5 Characterize Intervention Delivery Context

In some cases introducing a “control event” against which to compare the experimental event can help isolate the intervention from the context in which it is delivered. For instance, an intervention delivered on a mobile device is always delivered within the context of phone interaction. That is, the user is always using the phone when the intervention is delivered. It is possible that "using the phone" has it's own unique effect on the behavioral measure that may confound a comparison done against a “not-using-the-phone” baseline. Thus, using "phone use" events as a baseline against which to compare "phone use and intervention delivery" strengthens the chance that the observed effect is a result of the intervention itself and not the result of frequently concurrent contextual forces. For example, by looking at all times the phone was viewed in the mAvatar dataset, the average context of phone use can be characterized.

In figure 21, we see a notable increase in steps leading up to phone usage. It is possible that this increase - though it preempts avatar viewing - is indeed caused by the avatar. Consider, for instance, the unanimously reported case of participants viewing the phone with the explicit purpose of seeing how the avatar would be affected by their behavior. Thus a peak in physical activity may indeed be driven by the desire to illicit a response from the avatar, which is viewed only a few minutes later. This interpretation is quite speculative and other features of figure 21 are not so easily explained. It is clear, however, that this is not a flat baseline that we may expect to find on average, and exploration of dynamics surrounding the active and sedentary avatar viewings ought to subtract this baseline to account for the overlapping of this context-driven (rather than event-driven) signal.

4.3.6 Comparing Event Types

Aforementioned methods used to provide a contextual baseline of comparison for events can also be applied to allow for a comparison between two event types. By treating one event as the baseline, differences between the events can be visualized. Using this paradigm, nearly equivalent event responses will have a near-zero difference. Positively-valued areas of the resulting chart indicate times when the "experimental event" had a greater positive effect on the target measure, or, conversely, that the "control event" had a greater negative effect on the target measure.

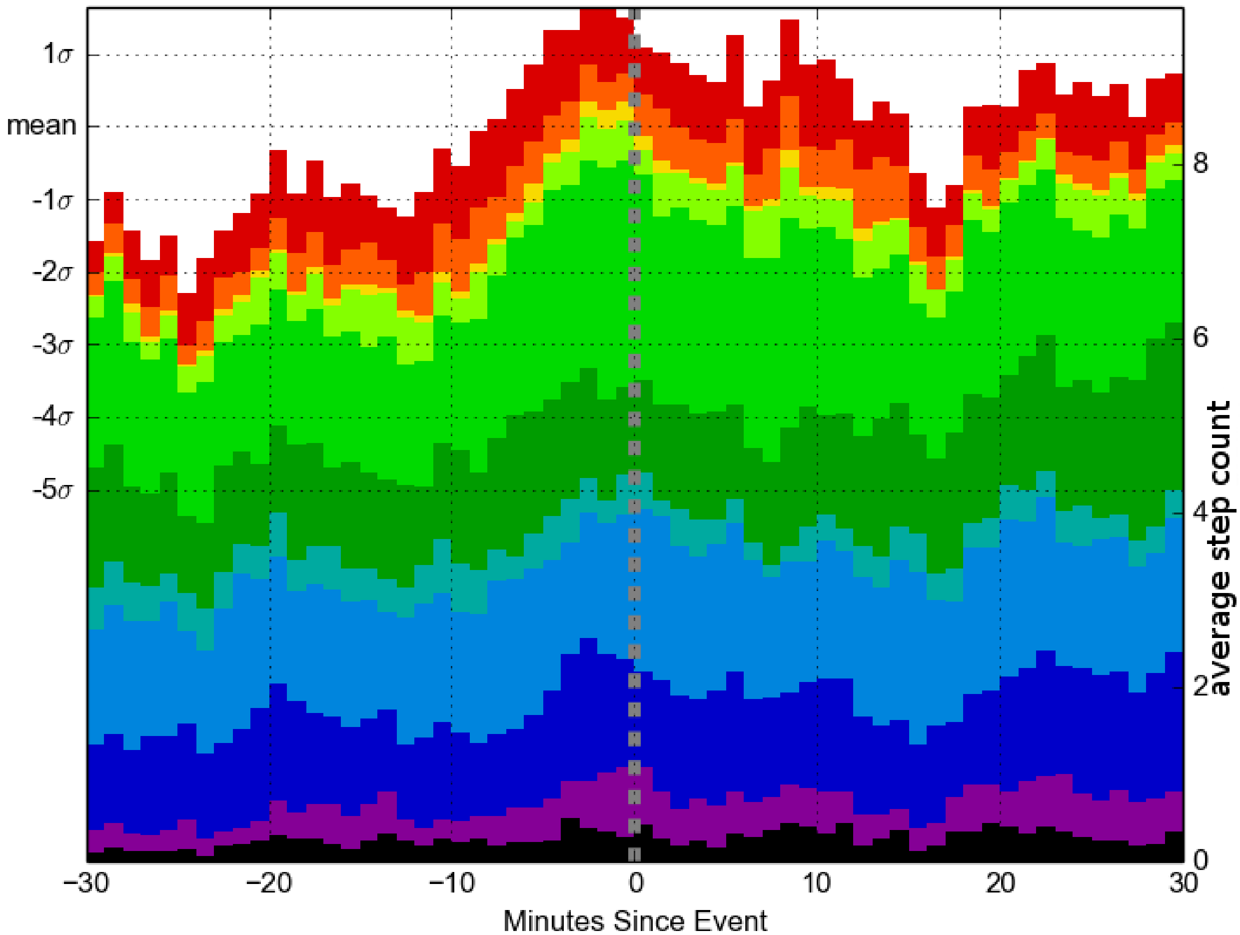

Figure 21: Stackplot of step count aggregates from the mAvatar dataset.

Shows 30 minutes surrounding 1673 phone-view events

(individual event segmentation removed due to large number of events).

The mAvatar dataset contains two types of intervention which may be interesting to compare: 1) active-avatar viewing, 2) sedentary-avatar viewing. In this case, the two event types are theoretically opposite in effect, meaning that the sedentary-avatar effect should resemble a mirrored version of the active-avatar effect. Thus, the difference should accentuate the intervention's effect signature and better isolate the behavioral response from noisy data.

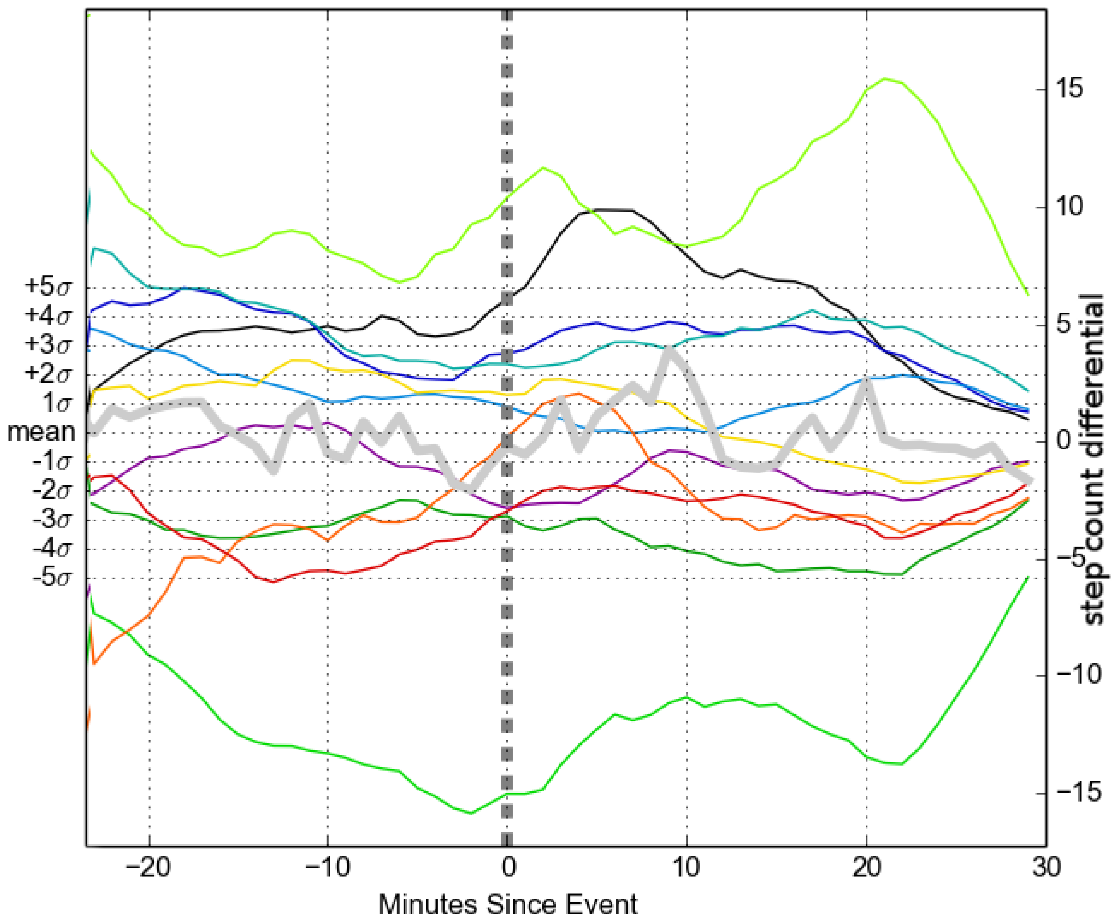

Figure 22: Active-event series average minus sedentary-event series average. Smoothed over a 15m rolling window. (average across participants shown in bold)

Even with two oppositely-polarized events, however, figure fails to show the dramatic effect a researcher might hope for. In this case, study investigators attribute the apparent lack of effect to an ambiguity in study design which led to the two opposing conditions mentioned previously, and figure 21 may indeed suggest this subgrouping within the data in the individual participant series.

4.4 Conclusion

The presented visualization methods reveal important insights into the intervention dynamics recorded in these datasets. Though this is an important first step towards better intervention efficacy analysis, there remain many open questions facing JiTAI developers. Better application of statistical measures to evaluate the findings shown here are needed to establish measures of significance for an observed effect response signature, and Predictive models which take into account the dynamics of intervention effect are needed to enable these statistical methods. However, the presented visualization methods provide the important “first look” at JiTAI data and the corresponding python scripts published publically at github.com/PIELab/interventionViz lay down a foundation upon which new modeling and analysis efforts can build.

CHAPTER 5: COMPUTATIONAL HUMAN BEHAVIOR MODELS[2]

This chapter outlines open challenges facing the development of JiTAIs and discusses the use of modeling as a common ground between behavioral scientists designing interventions and software engineers building applications. We propose that Computational Human Behavior Modeling (CHBM) has the potential to 1) help create better behavioral theories, 2) enable real-time ideographic intervention optimization, and 3) facilitate more robust data analysis techniques. First, a small set of definitions are presented to clarify ambiguities and mismatches in terminology between these two areas. Next, existing modeling concepts are used to formalize a modeling paradigm designed to fit the needs JiTAI development methodology. Last, potential benefits and open challenges of this modeling paradigm are highlighted through examination of the model-development methodology, run-time user modeling, and model-based data analysis.

Researchers theorize that an intervention which can be tailored based on the user and context may be an elegant solution to empower self-management of unhealthy behaviors like substance abuse, overeating, sedentary behavior, and more [42]. These persuasive technologies aim to utilize contextual information to deliver personalized interventions at the optimal moment in time. Real-time monitoring of data to identify states of special vulnerability to poor behavioral decisions or receptivity to intervention at any given moment is possible [44], but "a major gap exists between the technological capacity to deliver JITAIs and existing health

[2] This chapter has been adapted from an article published and presented at the International Conference on Persuasive Technology. Murray, T., Hekler, E., Spruijt-Metz, D., Rivera, D. E., & Raij, A. (2016, April). Formalization of Computational Human Behavior Models for Contextual Persuasive Technology. In *International Conference on Persuasive Technology* (pp. 150-161). Springer International Publishing. Permission to reproduce here is included in Appendix A.

behavior models." [42] Proof-of-concept applications have demonstrated the ability to adapt interventions to users [45, 46] and context [47, 48], but using current methods the complexity of the behavioral model underlying a JiTAI application grows exponentially as the complexity of the intervention design increases. Current methods for conceptualizing the human system take a piece-wise, descriptive approach, examining a phenomenon in detail, but often overlooking how the model fits into the bigger picture. Ultimately these conflicts arise because the needs of a persuasive technology are very different from the needs of extant behavioral research. While the latter places emphasis on the study of the human system's intricacies, the former needs a model which provides generalized insight and specific numerical predictions. Behavioral theories traditionally focus on nomothetic and static insights that do not offer the granularity and specificity to support the full potential of JiTAIs [50]. These methods are sufficient for analysis of traits which do not change much over days or weeks, but data collection and intervention delivery timing is now available to the microsecond for physiological data and at the at the minute-level for behavioral features and psychological constructs.

The current development process for JiTAI-like persuasive technologies requires close collaboration between behavioral scientists and application developers as they struggle to code-ify the model from extant behavioral theories for each individual experiment. The models used by a programmer to describe a user and the models used by behavioral scientists to describe a participant have certain key differences which can complicate the process of JiTAI design. In this chapter we present a hybridization of the two modeling paradigms designed to emphasize the strengths of each approach. As a part of this set of interdisciplinary terms, we introduce the concept of a Computational Human Behavior Model (CHBM) to describe this new class of models which aim to satisfy the demands of persuasive technology. Following definitions, we propose that by formalizing the CHBM underlying persuasive applications, it will be possible to create better behavioral theories, enable real-time ideographic optimization, facilitate more robust data analysis, and reduce application development time. In this section we present a look at how the concept of a CHBM would be applied to address open issues holding back JiTAIs and we highlight the remaining issues which must be addressed to make Just-in-Time Adaptive Interventions a reality.

5.1 Selected Definitions

This section presents definitions and design considerations relevant to human-behavior modeling from a theory-agnostic standpoint so that different modeling paradigms can be described under a common foundation. This set of definitions draws from both the area of HCI user-modeling and the extant paradigms of human behavior modeling in behavioral science in an attempt to synthesize a pragmatic language for use in the development of persuasive technology by behavioral scientists and application developers alike.

“Treatments” are defined by M.C. Kaptein as the set of messages or feedback a user receives from a persuasive application [105]. The term treatments seems synonymous with interventions in usage, but a single treatment should be used to unambiguously represent a single instance of user-interaction, whereas a single intervention may represent a set of interactions given as a dose.

“Just-in-Time (JiT)” is a cross-disciplinary concept defined in the context of behavioral interactions by Nahum-Shani et al. as "the effective provision of timely support, operationalized by offering the type of support needed, precisely when needed, in a way that minimizes waste (i.e., defined as anything that does not benefit the person) and accommodates the real-life setting in which support is needed." [42] Thus, for an intervention to be considered Just-in-Time (JiT), it must attempt to deliver treatment immediately before or after an event associated with the target behavior. For example, a smoking cessation JiT intervention may deliver an intervention in response to increased craving. It is important to note here that the term “event” is used to represent any exact set of circumstances over any predefined length of time. The targeted event can therefore represent not only behavioral events (such as a jog, smoking a cigarette, commuting to work), but also an interval of availability, a “meaningful moment”, or any “optimal time” defined by a match between a set of observed datapoints and a set of datapoints which define the event archetype.

“Adaptive” interventions must utilize dynamic (time-varying) “information from the person (e.g., changes in psychological distress, response to an intervention, intervention adherence) [...] to make intervention decisions repeatedly in the course of the intervention (e.g., changing the type, dosage, or timing of intervention delivery).” [42] An adaptive intervention is one that responds in real-time to the changing needs of the participant by tailoring the treatment itself based on situational context or the recent behavioral history of a user. For example, a weight loss trial might attempt to remove soda from a participant's diet and then move on to the next goal if the intervention was successful. Similarly, consider the use of step goals to increase physical activity for an obese individual; the goals may start off at a realistic level (1000 steps/day) and then build up slowly as the individual's ability progresses.

“Individualization” is defined by Nahum-Shani et al. as the “use of [static] information from the individual to make decisions about when, where and how to intervene.” [42] Thus, an intervention is individualized if “relatively stable information from the person (e.g., gender, baseline severity of symptoms) is used to make intervention-related decisions (e.g., to offer intervention package A or B)” [42] For example, a stress-relief intervention regimen may utilize relaxing music treatment based on the participant's favorite songs at study initialization, or a participant's favorite color may be used as the basis for the user interface color palette.

5.2 Computational Human Behavior Models

The following specification will allow for the formal description of a CHBM, providing a standard approach to describing, designing, and visualizing human behavior models for persuasive applications. A Computational Human Behavior Model (CHBM) is defined here as a mathematical, explicit model which describes how context is transformed into a behavioral outcome through the internal state of the human system. In summary, a Computational Human Behavior Model (CHBM) should have 1) a set of context, state, and behavior variables, 2) a set of computations which define behavior variables as a function of state which is itself a function of context, 3) a logical abstraction which allows researchers to internalize the model's behavior such that they will be better able to estimate control of the human system in general, and 4) guidelines regarding the applicable population and time-scale of the CHBM. The following section details each of these CHBM components, followed by a methodology which makes use of a graph representation to create and describe a particular CHBM.

5.2.1 Characteristics of a CHBM

5.2.1.1 User Features: Context, State, Behavior

A distinguishing feature of a CHBM is the separation of the participant definition into environmental context, internal state, and behavior variables. In reality, an individual represents an inseparable component within the larger environment, but this simplification segments out the human system for definition.

Dey et al performed an extensive literature search to define an agent's context as: “any information that can be used to characterize the situation of entities (i.e., whether a person, place, or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves. Context is typically the location, identity, and state of people, groups, and computational and physical objects.” [106] In most cases, however, it is sufficient to define context as a set of selected information from the environment available for inflow into the human system, but contextual information from the environment may be summarized and represented in countless ways. In reality, consider context to be everything that is observed by the senses. Some of this information will alter the internal state of the human system, but some may not. When building a model based on theory alone, modelers should make the selection and summary of contextual constructs to be as generalizable, extensible, and reusable as possible. When utilizing a model to simulate a particular experiment, efforts to connect available data to that which is available during the experiment may be needed, and contextual information not available empirically may need to be simulated. The environmental context influences the human system, which has an internal state represented by a set of internal state variables. In reality, internal state includes all information stored in the chemical and physical arrangement of our bodies. In order to make the model tractable, the mass of information is summarized into a set of meaningful constructs. Information flowing into a CHBM comes from the environment around an individual (the context) as an inflow which is independent of the individual's state in this instant. Similarly, information flowing out of a CHBM (as behaviors) represents actions the individual is taking to impact the environment.

As an example, consider a model of user physical activity level wherein an intervention acts to suggest physical activity as measured by a pedometer. In this example context might include a signal of intervention delivery and the location of the subject. The internal state could be represented by three constructs: 1) calling to exercise - a quickly decaying call to act if the subject is motivated and able (this is the construct targeted by the intervention), 2) physical health of the subject, 3) ability to exercise (based on location and current physical health of subject). Behavior, measured as a step count, is then determined through some function of calling to exercise and ability to exercise.

5.2.1.2 Relationships Between User Features

The relationships between context, state, and behavior variables in a CHBM must be defined computationally. The functional form of these computations is not constrained in this definition, theoretically allowing for the representation of any inter-variate relationship. There are numerous benefits to keeping the functional form of these relationships simple and homogenous across variables. Last, a simple formulation is more easily understood, allowing for a straightforward interpretation and abstraction of the model behavior.

5.2.1.3 Heuristic Interpretation

Statistical models trained on data do qualify as CHBMs in that they can define the relationships between state and context, but typically do not incorporate a logical abstraction of cognition and instead treat the internal state as a “black box”. This abstraction is essential when considering the process of JiTAI design, since the search-space available to a JiTAI designer can only be approached through heuristics guided by an understanding of how the human system will generally behave under given conditions. Though mathematical equations themselves reveal the nature of the system, naming and describing the interpretation of specific constructs or coefficients which play pivotal roles in the model can aid in the process of internalizing model behavior.

5.2.1.4 Model Metadata

While a CHBM should strive to be as broadly applicable as possible, this inevitably comes at the cost of increased complexity which can make the CHBM's nomothetic abstraction(s) intractable; there is a balance to be struck between a CHBM's inclusivity and the clarity of the abstraction. For this reason, it may be important to specify the circumstances in which a given model is valid. It may be useful to craft a highly detailed model of a particular population, but the added complexity in this model may not justify its use in a more general population. This is not analogous to the issue of overfitting in machine learning, as the model can remain accurate across the population; the primary reason for limiting the number of variables or the functional complexity of relationships is to preserve the heuristic understanding of the model.

5.2.2 Creating a CHBM

A network graph is an effective abstraction to describe the relationships (represented by arrows or “edges”) between variables (represented by the graph's "nodes") in a CHBM. In this case a directed graph wherein edge arrows represent the flow of information between nodes is used. Thus, a directed graph edge from node A to node B indicates that information flows from node A into node B. This relation can be read as "A influences B", "A informs B", or similar. This choice of notation is in agreement with graphs used in information theory, communications models, and behavioral science. In contrast, some graphing paradigms (such as probabilistic graphical models and software design) prefer to use notation wherein an edge is used to represent dependency.

While the network graph shows the connectivity of a model, it fails to indicate the meaning of each connection. In the majority of existing applications, the mathematical form of the relationship is implied or else it is neglected completely. For instance, path diagrams from the behavioral sciences frequently denote causal dependence and do not specify the functional form of the causal relationship. Adding even further to the confusion is the notion that these graphs are often developed using different statistical analyses which may make other assumptions about the functional definition of inter-variate dependency. The most common analyses assess linear relationships between variables, and thus it is perhaps reasonable to assume that this is the intention of most authors.

Assuming this is the case we can return to our simplistic example in Graph 1 and interpret the implied relationship as:

$$B(t) = coeff_{ab}A(t) + const_b$$

In this formulation $coeff_{ab}$ represents the correlation coefficient which relates A to B, and $const_b$ represents a scalar constant. For nodes with multiple inflow edges, such as node B in the following graph:

$$A => B <= C => D$$

Continuing with our assumption that node interrelations act as linear sums, the resulting formulation is simply a sum of the inflows:

$$B(t) = coeff_{ab}A(t) + coeff_{cb}C(t) + const_b$$

Using this formulation, the general form of the CHBM is expressed via the network graph alone. The general solution of an CHBM does not require definition of the constants, but a simulation cannot be run until some numerical value is assumed. These constants often have theoretical significance in that they often have meaningful influence upon system behavior. Scaling-coefficients, for instance allow for relative weighting of each inflow. Similarly, the coefficients of a dynamical equation define how quickly variables react to a change "upstream".

This linear, homogeneous-graph representation is useful, but also very limited. One important feature which this formulation does not take into account is the dynamics of the relationship. For instance, the above linear model assumes that there is no delay between variables. This assumption is fine for some applications, but this is a very poor assumption for human behavior models.

Differential equations based on a fluid-flow analogy can be used to describe the relationship between variables as described by Dong et al. [107]. Using the differential formulation our equation for B in Graph 1 becomes:

$$B(t) = coeff_{ab}A(t-\theta_{ab}) - \tau_b \frac{dB}{dt} + const$$

Just as before, our general model is not expressed entirely through the graph, and an ideographic example is specified by providing table of coefficient values. Our table is now quite a bit larger, but these coefficients have meaningful definitions which relate to our theory. While this formulation offers a huge improvement over the linear formulation, we can still imagine relationships which it cannot express.

It should be noted at this point that although the linear formulation is too simple to express the dynamics of the differential formulation, the differential formulation is capable of expressing linear relationships. This is accomplished by setting coefficients of dynamical components to zero. One might think, then, that there is some general formula which could express any functional form, and that this form should be used to express the relationships between variables in all CHBM graphs. While such formulations do exist (such as Taylor or Fourier series approximations or even ANN-based relations), this usage tends to make the model difficult to understand and to simulate with. Linear and differential formulations are in such widespread use because of the relative ease with which we can understand and solve them. Additionally, the table of coefficients needed to express an idiographic case of the model quickly becomes prohibitively large, and the effect of each coefficient on the outcome is not intuitively meaningful.

Let us now consider the case where a graph-wide assumption is not made. That is, we will specify the functional form of each node individually so that each edge on the graph may be linear in form while another may be differential. This has the benefit of allowing for both complex relationships between variables as well as simplistic ones. In this way one could craft a model in which two variables are linearly related and a third is dependent on the variance of another variable (a particularly odd formulation, but one which is relevant to behavioral theory). Unfortunately, this approach also means that a table of formulations must now be included with our graph to show the meaning of each edge in the graph. Consider for example table 3 below to describe the relations in Graph 2.

Table 3: Functional form at each node.

node	formulation
B	$coeff_{ab}A(t) + coeff_{cb}C(t) + const_b$
D	$coeff_{cd}C(t-\theta_{cd}) - \tau_d \frac{dD}{dt} + const_d$

If a fixed number of functional forms is adhered to, the graph can be made to visually represent these functional forms through the use of different node icon shapes. This approach quickly begins to resemble applications which use flow-based programming. Indeed, they are quite similar in their approach, and the specification of a CHBM is quite similar to the writing of a program.

In conclusion, we propose that an CHBM should be specified using the following rules 1) use a graph-wide formula assumption if possible, else specify formulations for each node individually, 2) when choosing a formulation, consistency between nodes is most important, 3) when choosing a formulation, simplicity and clarity is second only to consistency.

5.3 Benefits of CHBM-enabled JiTAIs

This section discusses the utility of a CHBM throughout the lifecycle of a JiTAI application. Hypothetical situations are posed to highlight the potential value of CHBM use in the JiTAI development process and show open challenges through establishment of a target user group model, application design, application implementation, data analysis, model personalization, and model iteration.

5.3.1 A Priori CHBMs

Prior to development of a JiTAI, a mental model of the target user group is established. This a priori model represents the researcher's understanding of the user group, and the design of the intervention utilizes the model in order to predict user actions. This level of detail to which this model is documented varies greatly between applications, and in some cases the causal descriptive model has little grounding in existing behavioral theory [108]. Nevertheless, a vague description of expected user behaviors and interactions with the persuasive technology still represents a user model. Existing JiTAI-like applications may not have a CHBM, but they always (sometimes informally) imply a CHBM. This section highlights the benefits of defining a CHBM explicitly, rather than relying on implicit behavioral theory.

5.3.1.1 Model Building

When model-building for a JiTAI, the planned system and underlying model of human behavior becomes very complicated, and user responses may be difficult to predict through thought experiments. Without a concrete framework to describe the model, user behavior becomes oversimplified, giving an even less accurate picture of the complex human system. When a model is under-developed, the application development process will open unaddressed questions and simple assumptions will be made. For instance, delivery of a treatment may be limited to the waking hours or to the weekdays, but this will not be reflected in the described user model. The mismatch between the documented theoretical model and the actually implemented model further muddle the process of study replication and analysis.

In addition to those assumptions knowingly made by application developers, causal descriptive modeling often contains implicit assumptions which are easily overlooked. For instance, the delay between a cause and effect is frequently neglected, that is: how quickly does a participant's behavior respond to an treatment? The process of defining a more detailed a priori model itself can lead to new insights and research questions by eliminating these oversights and forcing critical thinking on the assumptions being made.

5.3.1.2 Intervention Design

When designing intervention options for a JiTAI application, researchers will consider how a treatment influences the participant in the context of the chosen user model. When using a CHBM, this means quantifying the treatment's effect on user context. For instance, consider an intervention which provides information about the health repercussions of sedentary behavior. Assuming our CHBM uses an adaptation of the Theory of Planned Behavior [109], this intervention targets \emph{behavioral belief} regarding sedentary behavior. Since behavioral belief is part of the internal state and the treatment should be defined as part of the user's context, a context variable should be included in our model to represent external influences on behavioral belief from the environment. After defining the expected effect of a single treatment, the CHBM can then be used to predict a detailed account of user response. The use of simulations such as this in the process of designing controls is well-explored in many other areas, but is nearly unheard of in behavioral science. This is in part due to the prevalence of abstract causal descriptive models and the novelty of CHBMs, but there remain several important issues highlighted below which have not yet been addressed in this space.

5.3.1.3 Benefits of CHBMs in Persuasive Design

1) By using a CHBM with dynamical equations, the dynamics of relationships between variables can be explicitly described as a part of the model.

2) The use of an explicit a priori model for intervention design helps researchers formulate testable research questions and experiment designs.

3) The additional pre-study detail removes post-study modeling assumptions that can dilute the underlying behavioral theory or invalidate study results.

4) The process of defining a CHBM itself can lead to new insights and research questions which are almost entirely unaddressed by existing theory.

5.3.1.4 Open Questions for CHBM-Empowered Persuasive Design

1) The process of defining a CHBM requires detailed knowledge of both the underlying behavioral theory and the mathematics. Relatively few researchers today possess the necessary skillset.

2) Modeling software exists for other engineering domains, but is not directly applicable to the problem of CHBM development.

3) Software for running simulations to test the function of an \emph{a priori} CHBM is non-existent.

4) Methodologies for creating an a priori CHBM are not fully established, and mappings from existing causal descriptive models may be model-dependent.

5) The definition of a treatment's effect on a user is a subjective process. That is: how is one to know what amount of behavioral belief a specific "sedentary activity fact treatment" imparts?

6) The running of a single simulation implies a generically applicable user model, but there are likely to be multiple different responses to a single treatment which may depend on on other contextual variables. In order to get a more realistic look at user responses to an treatments, many simulations with varying parameters set to match the expectations of the researchers should be run and analyzed; this would require a CHBM simulation software suite that does not yet exist.

5.3.2 CHBMs at Run-time

In this section methods in which CHBMs may be used in the persuasive technology itself are discussed. Options include model-based intervention optimization, timing, and online ideographic modeling. A crucial step in the development of a persuasive technology today is to establish a set of decision rules based on behavioral theory which codify the circumstances in which a treatment should or should not be delivered. For instance, a treatment might be delivered only during the daytime, right before a meal, only in a particular location, or in response to a behavioral event such as cigarette use. Establishing a set of decision rules for a small number of conditions is feasible for a simple intervention, but as the number of conditions increases the number of rules required increases combinatorially. Even worse, when making use of adaptive interventions this set of rules must be expanded even further to map between all possible contexts and intervention permutations. Relying on simple decision rules loosely guided by existing theory to define the optimization of intervention delivery to control a complex system inevitably leads to under-optimized interventions, over-simplified models, and weakened data. An additional problem with this approach is the use of a binary state (i.e. rule satisfied or not) to optimize delivery over a continuous time. Because of this oversimplification, rules which govern the behavior often become part of the theory underlying the application and are clumsily expressed as decision rules. In contrast, optimization of treatment delivery using a CHBM can be done algorithmically to minimize the area between the desired and observed target behavior.

Because CHBMs are computational in nature, prediction of behavior is possible given information about the user's present and future context. Furthermore, because the behaviors in computational models are quantitative, an application could search available treatment options to find one which produces the ideal amount of a target behavior. That is, given three treatment options (A, B, C) with known effect on user context, the model can be run at t+1 for each option, and the optimum result can be chosen. Methods for model predictive control are a well studied topic of control systems engineering, but many methods cannot be applied to generic formulations. Without a constrained form to guide optimization, all possible options must be explored with equal feasibility in a brute-force search. With sufficient computational power this is effective for simple problems, but this approach becomes increasingly infeasible as the number of options and the number of future steps to be considered increase. If the functional form describing variable relationships is constrained appropriately, however, mathematical optimizations methods can greatly simplify this problem. Applications of model-predictive control over intervention delivery have been explored for gestational weight gain [107], smoking cessation [52], and fibromyalgia treatment [51] by limiting the functional form of the CHBM specification to a differential equation based on a fluid-flow analogy. In this way, application creators can implement software utilizing the advanced understanding of behavioral science described by the CHBM, without direct knowledge of the underlying behavioral science.

5.3.2.1 Benefits of CHBMs for Persuasive Applications

1) Using a CHBM enables the use of optimization algorithms instead of decision rules. This change is needed to apply complex control over target behaviors.

2) CHBMs can be adapted to fit a user's needs at run-time, establishing an idiographic model of each participant from the generalized CHBM.

5.3.2.2 Open Questions for CHBM-enabled Persuasive Applications

1) Optimization of intervention delivery can be computationally expensive unless the functional form of modelling is restricted, and it is not yet clear what formulations are most appropriate for behavioral construct relationships.

5.3.3 CHBMs Post-Study

Another rising challenge for persuasive technology researchers is the increasing complexity of data analysis methods needed to handle large amounts of "in the wild" data. Techniques designed to simplify construct relationships using statistical inferences between distinct groups of measurements cannot address emerging research questions which span the full spectrum of participant demographics, situational context, and time-scale. Contemporary approaches apply data mining and machine learning techniques to fit more advanced models to study data and identify key factors, but findings revealed in these exercises can be difficult to generalize and interpret. For example, "even if empirical evidence suggests that a given factor (e.g., psychological distress) marks state of vulnerability to a specific proximal outcome (e.g., it is highly predictive of poor state coping capacity), there is often insufficient empirical evidence concerning the cut-point of this factor that can inform the selection of one intervention option over another." [42]. By using a model as the hypothesis of an experiment rather than focusing solely on a particular relationship between two variables in specific conditions, research findings can be generalized more easily to practical persuasive applications. Methods for evaluating models, rather than evaluating correlation between two variables should be increasingly focused upon in the analysis of behavioral data. While analysis of correlation between variables looks at the statistical relationship between groups of data points, the evaluation of a model involves comparing the experimental data to the predictions of the model. CHBMs can be used with contextual data to produce a time series of expected behavioral outcomes throughout the study. The simulated "theoretical data" can then be directly compared to the "observed data" to observe how the theory differs from the reality. The process of comparing theoretical predictions to empirical data can be repeated with simulations from alternate theories and a goodness-of-fit metric can be used to evaluate the hypothesis against alternatives. Additionally, unification of existing behavioral models into this common paradigm would enable better collaboration between proponents of different theories.

5.3.3.1 Benefits of CHBMs Post-Experiment

1) Analysis of experimental data can shift focus from individual construct relationships to a larger view, evaluating the model as a hypothesis.

2) Comparison between different theories can be informed by a comparison of their respective models using a goodness-of-fit metric against empirical data.

3) The use of CHBMs makes re-use of theory and therefore collaborative improvement on existing theories easier, reversing the existing paradigm which has lead to a dizzying multitude of fragmented theories and sub-theories.

5.3.3.2 Open Questions for CHBM Post-Experiment Methods

1) Methods for fitting a model to experimental data require restrictions on the functional form of the relationships between variables, and the optimum functional form is not yet obvious.

2) Methods for evaluating the goodness-of-fit between empirical and simulated data exist, but cutting-edge software for exploring the intricacies of data mismatch may be difficult to apply to this use-case.

5.4 Conclusion

In this chapter we have offered supporting terminology, the CHBM formalization, and a set of open challenges to promote the interdisciplinary discussion needed to push forward the emerging field of JiTAI engineering. The progression of behavioral science towards computational modeling has progressed more slowly than in other scientific domains because of the limited amount of detailed, time-intensive contextual and behavioral measures available. This progression from causal descriptive modeling to causal explanatory modeling and increased mathematical rigor is a natural progression which parallels historical trends in the natural sciences. Now that behavioral and contextual data is becoming accessible, we should expect to see a similar paradigm shift in the behavioral sciences. It is our hope that this formative work towards Computational Human Behavior Modeling and the methods highlighted here act as a jumping-off point for others on the forefront of this impending paradigm shift who can use these methods to unlock the power of context-aware persuasive application driven by CHBMs.

CHAPTER 6: DESIGNING SOFTWARE TO AID DEVELOPMENT OF CHBM[3]

Computational Human Behavior Models (CHBMs) provide a mathematical model "which describes how context is transformed into a behavioral outcome through the internal state of the human system" [91]. CHBMs are a robust method of defining Just-in-Time Adaptive Intervention (JiTAI) behavior, but as the level of intervention tailoring increases, methods of modeling the relationships between sensor/EMA [49] data, user behavior, and application behavior will become increasingly important. The modeling methods of CHBMs are unfamiliar to behavioral scientists, and this remains a significant roadblock for the advancement of JiTAI systems. This chapter attempts to address this roadblock through the creation of methods and software which help behavioral scientists use CHBMs in their research. A summary of our iterative methodology for creating the BehaviorSim Model Builder is presented. BehaviorSim acts as a behavioral-scientist-facing software for development of computational models of human behavior for use in JiTAIs. We present insights gained at each stage of the development process, followed by a discussion section which formulates generalizable knowledge from our specific lessons learned that may be of use to others designing JiTAI development support software, or those targeting behavioral scientists as a user group.

[3] This chapter has been adapted from an article published and presented at the International Conference on Applied Human Factors and Ergonomics. Murray, T., Hekler, E., Spruijt-Metz, D., Rivera, D. E., & Raij, A. (2017). Lessons Learned in Development of a Behavior Modeling Tool for Health Intervention Design: BehaviorSim. In Advances in Applied Digital Human Modeling and Simulation (pp. 279-290). Springer International Publishing. Permission to reproduce here is included in Appendix A.

6.1 Methodology

6.1.1 Survey of Behavioral Scientists

A preliminary survey was given to a group of behavioral scientists in order to gauge the general perceptions and opinions on the development of behavioral models to support JiTAIs. In this survey we focused on a few key elements of the model building process to greatly simplify and shorten the modeling exercise. Contextual and behavioral outcomes based on physical activity were given, and user efforts were focused on defining the inner workings of the human system within these constraints. Participants were asked to describe the human system by sketching a time-series to represent their expectations, listing relevant constructs, and describing their constructs as they related to outcomes. Participants were also asked to complete survey items about the barriers facing modeling and simulation in behavioral science.

Approximately 50 surveys were distributed following presentations on behavioral modeling and simulation at the 35th Annual Conference of the Society of Behavioral Medicine. Out of these 50, 12 surveys were returned. In general, users had trouble with even the simplified modeling exercise. We also believe the low response rate to be indicative of the difficulty of the questionnaire, as it seemed as though all 50 participants who initially accepted the survey did attempt to complete it, but were unsatisfied with their answers and did not submit their responses. Of those few submitted, most did not stray far from the given example, and others provided very different solutions which (although helpful for conveying an abstract description of their model) could not be reconciled with the modeling paradigm presented. That is, the solutions provided abstract descriptions of the model, but they did not convey enough detail to form a CHBM. Participants seemed to find the sketching of time-series particularly challenging, and in the survey questions participants reported that the mathematics and programming concepts required for developing simulatable models were overwhelming. However, nearly all participants expressed a desire for increased collaboration between disciplines and a need for software tools to help them apply and validate these methods. These findings confirmed the need for modeling software tools to bridge the gap between systems theory and behavioral scientists.

6.1.2 BehaviorSim Model Builder v1

Using findings from the user study we developed proof-of-concept software to aid behavioral researchers with the task of building a computational behavioral model. The software, called the behaviorSim Model-Builder, took a step-wise approach towards the model-building process.

First, users are asked to list environmental inflows, internal state variables, and behavioral outflows of the model explicitly during the "think" stage. The “think” stage allows users to list the sensor measures as “context” or “behavioral”. “Context” is a measurement of the environment (e.g. location), and “behavioral measure” is a measure of a participant‘s conscious or unconscious actions. “Contracts” are variables used to represent everything in between context and behavior which are not directly measured.

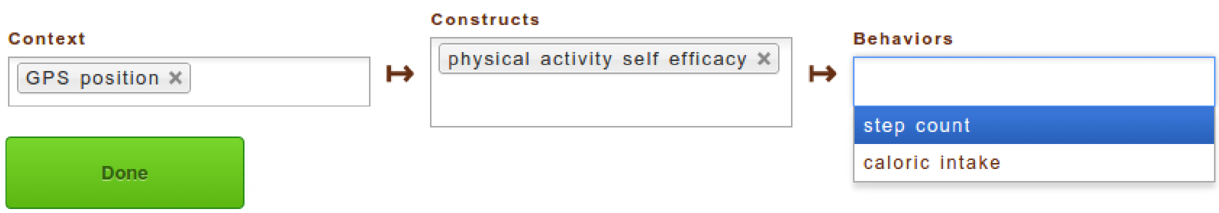

Figure 23: BehaviorSim v1 “think” user interface. Users list contextual variables, internal state variables (constructs), and behavioral measures to be used in their model.

Next, users are prompted to define the connections between nodes, "drawing" the model's structure. This is accomplished by specifying the connections using a simple Diagram

Specification Language (DSL) to denote connections between the context, state, and behavior variables given in the previous step.

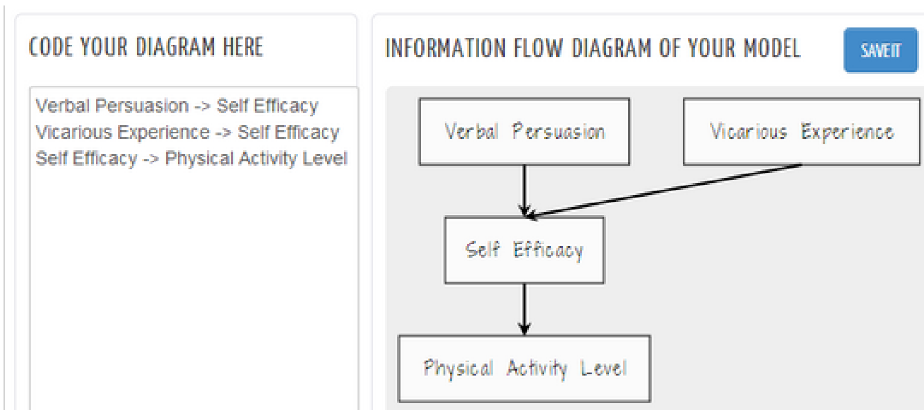

Figure 24: BehaviorSim v1 “draw” section. Shows DSL input box and information-flow graph.

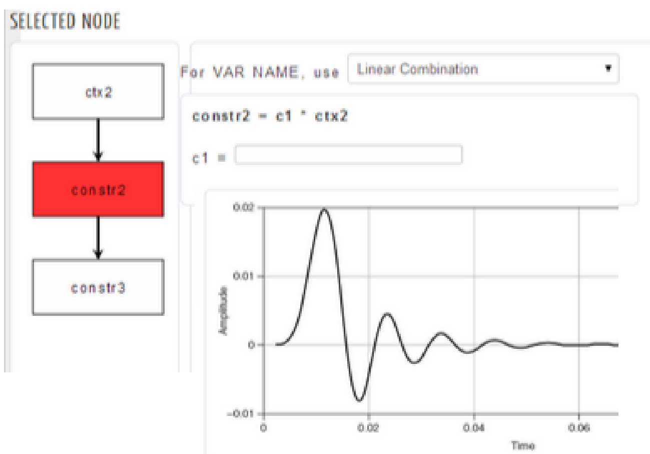

Figure 25: BehaviorSim v1 “specify” user interface. Showing node as a function its inflows.

Finally, users are required to "specify" the functional relationships at each node's inflow(s). Nodes are highlighted one-at-a-time and the relevant section of the graph including only the node in question along with its direct inflows and outflows is shown. Users are asked to select a functional form which should be used to compute the highlighted node from its inflows. Users are also asked to specify a specific set of constants to use in a test instance of the model.

These constants are used to compute a time-series representing the signal generated for this simulation instance.

As a very simple example, consider the following model of physical activity (PA): firstly we can name social pressure (SP) as an environmental inflow, normative belief (NB) as an internal state variable, and step count (SC) as a behavioral outflow; next we specify connections SP -> NB -> SC; lastly, we can specify that the connections (SP->NB and NB->SC) both represent simple linear relationships.

The model builder was reviewed by an expert panel of 2 behavioral scientists and 1 human-computer interaction expert. Though the steps in the outlined model development process seemed appropriate, it quickly became obvious that a step-wise design is not optimal. Users who are forced to explore the process step-by-step have difficulty understanding how earlier choices related to later results, and feel constrained by previous choices rather than backtracking to revise the model. This design does not allow for quick iteration on models, and requires the user to maintain a great deal of planning information internally. Though the information flow diagram (figure 24) employed in this version worked well to convey information about the model to the users, the graph was also assumed to be interactive and reviewers made attempts to modify the graph by clicking. Similarly, reviewers attempted to select nodes in the specification stage (figure 25) by clicking on them. Our review concluded that a less constrained approach to the stages of the modeling process was needed, and a greater focus on the graphical model could greatly improve user experience. Furthermore, reviewers felt that the rift between behavioral scientists and the modeling methods had not been adequately addressed; more was needed to communicate the treatment of context, constructs, and behavioral measures as time-series in the “specify” stage.

6.1.3 BehaviorSim Model-Building Tutorial

After reviewing v1 of the BehaviorSim Model Builder tool, a tutorial was designed to help bridge the knowledge gap for new modelers looking to use the tool. In theory, the tutorial would help users see the bigger picture before diving into the stepwise process. The tutorial was implemented as a walk-through of a simple example model's internals. The tutorial introduced a hybridized information-flow and time-series graph, wherein each node of the graph contains a time-series spanning a common time-frame. A user interface for adjusting model parameters and updating time-series values instantaneously was also overlaid onto this hybrid graph (see figure 26). This real-time parameter tweaking enables some degree of reconciliation with expectations of the data.

The same expert panel review process was used for the evaluation of the tutorial. Through this evaluation it became clear that, although we had taken a step in the right direction, an even more explicit definition of terms was needed in order to clarify persistent disciplinary differences. Reviewers also wanted better explanation of model input parameters and of the functional definition of the system. This tutorial included a specific scenario encoded as a set of time-series which defined the environment over time. Reviewers were not content with the hard-coded environmental inputs and wanted to be able to define how the contextual inflows changed over time. Though the hybrid graph was found helpful in conveying the connection between path diagram nodes and time series, the shared time-axis was not obvious, and reviewers expressed a need for more explicit x and y axes as well as a better explanation as to what "10 units of self efficacy" actually means. The time-series view was found to be both critical for the development of an accurate model, and valuable as a pedagogical exercise for users trying to internalize model formulations.

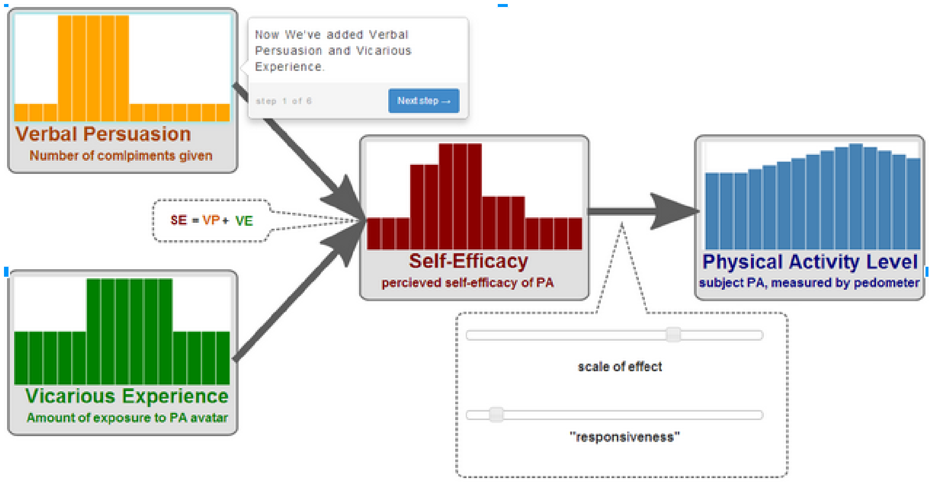

Figure 26: BehaviorSim tutorial merging time-series and information-flow graph.

6.1.4 BehaviorSim Model Builder v2

Using what we had learned so far, the BehaviorSim Model Building tool was re-designed and re-assessed. In this version all steps of the modeling process (think, draw, specify) are unified into a single-page application (see figure 27), allowing users to see how choices influence the model in real-time. This design allows users to iterate on their design more easily. The time-series charts popular in the tutorial were added as a "mini-simulation" to help users to visualize how variables change over time according to their model formulation. To address the terminology gap which plagued v1, a set of tool-tips were added which revealed detailed definitions for key terms used in the user interface. In addition, the second version incorporates findings from the v1 tutorial, adding a "miniature simulation" to the application to allow for "reconciliation" with model expectations.

In this version of the tool, users declare constructs and define the structure of their model simultaneously. In contrast to version 1, where the construct type had to be input by the user, the type of each node (contextual input, internal state, or behavioral measure) is inferred from the number of inflows and outflows. Source nodes are assumed to be contextual input, sink nodes assumed to be behavioral measures, and all others are treated as internal state nodes.

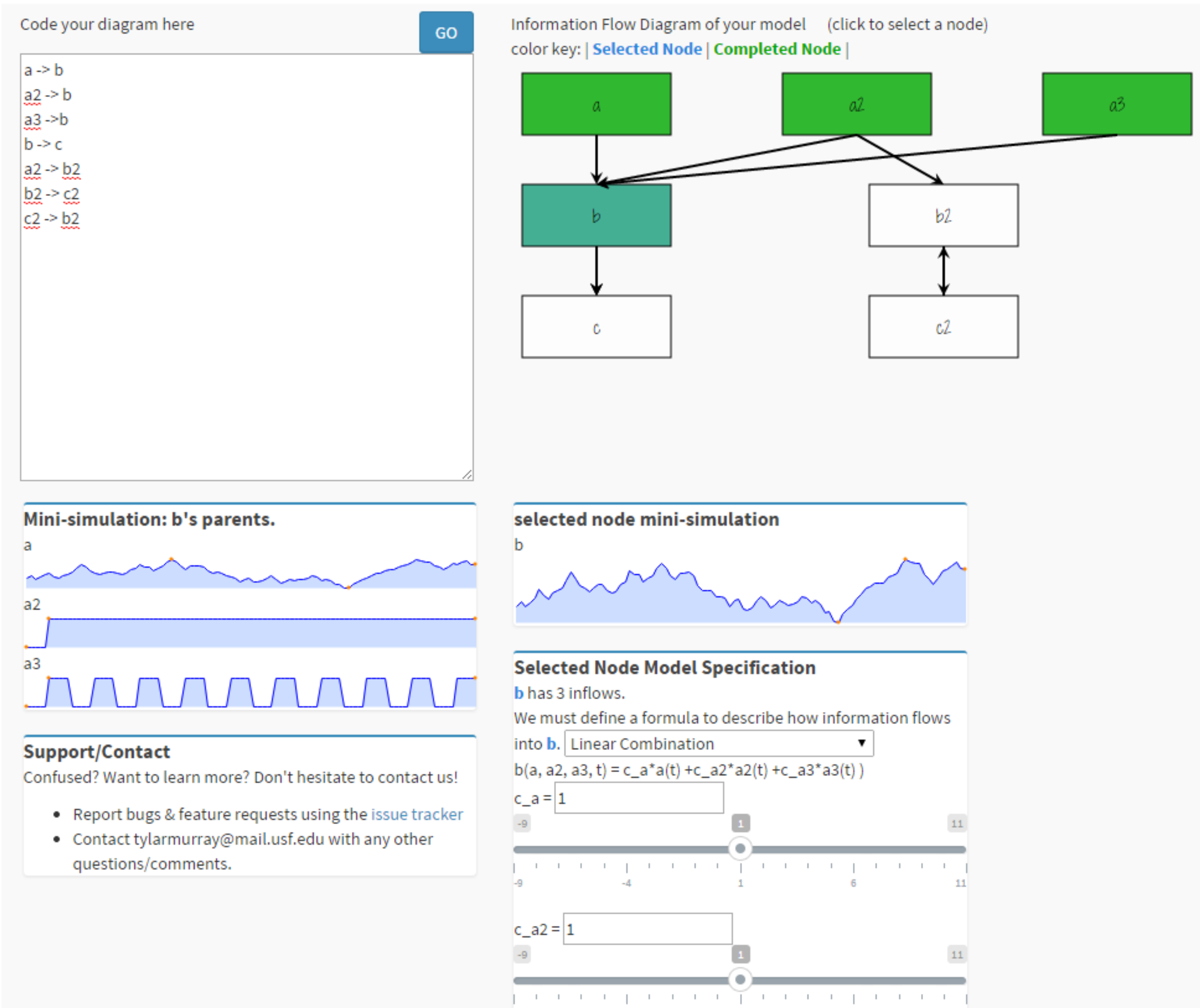

Figure 27: BehaviorSim Model Builder v2 combines elements into a single view.

The "miniature simulation" concept allows users to specify hypothetical contexts in which to explore the model dynamics, without the need to specify the full model. Users can specify environmental inflows for the simulation, choosing from adjustable presets (square wave, step function, random walk, or constant value). Selecting nodes on the graph by clicking, users can traverse the graph in any order. Time-series plots of inflows as well as the resulting outflow at each node are provided using the miniature simulation model instance. Internal state and behavioral measure nodes are specified similarly to environmental inflows through customization of function presets such as "linear combination" and "fluid flow analogy" [110].

To evaluate this design, this version was used as part of a structured exercise and interview outlining the design of a JiTAI to combat obesity. Based on recent work done to define the JiTAI design use-case [42], participants were asked to walk through a series of steps including “identify the distal outcome of your JiTAI”, “list the key factors affecting your distal outcome at the hour-to-hour level”, and “what tailoring variables will you use in your JiTAI”. Intermixed with these discussions, participants worked together with the staff to express these ideas as a CHBM using the BehaviorSim Model Builder. As the exercise progressed, the staff took an increasingly passive role, ending with a fully unassisted modeling task. A think-aloud protocol was applied while the software was in use, and the concluding interview gives insight to what users find most valuable, least valuable, and most in need of improvement. Preliminary findings from this exercise completed with 4 behavioral researchers highlight both the strengths and the remaining weaknesses of our tool.

Though the single-page design of version two did seem to allow for increased ability to iteratively explore models, reviewers now found the user interface somewhat overwhelming. Upon starting the review, users often felt unsure where to start. Furthermore the connection between the information flow graph and the related interface elements below was not obvious. Reviewers did identify the relationship after some exploration, however. Additionally, the common user interface for specifying constructs regardless of node type broke down the distinction between environmental inflows, state variables, and behavioral outputs. This lead to to some confusion when specifying the various types of nodes. Further contributing to this problem, the meaning of the mini-simulation was not always clear to reviewers, though the inclusion of time-series graphs was found helpful for understanding the model functions and their parameters once explained. Nodes on the graph were made to change color when the specification process was complete, but reviews revealed that this was not a significant enough indicator of node "completeness", and the user is sometimes unsure when they should feel free to move to the next node. Inclusion of a "next node" button which appears upon node specification completion may be all that is needed to help alleviate this issue.

6.2 Discussion

Our findings above reveal specific weaknesses in our design, and through analysis of these findings we present the following design guidelines for any software made to empower JiTAI developers. Firstly we outline a rough JiTAI developer user persona based on our assessment of the general population of researchers in behavioral intervention design. Next, we propose user stories and use-case details for the task of JiTAI design and evaluation. Lastly, we provide some generic design guidelines which we have found to be particularly relevant in this design space.

6.2.1 JiTAI Developer User Persona

In general, JiTAI developers are behavioral researchers who see the powerful potential of ubiquitous computing for high-frequency data collection, automated analysis, intervention deployment, and personalization. It is important to note that the research questions of a JiTAI researcher often differ significantly from the questions a behavioral scientist might typically have. The "traditional" way of modeling for behavior change relies primarily on statistical data analysis techniques to find relationships between variables on large time-scales. In contrast, the JiTAI researcher needs to translate these relationships into a small-time-scale model which provides guidance regarding which interventions are most effective at which specific time(s).

The JiTAI developer wants to turn a patient story into a set of equations that can be handled by an automated system. However, JiTAI developers typically do not have the level of familiarity with modeling systems to define computational psychological models mathematically. This is supported by our findings wherein we encounter more difficulty than expected using time-series as a common ground. Furthermore, the psychological models commonly used are ill-specified at the (small) timescales of greatest interest, and often do not fit commonly used modeling paradigms - making mathematical definition a unique challenge for even a systems engineer.

The JiTAI developer wants to deploy and test a hypothesis by comparing model predictions to experimental data. Statistical analysis techniques typically used to assess control vs experimental group differences are much less applicable to this problem, but the JiTAI developer often has little experience applying goodness-of-fit metrics.

6.2.2 Adaptive Interface

The science of JiTAIs is young, and (as our user persona shows) behavioral scientists looking to work with JiTAIs are likely to run into many new concepts. The potential complexity of a JiTAI system, however, may benefit from the use of advanced and specialized graphs, charts, and user interfaces. Additionally, we found overlapping terminologies to be a common pain point; meaning that new users may not recognize the need to investigate the definition of a term. Thus, the needs of a novice user versus an expert user may be very different. Because of this, software to support JiTAI development needs to promote the development of expertise in both the system and the relevant concepts through steady changes to the user interface [111]. In our case, a guided walk-through of the software interface was sufficient, but we believe that a more graded approach would be more effective.

6.2.3 Enable Quick Iterations

The value of iterating on a design spans many domains, and is very applicable to the development of JiTAIs. Through our studies we have found that the development of even a simple JiTAI requires many iterations. Thus, a software to aid in JiTAI development must allow for quick and easy modification, comparison, and reversion. A comparison between the usage of our multi-staged model builder versus the single-page application showed a dramatic increase in the number of model iterations along with reported user comprehension. Iterations on the model tended to follow a moment of realization or the learning of a new concept. Thus, allowing for quick iterations allows for the user to more quickly apply newly gained expertise, yielding a better model and increased understanding.

To encourage iteration in the JiTAI development process, assessment tools available part-way through the process (like the mini-simulation time-series) allow users to test their mental model of the system against its digital representation. Allowing for more assessment points throughout JiTAI development allows users to identify problems early and iterate before the error cascades further through the process.

6.2.4 High-Level Visuals to De-internalize Models

Traditionally, psychological models of human behavior are meant to be guidelines for thinking about human behavior. When using these models, the researcher must internalize the model and think through the participants' state. With JiTAI models, internalization of the full system becomes impossible due to the rise in specificity and complexity. Thus, JiTAI development software must provide visualizations of the system to ease cognitive load on the user. Focus plus context displays [112] can be used to allow users to delve into the specifics of a portion of the model without losing the larger context. In the behaviorSim Model Builder, we focus on the specification of a single variable at a time, and highlight this variable’s context in an information-flow path diagram of the model structure. This dual-viewing-area approach works well for comparing variable details, but the use of a zoomable interface such as is employed in some flow-based programming [113] tools may help alleviate the noted disconnect between specification and overview UIs.

6.2.5 Customized Interface

Though our JiTAI developer persona yields widely applicable general user stories, it is also important to recognize the diversity of the JiTAI developer user group. JiTAIs are applicable to any area of behavior change; just a few popular proposed JiTAI applications include management of eating behaviors, physical activity, smoking cessation, drug abuse, PTSD, and stress. Within each of these many application domains are a myriad of behavioral theories - further adding to the diversity of the user group. Each of these sub-user-groups may have slightly different needs as they develop a JiTAI. Furthermore, a JiTAI development software requires a standardized behavioral model or JiTAI format, and with that comes the opportunity to enable easy sharing and searching of JiTAI designs. Thus, personalization of the software interface - to adapt the process or to offer relevant information [114] - can greatly improve user experience in this domain.

6.3 Conclusion

In the quantified self era, we can now capture detailed, high frequency, context-specific measures of human behavior. Access to such data has the potential to change personal health, if only we could make sense of the hidden insights held within these datasets. Just-in-Time Adaptive Interventions are one application which stands to benefit from these insights and which may have great impact in applied behavioral health. One approach is to apply systems-thinking to help model and understand the data. The challenge is that health professionals and scientists do not usually have the experience or tools to apply systems thinking to health challenges.

In this chapter, we describe preliminary work on understanding how HCI and user-centered iterative design can be used to transform these data into positive behavioral health outcomes. We ran several rounds of user-centered iterative design, and identified a driving user persona and design guidelines for next-generation tools for behavioral health. In particular, our qualitative analysis indicates behavioral scientists need: 1) ways to gain expertise in systems-thinking, 2) rapid iteration through multiple theoretical designs, 3) managed cognitive load when analyzing complex models using visualization of systems and their dynamics, and 4) personalization of such tools to the behavioral problem at hand, given the great diversity and complexity of human behavior. Taking these guidelines into account, we are evolving the BehaviorSim system to better enable behavioral health researchers and practitioners to leverage high frequency, context-specific measurements and design predictive, preventive, personalized and participatory health interventions.

CHAPTER 7: CONCLUSION

The mass of personal health data available for analysis is quickly growing with no end in sight, and research is struggling to keep up. Simultaneously, an increasing percentage of health care expenditure is spent to manage chronic conditions which are often better treated through improved behavioral habits. The need for innovative, patient-driven health care continues to grow with the rising cost of health care. Existing research suggests that Just-in-Time Adaptive Interventions (JiTAIs) have the potential to leverage insights encoded in the increasingly available health and behavioral data, however, this area of research remains largely uncharted. The formative work presented in these chapters represent a significant step towards bringing engineering methodology to human behavior modeling and simulation.

The many benefits of Computational Human Behavior Models (CHBMs) have been enumerated and a vision of the potential utility brought to JiTAI design, implementation, and data analysis has been presented. It is made clear in this work that the development and use of Computational Human Behavior Models (CHBMs) is critical for the continued advancement of JiTAIs. CHBMs are the only known paradigm which allows for relatively concise codification of the complex application behaviors needed in order for mHealth applications to leverage contextual and historical user data and deliver optimally tailored and perfectly timed interventions. The formal definitions for terms relevant to CHBMs and the demonstration of their application can serve as a foundation upon which future research can build.

Carrying forward the motivating example of user-avatar-based interventions as previously presented, the real-world study participants of the mAvatar study could be modeled.

The study design attempts to influence users to raise or lower their ‘attitude’ of physical activity on alternating days and observes the resulting physical activity levels. This could be modeled using a square wave input as the user’s environment is changed by the intervention. Comparison of the physical activity level measured in real-world results with the predicted physical activity levels from the simulation could provide validation of simulation methods. Furthermore, the simulation could be improved through comparative analysis similar to that presented in the InterventionViz chapter. Insights regarding the dynamics of the intervention effect can be used to adapt the nomothetic model and systems which adjust a personalized model based on incoming data may be explored. Additionally, software may be developed to ease the transition to dynamical modeling methods and empower behavioral scientists to design JiTAIs. Guidelines presented in this work serve to inform the design of this software, and help to characterize the user base and their needs. These next steps are likely to have profound impact on personal health management and the field of behavioral science. Furthermore, these advancements depend on foundational works such as presented here.

REFERENCES

[1] Keeney RL. Personal decisions are the leading cause of death, in Operations Research. 2008. p. 1335 + .

[2] Schroeder, S. A. (2007). We can do better—improving the health of the American people. New England Journal of Medicine, 357(12), 1221-1228.

[3] Wang, Y., Beydoun, M. A., Liang, L., Caballero, B., & Kumanyika, S. K. (2008). Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic. Obesity, 16(10), 2323-2330.

[4] Waters E, et al. Interventions for preventing obesity in children. Cochrane Database of Systematic Reviews. 2011 [PubMed]

[5] Shaw KA, et al. Exercise for overweight or obesity. Cochrane Database of Systematic Reviews. 2006 [PubMed]

[6] Stead Lindsay F, Lancaster T. Group behaviour therapy programmes for smoking cessation. Cochrane Database of Systematic Reviews. 2005 [PubMed]

[7] Niederdeppe J, Farrelly MC, HavilandML. Confirming "Truth": more evidence of a successful tobacco countermarketing campaign in Florida. Journal Information. 2004; 94(2). [PMC free article] [PubMed]

[8] Franco, O. H., de Laet, C., Peeters, A., Jonker, J., Mackenbach, J., and Nusselder, W. Effects of physical activity on life expectancy with cardiovascular disease. Archives of internal medicine 165, 20 (2005), 2355–2360.

[9] Dunn, A. L., Trivedi, M. H., and O'Neal, H. A. Physical activity dose–response effects on outcomes of depression and anxiety. Medicine & Science in Sports & Exercise (2001).

[10] Yanbaeva, D. G., Dentener, M. A., Creutzberg, E. C., Wesseling, G., and Wouters, E. F. Systemic effects of smoking. Chest Journal 131, 5 (2007), 1557–1566.

[11] Ross, R., Dagnone, D., Jones, P. J., Smith, H., Paddags, A., Hudson, R., and Janssen, I. Reduction in obesity and related comorbid conditions after diet-induced weight loss or exercise-induced weight loss in mena. randomized, controlled trial. Annals of internal medicine. 133, 2 (2000), 92–103.

[12] Cooper, H., Robinson, J. C., and Patall, E. A. Does homework improve academic achievement? a synthesis of research, 1987–2003. Review of educational research 76, 1 (2006), 1–62.

[13] Norvilitis, J. M., Szablicki, P. B., and Wilson, S. D. Factors influencing levels of credit-card debt in college students. Journal of Applied Social Psychology 33, 5 (2003), 935–947.

[14] Ekelund, U., Gorely, T., Griffiths, M., Jago, R., Oppert, J. M., Raats, M., ... & Vicente-Rodríguez, G. Sedentary Behaviour and Obesity: Review of the Current Scientific Evidence.

[15] Popkin, B. M., Adair, L. S., & Ng, S. W. (2012). Global nutrition transition and the pandemic of obesity in developing countries. Nutrition reviews, 70(1), 3-21.

[16] Gurman, T. A., Rubin, S. E., & Roess, A. A. (2012). Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. Journal of health communication, 17(sup1), 82-104.

[17] Lau, P. W., Lau, E. Y., Wong, D. P., & Ransdell, L. (2011). A systematic review of information and communication technology–based interventions for promoting physical activity behavior change in children and adolescents. Journal of medical Internet research, 13(3), e48.

[18] Hou, S. I., Charlery, S. A. R., & Roberson, K. (2014). Systematic literature review of Internet interventions across health behaviors. Health Psychology and Behavioral Medicine: an Open Access Journal, 2(1), 455-481.

[19] Bravata, D. M., Smith-Spangler, C., Sundaram, V., Gienger, A. L., Lin, N., Lewis, R., ... & Sirard, J. R. (2007). Using pedometers to increase physical activity and improve health: a systematic review. Jama, 298(19), 2296-2304.

[20] Newton KH et al. Pedometers and text messaging to increase physical activity: adolescents with type 1 diabetes. Diabetes care. 32(5): 813-5. 2009

[21] Adams, M. A., Sallis, J. F., Norman, G. J., Hovell, M. F., Hekler, E. B., & Perata, E. (2013). An adaptive physical activity intervention for overweight adults: a randomized controlled trial. PloS one, 8(12), e82901.

[22] A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Franklin VL, Waller A, Pagliari C, Greene SA. Diabet Med. 2006 Dec; 23(12):1332-8.

[23] "Sweet Talk": text messaging support for intensive insulin therapy for young people with diabetes. Franklin V, Waller A, Pagliari C, Greene S Diabetes Technol Ther. 2003; 5(6):991-6.

[24] Mobile phone text messaging in the management of diabetes. Ferrer-Roca O, Cárdenas A, Diaz-Cardama A, Pulido P J Telemed Telecare. 2004; 10(5):282-5.

[25] Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin RB, Jones M. Tob Control. 2005 Aug; 14(4):255-61.

[26] Shapiro JR et al. Use of text messaging for monitoring sugar-sweetened beverages, physical activity, and screen time in children: a pilot study. Journal of Nutrition Education and Behavior. 40(6): 385-91. (2008)

[27] Bauer S, Percevic R, Okon E, Meermann R, Kordy H. Use of text messaging in the aftercare of patients with bulimia nervosa. European Eating Disorders Review. 2003;11:279–290.

[28] Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): response rate analysis and focus group evaluation from a pilot study. Anhøj J, Møldrup C. J Med Internet Res. 2004 Dec 2; 6(4):e42.

[29] Chang, K.-h., Fisher, D., Canny, J., and Hartmann, B. How's my mood and stress?: an efficient speech analysis library for unobtrusive monitoring on mobile phones. In Proceedings of the 6th International Conference on Body Area Networks, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2011), 71–77.

[30] Lu, H., Frauendorfer, D., Rabbi, M., Mast, M. S., Chittaranjan, G. T., Campbell, A. T., Gatica-Perez, D., and Choudhury, T. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, ACM (2012), 351–360.

[31] Li, M., Rozgic, V., Thatte, G., Lee, S., Emken, B., Annavaram, M., Mitra, U., Spruijt-Metz, D., and Narayanan, S. Multimodal physical activity recognition by fusing temporal and cepstral information. Neural Systems and Rehabilitation Engineering, IEEE Transactions on 18, 4 (2010), 369–380.

[32] Emken, B. A., Li, M., Thatte, G., Lee, S., Annavaram, M., Mitra, U., Narayanan, S., and Spruijt-Metz, D. Recognition of physical activities in overweight hispanic youth using knowme networks. Journal of physical activity & health 9, 3 (2012), 432.

[33] Wyatt, D., Choudhury, T., Bilmes, J., and Kitts, J. A. 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 (2011), 7.

[34] Sazonov, E., Metcalfe, K., Lopez-Meyer, P., and Tiffany, S. Rf hand gesture sensor for monitoring of cigarette smoking. In Sensing Technology (ICST), 2011 Fifth International Conference on, IEEE (2011), 426–430.

[35] Klasnja, P., and Pratt, W. Healthcare in the pocket: Mapping the space of mobile-phone health interventions. Journal of biomedical informatics 45, 1 (2012), 184–198.

[36] Nahum-Shani, I., Qian, M., Almirall, D., Pelham, W. E., Gnagy, B., Fabiano, G. A., Waxmonsky, J. G., Yu, J., and Murphy, S. A. Q-learning: A data analysis method for constructing adaptive interventions. Psychological Methods 17, 4 (2012), 478.

[37] Ho, J., and Intille, S. S. Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM (2005), 909–918.

[38] Sohn, T., Li, K. A., Lee, G., Smith, I., Scott, J., and Griswold, W. G. Place-its: A study of location-based reminders on mobile phones. Springer, 2005.

[39] Jewell, S. Productivity via mobile phones: Using smartphones in smart ways. Journal of Electronic Resources in Medical Libraries 8, 1 (2011), 81–86.

[40] Gallego Vico, D., Huecas, G., and Salvachua Rodriguez, J. Generating context-aware recommendations using banking data in a mobile recommender system. In ICDS 2012, The Sixth International Conference on Digital Society (2012), 73–78.

[41] Elliott, J. J. Development of an energy-information feedback system for a smartphone application.

[42] Nahum-Shani, I., Smith, S. N., Tewari, A., Witkiewitz, K., Collins, L. M., Spring, B., & Murphy, S. (2014). Just in time adaptive interventions (jitais): An organizing framework for ongoing health behavior support. Methodology Center technical report, (14-126).

[43] Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. Health Psychology, 34(S), 1209.

[44] Hekler, E. B., Klasnja, P., Traver, V., & Hendriks, M. (2013). Realizing effective behavioral management of health: the metamorphosis of behavioral science methods. IEEE pulse, 4(5), 29-34.

[45] Dallery, J., & Raiff, B. R. (2014). Optimizing behavioral health interventions with single-case designs: from development to dissemination. Translational behavioral medicine, 4(3), 290-303.

[46] Beck, C., McSweeney, J. C., Richards, K. C., Roberson, P. K., Tsai, P. F., & Souder, E. (2010). Challenges in tailored intervention research. Nursing outlook, 58(2), 104-110.

[47] Brailsford, S. C., Desai, S. M., & Viana, J. (2010, December). Towards the holy grail: combining system dynamics and discrete-event simulation in healthcare. In Simulation Conference (WSC), Proceedings of the 2010 Winter (pp. 2293-2303). IEEE.

[48] Collins, L. M., Murphy, S. A., & Bierman, K. L. (2004). A conceptual framework for adaptive preventive interventions. Prevention science, 5(3), 185-196.

[49] Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. Annu. Rev. Clin. Psychol., 4, 1-32.

[50] Riley, W. T., Rivera, D. E., Atienza, A. A., Nilsen, W., Allison, S. M., & Mermelstein, R. (2011). Health behavior models in the age of mobile interventions: are our theories up to the task?. Translational behavioral medicine, 1(1), 53-71.

[51] Deshpande, S., Nandola, N. N., Rivera, D. E., and Younger, J. W. : Optimized treatment of fibromyalgia using system identification and hybrid model predictive control. Control engineering practice, 33, 161-173. (2014)

[52] Timms, K.P., D.E. Rivera, M.E. Piper, and L.M. Collins, "A hybrid model predictive control strategy for optimizing a smoking cessation intervention," Proceedings of the 2014 American Control Conference, Portland, Oregon, June 4-6, 2014.

[53] Pina, A.A., L.E. Holly, A.A. Zerr, and D.E. Rivera, "A personalized and control systems engineering conceptual approach to target childhood anxiety in the contexts of cultural diversity," Journal of Clinical Child and Adolescent Psychology, Volume 43, Issue 3, pages 442-453, 2014.

[54] Dong, Y., D.E. Rivera, D.S. Downs, J.S. Savage, D.M. Thomas, and L.M. Collins, "Hybrid model predictive control for optimizing gestational weight behavioral interventions," Proceedings of the 2013 American Control Conference, June 17-19, 2013, Washington, D.C.

[55] Picard, Rosalind Wright. ”Affective computing.” (1995).

[56] Mehrabian, A. Silent messages: Implicit communication of emotions and attitudes. Belmont, CA: Wadsworth. (1981)

[57] Fox, J. et al. Virtual Self-Modeling: The Effects of Vicarious Reinforcement and Identification on Exercise. Behaviors. Media Psychology, 12(1), 125. (2009)

[58] Hershfield HE et al. Increasing Saving Behavior through Age-Progressed Renderings of the Future Self. Journal of Marketing Research. 48(SPL):23-37. (2011)

[59] Yee N et al. Walk a Mile in Digital Shoes: The Impact of Embodied Perspective-Taking on the Reduction of Negative Stereotyping in Immersive Virtual Environments. Proceedings of PRESENCE. 24-26. (2006)

[60] Ahn, SJ.: Embodied Experiences in Immersive Virtual Environments: Effects on Pro-Environmental Attitude and Behavior. Stanford University. (2011)

[62] Turk M et al. Multimodal Transformed Social Interaction. International Conference on Multimodal Interfaces: ACM; 46-52. (2004)

[63] Bandura, A.: Social learning theory. Englewood Cliffs, N.J.: Prentice Hall. (1977)

[64] Behm-Morawitz, Elizabeth: ”Mirrored selves: The influence of self-presence in a virtual world on health, appearance, and well-being.” Computers in Human Behavior (2012)

[65] Belman, Jonathan et al ”Designing Games to Foster Empathy.” International Journal of Cognitive Technology 14.2 : 11. (2009)

[66] Hekler EB et al. A case study of BSUED: Behavior science-informed user experience design. Personal Informatics & HCI: Design, Theory & Social Implications, at the ACM SIGCHI Conference on Human Factors Systems. Vancouver, BC, Canada. p. http://personalinformatics.org/chi2011/hekler (2011)

[67] King AC et al. Mobile phone applications to promote physical activity increases: Preliminary results of the MILES pilot study. Annual conference for the Society of Behavioral Medicine. New Orleans, LA, USA; (2012)

[68] Lin J et al. Fish'n'steps: Encouraging Physical Activity with an Interactive Computer Game. International Conference on Ubiquitous Computing (UbiComp): ACM; 261-278. (2006)

[69] Consolvo, Sunny et al: Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. Proceedings of the 10th international conference on Ubiquitous computing. ACM. (2008)

[70] Kelleher, Christa et al. "Ten guidelines for effective data visualization in scientific publications." Environmental Modelling & Software26.6 822-827. (2011)

[71] Murray, T., Jaimes, L., Hekler, E., Spruijt-Metz, D., & Raij, A. (2013, November). A glanceable mobile avatar for behavior change. In Proceedings of the 4th Conference on Wireless Health (p. 16). ACM.

[72] Revett, Kenneth. Behavioral biometrics: a remote access approach. Wiley Publishing. (2008)

[73] Yee N et al. The Proteus Effect: Implications of Transformed Digital Self-Representation on Online and Offline Behavior. Communication Research. 36(2):285-312. (2009)

[74] Cunningham, Michael R.: Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. Journal of Personality and Social Psychology, Vol 50(5), 925-935. (May 1986)

[75] Hekler, E.B. et al. Mind the Theoretical Gap: Interpreting, Using and Developing Behavioral Theory in HCI. Research. ACM-CHI Conference. Paris, France. (May 2013)

[76] Fogg, B. J.: "A behavior model for persuasive design." Proceedings of the 4th international Conference on Persuasive Technology. ACM. (2009)

[77] Fogg, B. et al. "Behavior wizard: A method for matching target behaviors with solutions." Persuasive Technology : 117-131. (2010)

[78] Hekler, E.B. et al. Exploring Behavioral Markers of Long-term Physical Activity Maintenance: A Case Study of System Identification Modeling within a Behavioral Intervention. Submitted for future publication in Health Education and Research. (2012)

[79] Vasalou, A., Joinson, A., Banziger, T., Goldie, P., and Pitt, J. Avatars in social media: Balancing accuracy, playfulness and embodied messages. International Journal of Human-Computer Studies 66, 11 (2008), 801–811.

[80] Kumar, S., Chhugani, J., Kim, C., Kim, D., Nguyen, A. D., Dubey, P., Bienia, C., and Kim, Y. Second life and the new generation of virtual worlds. IEEE Computer 41, 9 (2008), 46–53.

[81] Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., and Merget, D. The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. CyberPsychology & Behavior 10, 1 (2007), 115–121

[82] Benford, S., Bowers, J., Fahlen, L. E., Greenhalgh, C., and Snowdon, D. User embodiment in collaborative virtual environments. In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press/Addison-Wesley Publishing Co. (1995), 242–249

[83] Consolvo, S., Klasnja, P., McDonald, D. W., Avrahami, D., Froehlich, J., LeGrand, L., Libby, R., Mosher, K., and Landay, J. A. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In Proceedings of the 10th international conference on Ubiquitous computing, ACM (2008), 54–63.

[84] Lane, N. D., Mohammod, M., Lin, M., Yang, X., Lu, H., Ali, S., Doryab, A., Berke, E., Choudhury, T., and Campbell, A. Bewell: A smartphone application to monitor, model and promote wellbeing. In 5th International ICST Conference on Pervasive Computing Technologies for Healthcare (2011), 23–26.

[85] Harries, T., Eslambolchilar, P., Stride, C., Rettie, R., and Walton, S. Walking in the wild–using an always-on smartphone application to increase physical activity. In Human-Computer Interaction–INTERACT 2013. Springer, 2013, 19–36.

[86] Bauer, J., Consolvo, S., Greenstein, B., Schooler, J., Wu, E., Watson, N. F., and Kientz, J. Shuteye: encouraging awareness of healthy sleep recommendations with a mobile, peripheral display. In Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, ACM (2012), 1401–1410.

[87] Klasnja, P., and Pratt, W. Healthcare in the pocket: Mapping the space of mobile-phone health interventions. Journal of biomedical informatics 45, 1 (2012), 184–198.

[88] Hekler, E., Buman, M., Haskell, W., Rosenberger, M., and King, A. Validity of android-based mobile phones as assessment devices for physical activity. 2010 health summit. Washington, DC (2010).

[89] Inc., F. Fitbit One activity and sleep tracker. https://www.fitbit.com/one. Accessed: 2014-03-19.

[90] LLC, S. S. L. Fitabase research platform. http://www.fitabase.com/. Accessed: 2014-03-19

[91] Murray, T., Hekler, E., Spruijt-Metz, D., Rivera, D. E., & Raij, A. (2016). Formalization of Computational Human Behavior Models for Contextual Persuasive Technology. In Persuasive Technology (pp. 150-161). Springer International Publishing.

[92] Plaisant, C., Milash, B., Rose, A., Widoff, S., & Shneiderman, B. (1996, April). LifeLines: visualizing personal histories. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 221-227). ACM.

[93] Wongsuphasawat, K., Guerra Gómez, J. A., Plaisant, C., Wang, T. D., Taieb-Maimon, M., & Shneiderman, B. (2011, May). LifeFlow: visualizing an overview of event sequences. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 1747-1756). ACM.

[94] Wongsuphasawat, K., & Gotz, D. (2012). Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization. Visualization and Computer Graphics, IEEE Transactions on, 18(12), 2659-2668.

[95] Fails, J. A., Karlson, A., Shahamat, L., & Shneiderman, B. (2006, October). A visual interface for multivariate temporal data: Finding patterns of events across multiple histories. In Visual Analytics Science And Technology, 2006 IEEE Symposium On (pp. 167-174). IEEE.

[96] Vrotsou, K., Ellegård, K., & Cooper, M. (2007, July). Everyday life discoveries: Mining and visualizing activity patterns in social science diary data. In Information Visualization, 2007. IV'07. 11th International Conference (pp. 130-138). IEEE.

[97] Wongsuphasawat, K., & Gotz, D. (2011, March). Outflow: Visualizing patient flow by symptoms and outcome. In IEEE VisWeek Workshop on Visual Analytics in Healthcare, Providence, Rhode Island, USA (pp. 25-28). American Medical Informatics Association.

[98] Van Wijk, J. J., & Van Selow, E. R. (1999). Cluster and calendar based visualization of time series data. In Information Visualization, 1999.(Info Vis' 99) Proceedings. 1999 IEEE Symposium on (pp. 4-9). IEEE.

[99] Gotz, D., & Stavropoulos, H. (2014). Decisionflow: Visual analytics for high-dimensional temporal event sequence data. Visualization and Computer Graphics, IEEE Transactions on, 20(12), 1783-1792.

[100] Spruijt-Metz, D., Wen, C. K. F., O'Reilly, G., Li, M., Lee, S., Emken, B. A., ... & Narayanan, S. (2015). Innovations in the Use of Interactive Technology to Support Weight Management. Current Obesity Reports, 4(4), 510-519.

[101] Bailenson, J. N., & Segovia, K. Y. (2010). Virtual doppelgangers: Psychological effects of avatars who ignore their owners. In Online worlds: Convergence of the real and the virtual (pp. 175-186). Springer London.

[102] Glass, G. V., Willson, V. L., & Gottman, J. M. (1975). Design and analysis of time-series experiments (Vol. 197). Boulder: Colorado Associated University Press.

[103] Tufte, E. R. (2006). Beautiful evidence. New York.

[104] Reijner, H. (2008). The development of the horizon graph.

[105] Kaptein, M. C. : Formalizing Customization in Persuasive Technologies. In Persuasive Technology (pp. 27-38). Springer International Publishing. (2015)

[106] Dey, A. K. (2001). Understanding and using context. Personal and ubiquitous computing, 5(1), 4-7.

[107] Dong, Y., D.E. Rivera, D.S. Downs, J.S. Savage, D.M. Thomas, and L.M. Collins, "Hybrid model predictive control for optimizing gestational weight behavioral interventions," Proceedings of the 2013 American Control Conference, June 17-19, 2013, Washington, D.C.

[108] Prestwich, A., Sniehotta, F. F., Whittington, C., Dombrowski, S. U., Rogers, L., and Michie, S. : Does theory influence the effectiveness of health behavior interventions? Meta-analysis. Health Psychology, 33(5), 465. (2014)

[109] Ajzen, I. The theory of planned behavior. Organizational behavior and human decision processes, 50(2), 179-211. (1991)

[110] Martin, C. A., Rivera, D. E., Riley, W. T., Hekler, E. B., Buman, M. P., Adams, M. A., King, A. C.: A Dynamical Systems Model of Social Cognitive Theory. In: American Control Conference (ACC), pp. 2407--2412). IEEE. (2014, June)

[111] Cockburn, A., Gutwin, C., Scarr, J., Malacria, S.: Supporting Novice to Expert Transitions in User Interfaces. ACM Computing Surveys (CSUR), 47(2), 31. (2015)

[112] Baudisch, P., Good, N., Stewart, P.: Focus Plus Context Screens: Combining Display Technology with Visualization Techniques. In: Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology pp. 31--40. ACM. (2001, November)

[113] Morrison, J. P. Flow-Based Programming: A New Approach to Application Development. CreateSpace. (2010)

[114] Kay, J., & Kummerfeld, B. (2012). Creating Personalized Systems that People can Scrutinize and Control: Drivers, Principles and Experience. ACM Transactions on Interactive Intelligent Systems (TiiS), 2(4), 24.

APPENDIX A: PERMISSION FOR REPRODUCTION OF MATERIALS

Below is the permissions for reproduction of material in chapter two.

Springer
the language of science

Title:	Avatar Interfaces for Biobehavioral Feedback
Author:	Tylar Murray
Publication:	Springer eBook
Publisher:	Springer
Date:	Jan 1, 2013

Logged in as:
Tylar Murray
LOGOUT

Order Completed

Thank you for your order.

This Agreement between Tylar Murray ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

License Number	3977491273486
License date	Oct 28, 2016
Licensed Content Publisher	Springer
Licensed Content Publication	Springer eBook
Licensed Content Title	Avatar Interfaces for Biobehavioral Feedback
Licensed Content Author	Tylar Murray
Licensed Content Date	Jan 1, 2013
Type of Use	Book/Textbook
Requestor type	Agency acting on behalf of other industry
Portion	Full text
Format	Electronic
Will you be translating?	No
Print run	9999
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Title of new book	Towards Computational Human Behavior Modeling for Just-in-Time Adaptive Interventions
Publisher of new book	University of South Florida
Author of new book	Tylar Murray
Expected publication date of new book	Dec 2016
Estimated size of new book (pages)	100

Below is the permissions for reproduction of material in chapter five.

Copyright Clearance Center RightsLink®

Springer the language of science

Title:	Formalization of Computational Human Behavior Models for Contextual Persuasive Technology
Author:	Tylar Murray
Publication:	Springer eBook
Publisher:	Springer
Date:	Jan 1, 2016

Logged in as:
Tylar Murray
LOGOUT

Order Completed

Thank you for your order.

This Agreement between Tylar Murray ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

License Number	3977500554884
License date	Oct 28, 2016
Licensed Content Publisher	Springer
Licensed Content Publication	Springer eBook
Licensed Content Title	Formalization of Computational Human Behavior Models for Contextual Persuasive Technology
Licensed Content Author	Tylar Murray
Licensed Content Date	Jan 1, 2016
Type of Use	Book/Textbook
Requestor type	Agency acting on behalf of other industry
Portion	Full text
Format	Electronic
Will you be translating?	No
Print run	9999999
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Title of new book	Towards Computational Human Behavior Modeling for Just-in-Time Adaptive Interventions
Publisher of new book	University of South Florida
Author of new book	Tylar Murray
Expected publication date of new book	Dec 2016
Estimated size of new book (pages)	100

Below is the permissions for reproduction of material in chapter six.

Copyright Clearance Center RightsLink®

Springer the language of science

Title: Lessons Learned in Development of a Behavior Modeling Tool for Health Intervention Design: BehaviorSim

Author: Tylar Murray

Publication: Springer eBook

Publisher: Springer

Date: Jan 1, 2017

Logged in as:
Tylar Murray
LOGOUT

Order Completed

Thank you for your order.

This Agreement between Tylar Murray ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

License Number	3977500704154
License date	Oct 28, 2016
Licensed Content Publisher	Springer
Licensed Content Publication	Springer eBook
Licensed Content Title	Lessons Learned in Development of a Behavior Modeling Tool for Health Intervention Design: BehaviorSim
Licensed Content Author	Tylar Murray
Licensed Content Date	Jan 1, 2017
Type of Use	Book/Textbook
Requestor type	Agency acting on behalf of other industry
Portion	Full text
Format	Electronic
Will you be translating?	No
Print run	9999999
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Title of new book	Towards Computational Human Behavior Modeling for Just-in-Time Adaptive Interventions
Publisher of new book	University of South Florida
Author of new book	Tylar Murray
Expected publication date of new book	Dec 2016
Estimated size of new book (pages)	100

APPENDIX B: IRB LETTERS OF APPROVAL

USF
UNIVERSITY OF SOUTH FLORIDA

DIVISION OF RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-5618

January 9, 2013

Tylar Murray, B.S.
Electrical Engineering
1712 Mulberry Drive
Tampa, FL 33604

RE: **Expedited Approval** for Initial Review
IRB#: Pro00009046
Title: Improving the Health of Adolescents with Mobile Avatars

Dear Mr. Murray:

On 1/9/2013 the Institutional Review Board (IRB) reviewed and **APPROVED** the above referenced protocol. Please note that your approval for this study will expire on 1/9/2014.

Approved Items:
Protocol Document:
Protocol

Consent/Assent Documents:
Assent Form.docx.pdf
Parental Permission Minimal Risk.docx.pdf
Please use only the official, IRB- stamped consent/assent document(s) found under the "Attachment Tab" in the recruitment of participants. Please note that these documents are only valid during the approval period indicated on the stamped document.

This study involves children; approved under 45CFR46.404: Research not involving greater than minimal risk. It was the determination of the IRB that your study qualified for expedited review which includes activities that (1) present no more than minimal risk to human subjects, and (2) involve only procedures listed in one or more of the categories outlined below. The IRB may review research through the expedited review procedure authorized by 45CFR46.110 and 21 CFR 56.110. The research proposed in this study is categorized under the following expedited review categories:

(6) Collection of data from voice, video, digital, or image recordings made for research purposes.

(7) Research on individual or group characteristics or behavior (including, but not limited to,

research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB. Any changes to the approved research must be submitted to the IRB for review and approval by an amendment.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

UNIVERSITY OF SOUTH FLORIDA

RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-7091

April 23, 2014

Tylar Murray, B.S.
Electrical Engineering
4202 E. Fowler Avenue, SVC1072
Tampa, FL 33620

RE: **Exempt Certification**
IRB#: Pro00017068
Title: Behavioral Modeling and Simulation Interactive Presentation

Study Approval Period: 4/23/2014 to 4/23/2019

Dear Mr. Murray:

On 4/23/2014, the Institutional Review Board (IRB) determined that your research meets USF requirements and Federal Exemption criteria as outlined in the federal regulations at 45CFR46.101(b):

(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

Approved Documents:

studyProtocol.pdf

InformedConsentScript.pdf

As the principal investigator for this study, it is your responsibility to ensure that this research is conducted as outlined in your application and consistent with the ethical principles outlined in the Belmont Report and with USF IRB policies and procedures. Please note that changes to this protocol may disqualify it from exempt status. Please note that you are responsible for notifying the IRB prior to implementing any changes to the currently approved protocol.

The Institutional Review Board will maintain your exemption application for a period of five years from the date of this letter or for three years after a Final Progress Report is received, whichever is longer. If you wish to continue this protocol beyond five years, you will need to submit a new application at least 60 days prior to the end of your exemption approval period. Should you complete this study prior to the end of the five-year period, you must submit a request to close the study.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

UNIVERSITY OF SOUTH FLORIDA

RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-7091

4/30/2015

Tylar Murray, MSES
USF Electrical Engineering
3931 Endicott Dr
New Port Richey, FL 34652

RE: **Exempt Certification**
IRB#: Pro00020226
Title: Usability Assessment of the behaviorSim Model Builder Tool

Dear Tylar Murray:

On 4/30/2015, the Institutional Review Board (IRB) determined that your research meets criteria for exemption from the federal regulations as outlined by 45CFR46.101(b):

(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

Approved Item(s):
Protocol Document(s):

Protocol-behaviorSimModelBuilderUsabilityStudy.pdf

Consent/Assent Document(s):

consentform.pdf

As the principal investigator for this study, it is your responsibility to ensure that this research is conducted as outlined in your application and consistent with the ethical principles outlined in the Belmont Report and with USF IRB policies and procedures.

Please note, as per USF IRB Policy 303, "Once the Exempt determination is made, the application is closed in eIRB. Any proposed or anticipated changes to the study design that was previously declared exempt from IRB review must be submitted to the IRB as a new study prior to initiation of the change."

If alterations are made to the study design that change the review category from Exempt (i.e., adding a focus group, access to identifying information, adding a vulnerable population, or an intervention), these changes require a new application. However, administrative changes, including changes in research personnel, do not warrant an amendment or new application.

Given the determination of exemption, this application is being closed in ARC. This does not limit your ability to conduct your research project. Again, your research may continue as planned; only a change in the study design that would affect the exempt determination requires a new submission to the IRB.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.