Intelligent Positive Computing with Mobile, Wearable, and IoT Devices: Literature Review and Research Directions

Uichin Lee[a,*], Kyungsik Han[b,*], Hyunsung Cho[a,1], Kyong-Mee Chung[c,1], Hwajung Hong[d,1,*], Sung-Ju Lee[a,1], Youngtae Noh[e,1,*], Sooyoung Park[a,1], John M. Carroll[f]

[a]*KAIST, Daejeon, Rep. of Korea*
[b]*Ajou University, Suwon, Rep. of Korea*
[c]*Yeonsei University, Seoul, Rep. of Korea*
[d]*Seoul National University, Seoul, Rep. of Korea*
[e]*Inha University, Incheon, Rep. of Korea*
[f]*Pennsylvania State University, Pennsylvania, USA*

Abstract

The use of mobile, wearable, and Internet of Things (IoT) technologies fosters unique opportunities for designing novel intelligent positive computing services that address various health and well-being issues such as stress and depression. As positive computing research is often cross-disciplinary, it is difficult to acquire holistic perspectives on the design, implementation, and evaluation of intelligent positive computing systems with mobile, wearable, and IoT technologies. To bridge this gap, we propose a conceptual framework and review the key components to provide guidelines for intelligent positive computing systems research. We also present several practical service scenarios and provide useful insights on opportunities and challenges. By critically reflecting on the literature and scenarios, we suggest several research directions on the core topics in intelligent positive computing systems research. In addition, we discuss concerns and challenges such as technology dependence, abandonment, side effects, privacy, and ethical issues.

*Corresponding authors: Uichin Lee (Framework & Core Technologies), Kyungsik Han (Scenarios & Research Directions), Hwajung Hong (Design & Evaluation). Email addresses: uclee@kaist.ac.kr, kyungsikhan@ajou.ac.kr, hwajunghong@snu.ac.kr

[1]These authors made equal contributions, and the authors' names were alphabetically ordered.

[2]This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017M3C4A7083529).

Keywords: Positive Computing, Persuasive Technologies, Mobile and Wearable Technologies, Internet-of-Things (IoT), Evidence-based Design and Intervention

1. Introduction

Recent advances of mobile, wearable, and Internet-of-Things (IoT) technologies have greatly changed computing paradigms. Earlier paradigms based on desktop environments were more focused on making computing services usable. Nowadays, however, the focus has shifted towards making computing services *persuasive*. For example, wearable activity trackers not only quantify users' physical activities, but also motivate users by visualizing data for self-reflection and enable users to share data for social facilitation. Moreover, voice-based artificial intelligence (AI) assistants can recognize user moods and recommend personalized play lists based on those moods.

This paradigm shift has also sparked a *positive computing* movement toward the design of information technologies that promote personal well-being and help to fulfill human potentialities, beyond improving efficiency and effectiveness in knowledge work [1]. Furthermore, recent advances of mobile, wearable, and IoT technologies foster novel opportunities for designing and developing novel intelligent positive computing services that address various health and well-being issues, such as mental and physical health. The major benefits of intelligent positive computing services include the facilitation of novel means of detecting human behaviors that might signal well-being problems, delivering therapeutic interventions in a timely fashion, and tracking responses for assessing the effectiveness of the interventions. Positive computing research requires cross-disciplinary collaboration among computing, design, human-computer interaction (HCI), and psychology fields. Therefore, acquiring holistic perspectives on this research domain is very challenging. The goal of this work is to bridge this gap by providing an integrative review of existing studies for researchers and practitioners who strive to design, develop, and evaluate intelligent positive computing systems using mobile, wearable, and IoT technologies.

Towards this goal, we propose a conceptual framework of intelligent positive computing systems that leverage mobile, wearable, and IoT technologies. The core component of the framework is the collection of sensor data from mobile, wearable, and IoT devices to extract basic context features, such as physical activities. This sensor data is analyzed to detect behavioral markers of well-being problems, such as lack of physical activity and depressive symptoms. After the marker detection, favorable moments for delivering proper intervention content to the users are identified with a careful selection of delivery device and modality. This core component includes a user feedback loop to improve the accuracy of algorithms and accommodate user preferences. Furthermore, the system design component provides the evidence-based guidelines, which are established based on behavior principles and systematic evaluations. The system evaluation component considers the effectiveness of the systems for human behavioral change, while determining how and why the system is used for design improvement.

With the consideration of the proposed conceptual framework, we review the literature of six core areas that are critical for intelligent positive computing systems research, namely (1) design methodologies, (2) mobile platform design, (3) behavior marker detection, (4) opportune moment detection, (5) device and modality selection, and (6) evaluation methodologies. As positive computing spans a wide range of disciplines, our goal is to provide an overview of the related studies and suggest practical guidelines for intelligent positive computing research involving mobile, wearable, and IoT technologies.

In addition, a set of positive computing service scenarios is proposed by using a scenario-based design method [2]. The conceptualization of service scenarios provides researchers and practitioners preliminary yet useful insights on possible opportunities and challenges for positive computing system design and implementation. Our scenarios are targeted to college students as a focal lens for designing positive computing systems as young adults are the early adopters of new information technologies. Moreover, many of them tend to be vulnerable to health and well-being problems, such as depression and addiction, on account of their developmental dynamics and relative independence from social roles and expectations [3, 4, 5, 6]. We conclude this paper by providing the research directions on the core topics in positive computing systems research, specifically system design, platform design, behavioral markers, opportune moments, device/modality selection, and evaluation. In addition, we discuss various concerns and challenges of positive computing systems that researchers and practitioners should consider such as dependence, abandonment, side effects, privacy, and ethical issues.

2. Mobile, Wearable, and IoT Technologies for Intelligent Positive Computing

Intelligent positive computing leverages mobile, wearable, and IoT devices such as smartphones, activity trackers, voice assistants, and smart sensors. These technologies support refined sensing and tracking of a user's status ranging from physiological signals, such as heart rates and skin temperature, to physical activities, social interactions, and their interaction with everyday objects. In addition, IoT devices provide *physical actuation* such as controlling light bulbs, door locks, and thermometers, as well as *virtual actuation* such as emailing users of sensed events. It is also possible to collect many kinds of sensor data through the web, such as weather and air quality, via open Application Programming Interfaces (APIs), which are known as virtual sensors [7].

Smartphones are equipped with various sensors (e.g., GPS, motion sensors, compass, ambient light, camera, and microphone). Mining sensor data facilitates an improved understanding of user contexts and detection of various events of interests [8]. For example, smartphone sensing provides location tracking with GPS, activity tracking with motion sensors (accelerometers and gyroscopes), and social interaction tracking through audio sensing (or call/SMS log analysis). Similar sensing features are also supported by wearable devices such as smart watches and activity trackers. For example, Apple Watch 3 and Samsung Gear S3 include a built-in GPS, barometer, heart-rate sensor, accelerometer, and gyroscope. A major advantage of wearable devices is their support of sensing physiological signals, such as heart rate, electrocardiogram (ECG), and skin temperature, which are useful for detecting stress and emotion [9]. Several wrist-worn devices enable researchers to access raw data such as Empatica E4, Shimmer3, and Philips DTI-2—unlike popular wrist-worn devices, these devices provide APIs for accessing raw data, but their cost is an order of magnitude greater. There are also sensing devices that can be attached to the smartphones as in AliveCor's KardiaMobile ECG that allows users to track their heart conditions.

IoT devices can be largely classified based on functionality. A popular trend is domestic and office environments. Major functionalities in this area include IoT hubs, voice assistants, lighting and switches, outlets, sensors (e.g., motion, temperature, moisture/water), door locks, thermostats, and speakers. Most IoT devices are standalone products that are controllable through local networks or the Internet. For example, users can wirelessly control connected door locks and thermostats through the Internet. Several platforms, such as SmartThings [10] and Sen.se Mother [11], provide the central hubs that wirelessly connect multiple sensing and actuation devices. Popular sensing mechanisms include motion and activity sensing, thereby enabling tracking user location and activities, as well as user interactions with everyday objects. Sen.se Mother [11] has motion tags called Motion Cookies that can be attached to any kinds of trackable objects (e.g., pillbox, keys). When a user carries this sensor tag, it automatically tracks the user's physical activities, such as step counting. There are also standalone tags for sensing such as TI SensorTag and Cao Wireless Tag. In addition to sensing, IoT devices with actuation features include smart bulbs, outlets, door locks, thermostats, and speakers, which may be connected to the central hubs for integrated control. Voice assistants include Amazon Echo and Google Home, which provide natural language support for information activities (e.g., Q&A) and device control (e.g., turning off the blubs). This kind of IoT device controlling is also supported by the IoT hubs, such as SmartThings Hub.

3. Conceptual Framework for Intelligent Positive Computing Systems Research

The major benefits of using mobile, wearable, and IoT technologies for positive computing are their novel means of measuring and tracking well-being problems and delivering intervention methods. The large amount of personal big data collected from these devices helps elucidate the user context. Thus, mining this big data provides new approaches of measuring/tracking well-being problems. Furthermore, mobile, wearable, and IoT technologies enable intervention to be enacted in a timely fashion and therapeutic responses to be tracked.

In Figure 1, we present our conceptual framework for intelligent positive com-

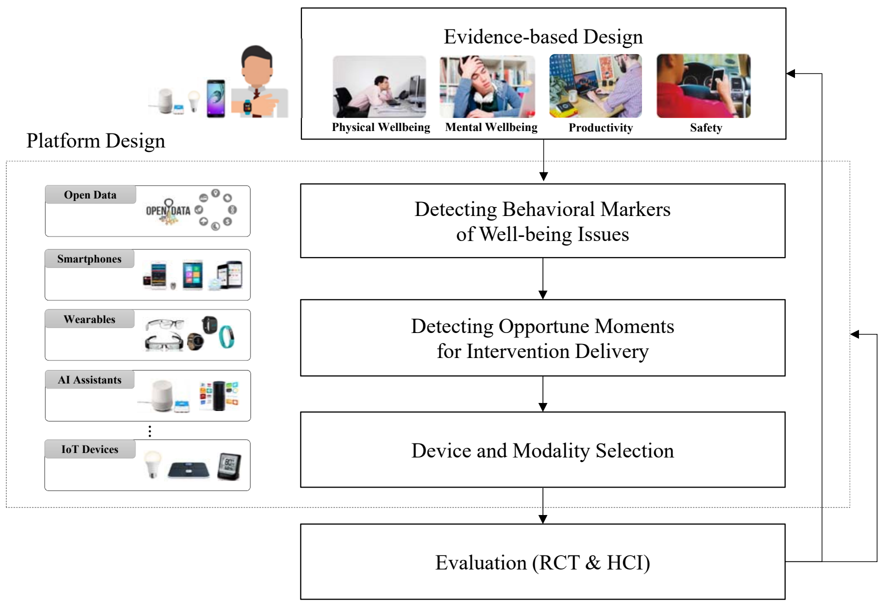


Figure 1: Intelligent positive computing framework for delivering in-situ, intelligent, in-time, intimate, and incorporated services (i.e., the so-called "five I's" of intelligent positive computing)

puting. The top of the figure illustrates the design of a positive computing system that aligns with guidelines for designing *evidence-based persuasive systems*. These guidelines are established based on behavior principles (e.g., operant conditioning, motivation) and systematic evaluations (e.g., long-term randomized field trials, user experiences). A positive computing system is used as a strategy for behavior change. Thus, the effectiveness for human behavior change must be properly evaluated. A well-known approach is to use a randomized controlled trial (RCT), which randomly assigns participants into experimental groups (e.g., control versus experimental group) without revealing the assignment information. With the objective of gaining design knowledge for better system design, human-computer interaction (HCI) researchers should additionally tailor evaluation procedures to understand *how and why* the system is to be used by its target users.

Positive computing platforms collect sensor data from mobile, wearable, and IoT devices, including virtual sensors from the web. The sensor data can be preprocessed to extract basic context features such as physical activities and semantic user locations. Behavioral markers of well-being problems are then detected by mining sensor data streams. For example, lack of physical activity and aggressive driving can be inferred by processing motion sensor data. Unique behavioral patterns of a depressed individual can be discovered via mining sensor data streams from multiple data sources, such as app usage and mobility traces [12, 13, 14].

Once a behavior marker is detected, it is necessary to identify opportune moments for interruption. For example, we can consider the points at which it is most appropriate to ask a user to perform certain activities that can help them to recover from a depressed state. Afterwards, we then can select the device and modality for information delivery. For an alert message delivery, for example, it might be desirable to consider which device to choose (e.g., smartphones or smartwatches), and which output modality to employ (e.g., auditory or visual output). These three steps—detecting behavioral markers and opportune moments, and selecting the device/modality—may include user feedback loop for incremental learning. A user's feedback can be employed to improve the accuracy of the detection algorithms and provide the opportunity for accommodating the user's preferences.

This conceptual framework provides in-situ, intelligent, in-time, intimate, and incorporating services—the so-called "five I's" of intelligent positive computing—using mobile, wearable, and IoT devices. *In-situ sensing* in a user's daily life provides quantified self and context data collection. Mining this personal big data enables *intelligent* identification of problematic situations *in time*. An intervention can then be *intimately* provided to the user by using always-on mobile, wearable, and IoT devices in a personalized fashion. In addition, *incorporating* a user's feedback continuously improves the service experience.

Figure 2 depicts an example service scenario of promoting active lifestyles by delivering a set of intervention methods, specifically addressing unproductive use of smart media, a lack of physical activities, and stressful tasks. In the first stage, we collect quantified self and contextual data from the users; this includes not only sensor data, but also user self-reported data (e.g., level of stress, emotional state). Various behavioral patterns that are related to well-being problems are monitored. In addition, data mining

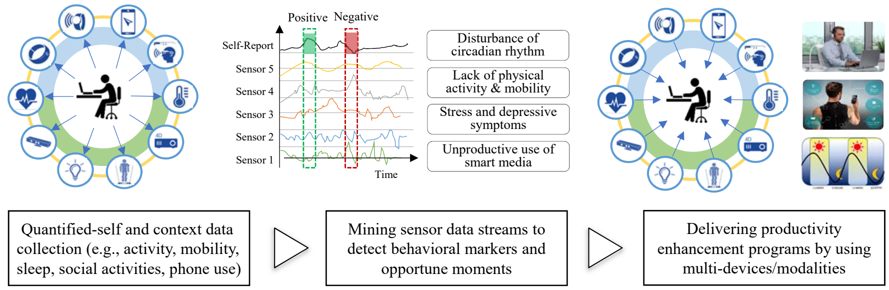


Figure 2: Example service flow of the conceptual framework

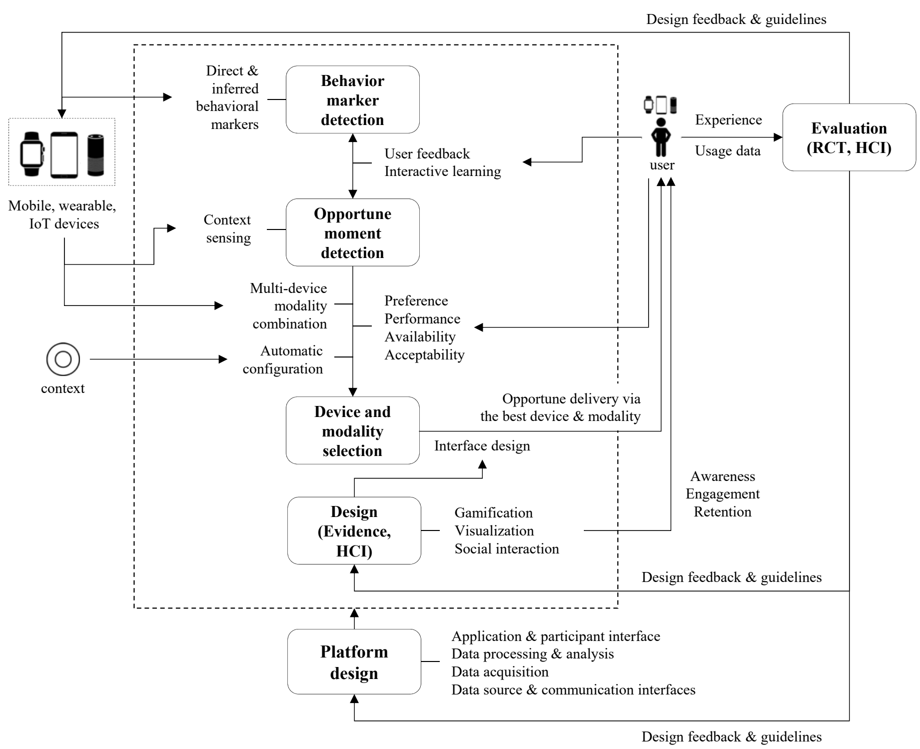


Figure 3: Key components of the proposed conceptual framework: (1) design, (2) platform design, (3) behavior marker detection, (4) opportune moment detection, (5) device and modality selection, and (6) evaluation

can be used for uncovering a person's unique behavioral patterns. Automatic detection of behavioral markers and opportune moments enables context-sensitive, timely delivery of intervention content to the users.

4. Literature Review of Key Components

We review the literature of the major components—design, technologies, evaluation—in positive computing systems research. Given the fact that positive computing research spans a wide range of disciplines, our goal is to provide an overview of the related studies and suggest guidelines for positive computing research with mobile, wearable, and IoT technologies. The key components of intelligent positive computing systems are illustrated in Figure 3.

4.1. Evidence-based Design for Positive Computing Systems

In the field of positive computing research, the target domain of human behavior has expanded from a single simple behavior (e.g., walking, running, Internet use) to a complex set of implicit behaviors (e.g., health, productivity, sustainability) [15, 16, 17, 18, 19]. Change in human behavior is a highly complex process and is affected by a tremendous number of both internal and external factors and their combinations [20]. We categorize factors presented in the existing literature, such as the health belief model [21], social cognitive theory [22], and theory of reasoned action [23], into five major factors that influence the likelihood that an individual will change their behavior.

The first factor is *perceived susceptibility*. An individual must feel susceptible to a negative condition with severe consequences. The second is *perceived benefit*, which is one's belief in the efficacy of behavioral change that outweighs the perceived barriers to tangible and psychological costs of the change. *Self-efficacy* is one of the most important factors influencing an individual's behavioral change. That is, a subject must have confidence in their own ability to take action under different circumstances. *Incentives* involving physical outcomes, social outcomes, or even self-sanctions also play a key role in leading behavior change. Moreover, it was found that perceptions of *social pressure* (i.e., social reinforcement) prompt individuals to perform the target behavior. Furthermore, these factors are continually interacting with each other, resulting in satisfaction with one's behavioral performance and maintenance of the behavior, or in dissatisfaction and possible termination of the behavior. Thus, researchers of positive computing systems must utilize these variables to understand a wide variety of human behaviors and consider them for designing persuasive approaches.

Meanwhile, there are many challenges remaining when designing effective persuasive systems. Prior research identified barriers to behavior change, including lack of motivation, resistance to change, and/or compliance. Thus, recent efforts have been made for those who suffer from psychological difficulties on account of the urgency of their problems [24, 25, 26, 27]. Some interventions were found to be effective via scientifically sound methods (e.g., cognitive behavior therapy), while others continue to be used without validation by sound evidence (e.g., psychodynamics) [28, 29]. Regardless of theoretical differences across these attempts, they all share the same component—the primary agency for behavioral changes in human beings [30]. Furthermore, human beings by nature are very flexible, sensitive, and responsive to the complexity of human behaviors [31].

Now, every design of a technical device, system, and program that is relevant to positive computing should consider various aspects of human behaviors which could be translated into design components [32, 33, 34]. Literature in HCI has been shown to heighten the awareness of design components, and many attempts to identify the critical design components have been made [35, 36, 37]. Researchers have suggested new approaches to providing *simple, accessible, scalable, and sustainable* regimens that meet user needs to promote positive behaviors [38]. Computational methods or interventions have been developed to promote physical and mental well-being [15, 39, 40, 38]. For instance, wearable technologies provide opportunities to monitor stress [41]. Some studies further suggest just-in-time interventions by analyzing a user's state in-situ [15, 42].

Although these interventions are found to be useful, research on them are dominantly based on either the synthesis of expert opinion or the results of short-term user studies, limiting its usefulness as a guide for designing *evidence-based persuasive systems*. To date, what is generally lacking in the literature, except in certain studies, is this systematic approach to identifying design components based on behavior principles and evaluations. Nevertheless, a notable study led to the proposal of a health-coaching system for stress reduction built upon successful behavioral approaches [38]. Incorporating behavior change theory, such as self-efficacy theory [22] and goal-setting theory [43], researchers have attempted to tackle a paradox of compliance; users fail to persist with given health-promoting behaviors because goals are too easy or too difficult. To address this concern, they suggested a smartphone-based adaptive coaching system that modifies goal difficulties based on a user's previous performance to provide more tailored and contextualized suggestions. In summary, system designers should extract principles from a theoretical framework and translate them into key technical features of the system as clearly illustrated in the behavioral intervention technology model [44].

Based on the learning principles, behavior theory (BT) can provide guidelines for the design of persuasive computing systems for behavior change. BT is one of the theories that were developed to explain human behaviors and methods following BT principles have been successfully used to modify human behaviors [45]. In BT, factors affecting human behavior are categorized into two parts, consisting of personal factors (e.g., personality, learning history) and environmental factors (e.g., settings, people, and systems) [46, 47]. This framework of behavior theory provides a conceptual lens on the design of persuasive computing systems because it provides guidance on the focus to maximize human behavior change [48, 49]. It has been well documented that one's personality is a stable condition that is resistant to change [50, 51]. Learning the history of a person is also the same; past experience cannot be modified [52]. However, environmental factors are malleable and can be modified [53, 54]. In addition, operant conditioning, one of the primary learning principles that explains how people learn what actions to take and not to take based on reward or punishment, provides a classification system (e.g., antecedent-behavior-consequence) to analyze environmental factors (for more information, see Cooper et al. [55].)

In the field of HCI, positive computing devices, programs, and systems are the major agency for behavior change in accordance with environmental factors. To maximize change in human behavior, design components should be analyzed and classified according to their functions that follow the principles of behavioral learning or operant conditioning. Although the necessity of each specific environmental factor as a critical design component should be empirically tested, a few can be identified from the analysis of design components commonly used in other research domains. For example, as antecedents, which are factors influencing behaviors, we should consider plots (e.g., story, contents, narrative, goal setting, self-monitoring), sensation (e.g., graphics, sounds, beauty), problem solving (e.g., challenge, intellectual demands), social interaction (e.g., social exchange, sharing ideas and products), diversity (e.g., change over time, exploration), and balance across design components [56, 57, 58], when designing a system. Ongoing efforts will follow the means of prioritizing or combining these components to maximize the effects of each system.

More importantly, the reward system, which is the consequence of the operant conditioning and the key mechanism of human behavior change, should be intensively focused on and extensively analyzed. Various types of reward systems, such as performance feedback, points, level system, and tangible rewards. [59, 34, 60, 61], have been successfully used. In BT, a reward is conceptualized as a component that increases the frequency of a behavior as a consequence of experiencing a certain stimulus [46]. Furthermore, the reward system is complex because it is influenced by type, timing, amount, and contingency of the stimulus, as well as by the motivational level of the person who receives the stimulus [55]. Only the careful consideration and combination of these components can guarantee the best design for the positive computing persuasive interaction platform.

4.2. Platform Design

In this sub-section, we present the major building blocks of a general-purpose software platform and then provide a detailed review of well-known software platforms.

4.2.1. Major Building Blocks

Recent advances in mobile technologies—smartphones with wirelessly connected wearable and IoT devices—enable myriad applications that can help people better manage their health, wellness, and productivity. These applications usually require collection of fast-growing raw sensor data (e.g., from an accelerometer, gyroscope, luminance, GPS, and so on) in order to extract the user's behavioral markers by processing and analyzing the collected data and then identifiying appropriate moments for interventions to better achieve their objectives (i.e., persuasive interactions). To this end, development of a general-purposed platform is highly necessary; however, the endeavor

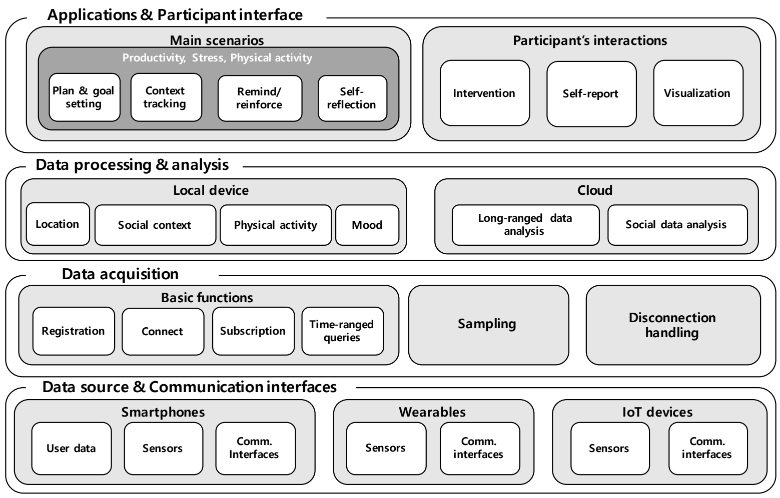


Figure 4: General-purposed platform architecture: A layered approach

has various challenges. These challenges include concurrent wireless connections to a wide variety of devices, handling of a large-volume raw sensor data influx, reliable and flexible storage, quick and accurate sanity check of incoming data, efficient data sharing between components, extensibility to a wide range of sensors, and effective user interactions with the platform.

In Figure 4, we conceptualize the platform components with a layered architecture. (1) *Data source and communication interface* includes sensor data (e.g., accelerometers, gyroscopes, magnetometers, GPS, etc.) from smartphones and other wirelessly connected devices, and communication interfaces deliver the collected data to smartphones or cloud for further processing. (2) *Data acquisition* includes implementation of an agent that performs basic functions (e.g., device registration, connection, subscription, and time-ranged queries), sampling rate of raw sensor data, and connection failure handling. (3) *Data processing and analysis* implements deterministic algorithms or machine learning methods for processing of raw sensor data to extract contextual features and behavioral markers. The processing and analysis can be performed in-situ at a mobile device or it could be offloaded to the cloud. (4) *Application and participant interface* includes intervention applications and participant's interactions. The application usually contains pipe-lined procedures (i.e., plan and goal setting, context tracking, reminding/reinforcement, and self-reflection). The participant's interactions involve user interventions, user's self-reporting to the platform, visualization of the user's daily (or longer) summary.

4.2.2. Review of Recent Platform Studies

Existing platforms can be broadly classified in two categories: vendor-specific and vendor-agnostic software platforms. First, vendor-specific platforms are more akin to data collection systems because they are solely developed by the vendors, such as Fitbit, Garmin, and Xiaomi. These types of platforms typically provide specific applications to support vendor-specific devices and support cloud-based back-end storage. For example, Fitbit trackers and Garmin wearables upload activity tracking data to a smartphone and optionally to back-end storage to overcome the memory shortage of wearables. However, such vendor-specific platforms do not allow accessing raw sensor data and thus, extensibility is quite limited.

Second, vendor-agnostic software platforms aim to support a broad spectrum of vendor agnostic devices and a rich suite of services. These software platforms can further be divided into commercial and research platforms. Examples of commercial software platforms include Apple HealthKit [62], CareKit [63], Google Fit [64], and Microsoft HealthVault [65]. These types of platforms typically provide a rich suite of services. For example, Google Fit supports a broad spectrum of wearable devices to track health-related information such as steps, time, distance, burned calories, and sleep. Apple's ResearchKit and CareKit further provide an ecosystem for developers to build apps that enable users to manage their well-being on a daily basis. Compared to vendor specific ones, these provide an optimized in-situ storage engine (i.e., smartphone) and a broad spectrum of queries over the back-end cloud. However, these types of platforms cannot fully support high-rate sensor data influx, data quality assessment, behavioral marker extraction, context-triggered user data collection, intervention design, self-report data collection, and privacy management.

In recent years, research communities made considerable efforts to build novel platforms that can deal with such limitations. Research platforms often have diverse objectives (e.g., time management, mood, stress, etc.) and procedures (i.e., behavioral markers extraction, intervention, reinforcement, etc.). Despite such diversity, they share the similar design challenges and major building blocks of a general-purposed platform as described in Figure 4. For example, ContextPhone [66] focuses on contextual information as an understandable resource for users. With the help of widgets, users can control the sensor data collection. Ohmage [67] is a smartphone-to-web toolkit designed to create and manage the experience of sampling-based data collection campaigns in support of mobile health pilot studies. Moreover, it is accessible in multiple platforms. Similarly, CenceMe [68] infers the physical-social context and shared information through back-end server processing to match commonly shared social contexts to raise social awareness. Momento [69] is integrated with a ContextToolkit server to analyze audio segments to detect proximity of people. A middleware approach can be adapted and extended to support extensibility. AWARENESS [70] focuses on privacy of users. The context in this platform is shared with previously trusted devices, and a smartphone user is the sole controller of privacy aspects. However, this may sacrifice the quality of context based on the extent to which the context is shared at a given time. To cope with extensibility, OpendDataKit [71] adopts a middleware design approach and allows developers to minimize their efforts on sensor-specific codes via reusable sensor drivers (downloading new sensor capabilities from an application market without any modifications) and provide the management of discovery, communication channels, and data buffers. AWARE [72] provides mobile data-logging tools, and it supports external sensor plugins to collect and abstract sensor data for context-aware service delivery. More recently, mCerebrum [73] significantly improves scalability of storage for high-rate sensor data and further provides several fine-tuned features, such as sensor duty-cycling, energy-optimized context inference with inference computation as a shared service, and sensor data quality assessment [74].

4.3. *Behavioral Marker Detection*

We present the concept of behavioral markers and illustrate the importance of this concept by describing various applications. We then provide a detailed review of well-known behavioral marker detection methods.

4.3.1. *Defining Behavioral Markers*

The concept of behavioral markers is closely related to that of the digital phenotype, which includes a set of observable characteristics of an individual through mobile and wearable devices such as activity trackers and smartphone loggers [75, 76, 77, 78]. A person's digital phenotype may cover various data sources, ranging from passive sensor data (e.g., self-trackers, smartphone logging) and social media use to active self-reporting (e.g., mood, stress). Similar to genotyping, which aims to find associations between genetic variants and disease (as in typical biomarkers), the goal of digital phenotyping is to uncover mappings between digital phenotypic variants and diseases of interests (e.g., depression). In this case, most of digital phenotype data are related to an individual's behaviors; thus, such digital phenotypic variants are called "behavioral markers."

Digital phenotyping provides major benefits to well-being care in terms of diagnosis, treatment, and management. Mobile and wearable devices can collect and analyze an individual's digital phenotype data in real time. This means that we can perform "continuous and unobtrusive measurement and inference of health, behavior, and other parameters" using mobile and wearable devices [77]. In other words, digital phenotyping helps to re-define the manifestation of well-being problems, provide alternative approaches for measuring such problems, deliver interventions in a timely fashion (known as just-in-time delivery), track therapeutic responses of delivered interventions, and enable proactive management of problems in well-being (e.g., remission/relapse monitoring, risk prediction) [75, 77, 79]. In particular, automatic identification of well-being problems is the key enabler of the just-in-time intervention [80]

Prior studies attempted to define behavioral markers based on digital phenotyping techniques [76, 78, 77]. Harari et al. [76] proposed a behavioral model with three dimensions: social interaction, daily activities, and mobility patterns. Behavioral markers in each dimension are further defined based on sensor data processing (e.g., duration of social interaction). Likewise, Mohr et al. [78] proposed a hierarchical feature model in which low-level sensor data are transformed into low-level features that constitute high-level behavioral markers. As in the case of the work of Harari et al. [76], most low-level features are human interpretable features such as location type, activity type, movement intensity, and phone usage. Given that existing models are predominantly based on human interpretable features, it is very natural to extend these models to include "contextual models." Here, "context" means "any information that can be used to characterize the situation of a person" [81]. As Schmidt et al. [82] defined, a context describes a situation and the environment in which a device/user is situated by a set of relevant features over several domains. For example, human factors may have feature sets of a user, social environment, and tasks; the physical environment may have feature sets of conditions, infrastructures, and locations. However, recent studies tended to extract various kinds of features for machine learning, and thus, pure contextual meaning of extracted features may be weak in reality, such as the entropy of app usage.

Behavioral markers in prior studies can be classified as two types: direct behavioral markers and inferred behavioral markers. Direct behavioral markers are the types of behavioral markers that are directly measurable using sensors based on prior knowledge of well-being problems. For example, lack of physical activity and aggressive driving can be directly measurable using motion sensors. However, in most cases, there is a lack of prior knowledge on detailed manifestation of well-being problems. For example, a person with depression may show different behaviors when the person falls into the depressed state; however, we do not know what kinds of behaviors are related to the depressed state. In this case, we can collect self-reported data and digital phenotype data to find meaningful behavioral markers related to the depressed state. In addition, we can use a standard diagnostic manual such as Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [83], which informs us of which behavioral features to extract as Wang et al. did in their recent studies on depression tracking [84].

4.3.2. Review of Recent Behavioral Marker Detection Studies

In the case of direct behavioral markers, motion sensors are mainly used to detect various types of physical activities such as movement [85], sleeping [86], eating [87], and agitation [88]. For example, SitCoach detects that a user is in a sedentary state if the arithmetic difference between consecutive accelerometer samples is smaller than a predetermined threshold [85]. Eating gestures can be recognized by applying machine learning on wrist motion data measured from smartwatches [87].

Prior studies on inferred behavioral markers examined various well-being problems, such as depression [12, 13, 14], social anxiety [89], bipolar disorder [90], smartphone addiction [91], and schizophrenia [92].[3] As a representative case, we review how smartphone GPS traces can be used to automatically identify the depressed stage of an individual [13]. First, Canzian and Musolesi [13] built a mobile app that was to collect GPS traces as well as self-reported data about an individual's depression level using PHQ-8, an eight-item questionnaire with the sum of items denoting the level of depression [93]. This app was distributed via Android Play Store. During the data collection period, users answered the PHQ-8 questionnaire daily, and both self-reported and mobility traces were transferred to the remote server for data analysis. For reliable data analysis, the dataset was collected for two months. After data collection, GPS traces were preprocessed to find a sequence of places visited. For any time interval, various daily mobility features could be extracted such as places visited, place variety, distance travelled, and regularity of daily routines. Thus, for a given user, we can prepare the vector for each day, i as follows: (PHQ score for day i, a set of mobility metrics for day i). This dataset enables us to investigate what kinds of mobility features are correlated to the PHQ scores. Furthermore, predictive analysis can be performed. For example, multiple regression can be used to understand the predictive power of various mobility metrics. If each day is labeled as a depressed or normal day based on the PHQ score threshold, we can run classification algorithms such as support vec-

[3] Schizophrenia is a mental disorder characterized by abnormal social behavior and failure to understand reality; thus, someone with this disorder may have difficulty in distinguishing between what is real and what is imaginary.

tor machine (SVM). Saeb et al.'s work [14] further considered both GPS traces and smartphone usage and found that mobility and phone usage were significantly correlated with depressive symptoms. Besides sensor data processing, many prior studies alternatively analyzed social media data (e.g., Twitter, Facebook) to identify various well-being problems such as suicidal ideation [94] and depressive states [12].

4.4. Opportune Moment Detection

We define the concept of opportune moments and highlight its importance through examples. In addition, we present a review of recent studies on opportune moments.

4.4.1. Defining Opportune Moments

Positive computing is intended to help users change behaviors or attitudes. It is thus imperative for the system to persuade its users to restrain undesirable actions and promote their desirable actions. Timely delivery of intervention with mobile, wearable, and IoT technologies requires that the system interrupt users to draw their attention from their current task. Interruption of users can be achieved using various forms of alert methods, including a visual, vibration, sound notification, or a combination of these for information delivery. Typically, these interruptions occur when an event is detected (e.g., behavioral marker detection) or it follows predetermined scheduling on a regular basis. However, it is well known that off-task interruptions often result in productivity loss, increased stress, and time pressure [95]. Thus, it is very important to consider interruptibility, which is a user's receptiveness to interruption or perceived burden of interruption [96, 97, 98, 99].

Prior studies found that a task changing moment is most suitable for interruption because that results in the lowest resumption lag and user annoyance [100]. In the mobile environment, Ho and Intille [97] showed that an opportune moment for interruption should consider the patterns of various user activities such as physical activities and social engagements. For accurate prediction, it is very important to carefully consider a user's contextual model, which can be built based on a user's current location, activities and interactions with other users, by using various built-in mobile sensors (e.g., GPS, accelerometer, microphone, application contexts and Bluetooth signalling).

Besides current contexts, we can also consider a user's past context information; for example, Choy et al. [101] showed that beyond immediate past, looking back on the current day can significantly help detect opportune moments.

The first step towards automatic identification of opportune moments is to collect sensor data and user feedback (or label) of interruption instances. We can use simple context sensing algorithms for automatic detection (e.g., activity change detection [102]), or apply machine learning methods by extracting various contextual features [103, 101]. An explicit way of collecting user feedback is to ask users to label every instance (e.g., Likert scale rating), which is challenging and laborious [102]. Thus, it is important to selectively ask users to label instances as in active learning or decision-theoretic modeling used in BusyBody [104]. Alternatively, we can use a passive way of labeling interruptible moments. For example, *InterruptMe* used a user's responsiveness to notification to judge whether a user is interruptible [103, 80].

4.4.2. Review of Recent Studies on Opportune Moments

Prior studies attempted to detect opportune moments in various contexts. "Let's FOCUS" is an app for helping students to self-regulate their smartphone use in classrooms [105]. This service automatically detects a context switching moment of arriving at a classroom with indoor localization and nudges a student to lock the phone for self-regulation. In their user study, Park et al. [106] uncovered several social contexts for interruption such as long silence and a user left alone, which can be automatically identified using built-in mobile sensors.

BreakSense aims to promote physical activity of office workers [102] by nudging users to engage in more physical activities when they start moving away from their desks. The system automatically detects their movements by using motion sensors, and then it sends a notification of asking them to take a short break challenge of indoor walking. Indoor mobility can be monitored with Bluetooth beacons, and completion of break challenges can be automatically checked. The field trial results revealed that nudging users at the opportune moments and challenge-based gamification served as major motives for active engagement.

4.5. Device and Modality Selection

We explain the basic concepts of device and modality selection. We then provide a detailed review of recent device and modality selection studies.

4.5.1. Introduction to Device and Modality Selection

We consider multi-device environments with mobile, wearable, and IoT devices which typically have multiple input and output modalities. Here, a modality means a single independent channel of sensory input or output between a computer and a person. It is likely that devices have different form factors, and their modalities vary widely (e.g., screen size, vibration patterns, actuation support). The primary effectiveness of any kinds of ubiquitous-technology-based intervention is the production of a successful delivery of the intervention to the users. In multi-device/modality scenarios, it would be important to understand what the most effective mechanisms are for information delivery. In addition, it is possible to learn about the user behaviors and their preferences for more effective delivery.

Device and modality selection is critical in intervention delivery. It is closely related to the types of information (e.g., text, picture/video, and audio) and comprises a required user interaction. Furthermore, there are several constraints to consider for the device and modality: user preference, attention performance, device availability, and acceptability. User preference means that users may have a preference in the device and modality selection. For example, a user may prefer to receive messages using smartphones instead of smartwatches. Attention performance is related to an output modality; for a given environment, it is the signal-to-noise ratio (or attention focus) of a given device and modaility pair. For example, in a noisy environment, it is difficult to perceive sound-based notifications. According to the multiple resource theory, it is possible to perform tasks simultaneously as long as they differ in their type of resource demand (e.g., visual and auditory), and incoming stimuli are filtered based on their level of relevance [107]. Acceptability refers to the extent to which the device and modality selection is personally or socially acceptable because information delivery with such selection may cause a distraction or a disturbance. Thus, we can formulate the device and modaility selection as follows: for a given intervention content (instruction and visual aids) and availability/attention/preference/acceptability constraints, it must be determined how we should render the content for a given set of devices and its modalities.

Several studies investigated various aspects of device and modality selection. Researchers examined notification awareness and accuracy by varying the types of output modalities: visual, auditory, tactile, and olfactory outputs. For example, a user study showed higher accuracy and user preference in visual and auditory outputs compared to other modalities [108]. When multiple modalities were available (e.g., audio and visual outputs), output modality-combinations could be more effective than single modality use [109]. For example, in recall tests, study participants showed that spoken text with pictures showed the highest recall performance. Various studies also reported on availability, acceptability, and preference issues [109]. Shirazi and Henze [110] identified the priority of device preference over different content types (e.g., messenger for smartphone, and calendar checking for smartwatches). Weber et al. [111] identified device preferences based on screen size and availability. Jeong et al. [112] determined that it is important to consider interaction availability (e.g., whether devices are nearby or interactable) and social acceptability (e.g., whether it is okay to interact under the given social circumstances). Automatic configuration of a mobile phone's notification modalities is closely related to our modality selection. For example, Sensay [113] is a context-aware mobile phone in which, by performing context sensing, it can automatically configure a phone's output modality for notification delivery. For example, during a meeting, it can change its notification mode to visual notification with LED signals.

4.5.2. Review of Recent Device and Modality Selection Studies

We then review three case studies on device and modality selection: (1) user preferences of smartwatch wearing behaviors [112], (2) automatic configuration of the output modality [113], and (3) the multi-device/modality combination [114]. In terms of wearable devices for intervention delivery, it is important to understand wearing behaviors. Jeong et al. [112] collected a longitudinal activity tracking dataset of 50 Apple Watch users. Participants showed the following patterns of diurnal usage: those who tended to wear during the work hours (58%), during active hours (38%), and on all days (12%). For example, work hour wearers tended to remove their smartwatches after work (or at home). For these kinds of people, it would be difficult to deliver content after work hours using smartwatches. In contrast, all day wearers were likely to wear their devices, even in bed, and a significant opportunity exists for anytime content delivery. Smartwatch usage was preferred when users desired immediate responses or engaged in multitasking. For example, a user desired receipt of a phone call while biking. However, its usage was nuanced in that, for some users, the capability of making immediate responses may have been a major reason for them not wearing the smartwatch at home. This case study clearly showed that carefully understanding wearing behaviors and user preferences is crucial for effective intervention delivery.

Automatic configuration based on context awareness is the key enabler for intelligence device and modality selection. The context-aware mobile phone, Sensay, modifies its ringer mode based on the user's state and environment [113]. Specifically, it uses multiple sensory data (e.g., light, motion, sound) for context recognition and it changes the output modality for notifications based on the user current state, namely, uninterruptible, idle, active, or normal. Although the study in [113] focused only on a single device, we can naturally extend this concept to multi-device environments [115]. For a given context, we can rank which device and modality is the most appropriate by considering the device availability, user preference, and attention performance [110, 111, 112].

The availability of multiple personal devices and shared IoT devices engenders novel opportunities for intervention. Lee et al. [114] studied how these devices, as interactive instrumental materials, can be used for behavioral changes by enabling context-aware just-in-time intervention. In their sleep intervention studies, they found that participants were able to configure multiple devices, such as smart plugs, smartphones, and speakers for behavioral changes. For example, when a phone is not charged on time, sad music will be played through a speaker. Since IoT devices are equipped with various actuation features, combining multiple devices and context data has enabled the design of novel intervention methods that can significantly improve the effectiveness of content delivery.

4.6. Evaluation of Positive Computing Systems

The primary goal of a positive computing system platform is behavior change. As a strategy for behavior change, the effectiveness of platform should be proved [116, 117, 32]. Use of a mobile, wearable, or IoT device, or its embedded program without empirical evidence is less acceptable in both business and real-life settings [60, 33]. The most recommended experimental design to prove the effectiveness of an intervention for human behavior is a randomized controlled trial (RCT) [61, 118]. For a study to be classified as an RCT, a random assignment of participants to experimental groups (e.g., control versus experimental group) and a double-blind design (e.g., a study in which both the participants and experimenters do not know to which the group participants belong) should be conducted. The RCT proves its effectiveness by statistically testing the significance of the dependent variables (DVs) (e.g., changes of scores before and after the intervention) between the experimental and control groups, while controlling extraneous variables (e.g., age differences, duration of device use) that may affect the DVs [119, 120]. The RCT research design has been adopted as an evaluation method for newly developed drugs, programs, devices, and systems in several disciplines, including pharmacology, psychology, education, economy, and political science [121, 122, 123, 124, 125, 126].

However, it appears that the RCT research design has not been widely applied to the study of the effectiveness of computational interventions for health monitoring yet [33, 60, 61]. Recently, several studies reviewed RCTs to investigate the effectiveness of computational interventions developed for health promoting behaviors [33, 60, 61]. Although works varied, depending on the topic, search period, and data selection criteria, researchers were able to identify only a handful of RCTs. For example, a review of studies exploring the effectiveness of using smartphone applications to promote physical activity concluded that 55% of studies used an RCT (11 studies out of 20) [32]. Other studies that explored web-based interventions for health enhancement found that only 14% of studies (11 studies out of 83) used an RCT [33]. It was estimated that approximately 21% of studies were published in the field of HCI. As the research topic was broadened, a lower frequency of using RCTs was found. Moreover, these RCTs accounted for only a negligible portion. Nevertheless, it should be noted that the number of these studies has been slightly increasing in the past five years. Additionally, many studies on HCI research methodologies clearly show the importance of adopting an RCT design [60, 61, 127, 34, 128, 33].

On the other hand, RCT design has limitations [129, 128, 130]. For example, conducting an RCT is very costly and resource intensive. Moreover, even when using an RCT in which all factors that may affect human behavior are controlled, the clear identification of the mechanism behind the target human behavior is not always guaranteed. Some HCI scholars contended that focusing on identifying how and why the target population uses the system is as important as examining the effectiveness of the system [129, 48]. They suggest that qualitative methods, including focus-group interviews and open-ended question surveys, have been frequently chosen to answer the hows and whys questions above [129, 48]. Thus, some of the limitations of an RCT can be properly addressed if researchers examine the dynamic aspects that affect the effectiveness of the system by incorporating qualitative studies. There are also several possible ways to compensate the weaknesses of RCTs. They include, but are not limited to, proper sampling of the target population, appropriate control groups (e.g., alternative intervention group instead of a waitlist control group), sensitive and objective outcome measures (e.g., physiological measures, big data), and evaluation for social validity and fidelity of intervention. The RCT study should cover an adequate amount of time to elicit human behavior change and should be retested at least 66 days [131] to check the continuity of the modified behaviors.

While an RCT can answer *whether* a specific, complete system engenders relevant changes in targeted behavior, it often reveals a minimal amount about *why* the system is or is not effective. Furthermore, a large-size RCT may not be suitable for the evaluation of technologies at early stages of development. According to [129], understanding user experience with the system and the underlying mechanism of a system's success or failure is exactly what HCI researchers should achieve to improve the design technologies. Beyond efficacy evaluation via RCT, HCI researchers should also consider performing either quasi-experimental or case studies with a focus of how and why the given system is being used from the user experience point of view. HCI researchers can define and even tailor outcome measures to the intervention strategies that a system employs. For example, researchers conducted an RCT with 77 participants over 12 weeks to evaluate a mobile and wearable system for promoting physical activity [132]. The researchers collected self-reported data via the International Physical Activity Questionnaire (IPAQ), accelerometer data, and data on changes in physical state (e.g., weight, body fat).

A field deployment study of the *TimeAware* that examines the effects of framing an individual's productivity [133] could be a notable example of how HCI research adequately incorporates RCT method into the evaluation. A total 24 participants were assigned to two different conditions (positive framing and negative framing) for an eight-week study, composed of a two-week baseline period, a four-week intervention period, and a two-week withdrawal period. Quantitative measures extracted from usage logs and qualitative findings from pre- and post-questionnaires addressed the researchers initial questions about how the framing strategy affects personal productivity. In addition, an evaluation study on an adaptive goal setting system for stress reduction recruited 65 participants [38] to examine how the system affect behavior change in the wild. Participants were randomly assigned to three different conditions and participated in the trial for a month. Since there is no universal measure for stress, the researchers combined multiple measures, including the Perceived Stress Scale (PSS), Depression, Anxiety, Stress Scale (DASS), and Cohen-Huberman Inventory of Physical Symptoms (CHIPS). After the test, perceptions of system usability were collected.

Although the above study examples employed a similar evaluation format, there remains no absolutely established standard evaluation technique and measure for behavior change technologies in the HCI field. HCI researchers should be able to tailor evaluation procedures to gain a deep understanding of how and why a system is employed by its target users particularly in the early stage of intervention technology development. At the same time, it is also important to consider an RCT experiment to show the effectiveness of an intervention technology, which contributes to accumulating *evidence-based design guidelines* for building effective positive computing systems.

5. Well-being Care System for College Students

Based on research opportunities and design considerations that we discussed in the previous section, we introduce detailed scenarios of the use of our mobile system and engagement in the context of positive and persuasive computing.

Scenario-based design is a method that focuses on describing the use of an information technology system in the development process [2]. It describes how people will use a system to accomplish work tasks and other activities through a sequence of actions and events, making envisioned possibilities more concrete. It provides researchers and practitioners preliminary yet useful insights on not only the opportunities and challenges of the system use, but also on the design work based on defining system operations (i.e., functional specifications).

Our target populations are college students. Earlier studies on tracking students' happiness and well-being in academia showed that many of them struggle with mental health issues [134]. For example, an annual survey by the University of California, Los Angeles, Higher Education Research Institute found that college freshmen reported feeling more stress and low "emotional well-being" but are increasingly spending more time surfing the web [3]. A report, based on a survey of over 1,000 first- and second-year university students, revealed that 82% of students at UK universities suffered from stress and anxiety and 45% experienced depression [4]. A study at the University of California, Berkeley [135], found that 47% of graduate students suffered from depression where the assessment factors included career prospects, overall health, living condition, academic engagement, sleep, and others.

A considerable amount of research has shown that stress, time management, physical activity, productivity, life satisfaction, and other factors are correlated [5, 6]. These are all primary elements handled by our system. In this paper, our scenarios specifically focus on *student well-being in school life* including (1) lowering stress, (2) encouraging physical activities, and (3) increasing productivity through the use of our proposed system of positive, persuasive computing. Figure 5 presents four scenarios of positive, persuasive computing for college students.

Scenario A relates to *physical activity* and *stress*. It describes a route recommenda-

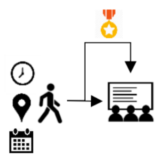
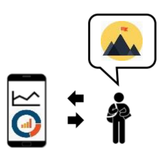
A. Route recommendation	
	Mary arrives at school 20 minutes before the class starts. On her way to the classroom (taking 10 minutes), she receives a notification from the smartphone suggesting a new route, which takes about 5 minutes more. According to the notification, 150 students have chosen the route to date, and their satisfaction with the route is 9.5 out of 10. As she is aware of her recent low physical activity, she decides to walk the recommended route. On her way, she receives a heart item as a bonus point.
B. Micro-break and user-generated content	
	After two hours of studying at the library, Mary receives a micro-break notification from her smartphone. As suggested, she stretches out and walks for five minutes. She then concentrates on studying again. After an hour, she finds course materials quite difficult to understand and gets stressed a bit. She turns on the smartphone, takes a photo of the library, adds hashtags, #hatestudying, #needrest and puts it on Instagram (the location of the library is also saved).
C. User status confirmation	
	Mary and her friends have a break at the café. When she is about to leave to prepare for the exam tomorrow, she receives a confirmation notification asking her current location, work, and stress level. She enters the information that the location of the café is correct, and she is taking a rest. She also marks that she relieves her stress at the current location.
D. Daily reflection & goal setting	
	Before going to sleep, Mary receives a summary of her daily activities such as the places visited, daily calorie consumption, types of app use, etc. She can confirm not only the information of her daily summary compared to other students, but also the score obtained so far. Today she achieved the goals she had planned, completed physical activities, received more points than others, and her rank went up, making her feel rewarded. However, she finds that she used SNS more than others than she thought. As she has an important exam early next week, she sets daily limits of SNS use and lets the smartphone notify her when she exceeds them.

Figure 5: Four scenarios of positive, persuasive computing

tion generated by the smartphone. We assume that Mary's schedule was already added to the smartphone database (*user feedback: schedule*). As one of the goals that Mary has set is increasing her physical activity (*user feedback: goal setting*), the recommendation is generated based on the remaining time for the class and Mary's current location (*behavioral markers*). The amount of time that will take from a recommended route should be less than the remaining time (*context-aware interruptibility*). Moreover, a summary page of the route, which includes the number of other users who employed the route and their satisfaction with it, can be offered together (*modality and interface design of interruption*). This may increase Mary's motivation of accepting the recommendation. A "heart item" conferred as a bonus point while following the route is likely to make her feel rewarded and increase engagement (*design methodology: gamification*). We expect that this may increase Mary's physical activity and lower stress, which may lead to increased productivity during the lecture.

Scenario B relates to *productivity* and *stress*. The smartphone knows that Mary is at the library. If Mary does not use the smartphone, it is reasonable to assume that she is studying (*behavioral markers*). After two hours of continuous study, the smartphone recommends that she take a short break (*context-aware interruption*). A short break is likely to refresh her mind and increase her study productivity. The next scenario considers the case in which Mary resumes use of the smartphone. As she finds the course material difficult and becomes stressed, she takes a photograph and posts it on Instagram with hashtags that mirror her current emotional state (e.g., #gettingstressed). Here, the smartphone can infer the correlation between her location, posted image, and stress level (*inferred behavioral markers*), which will be used to understand contextual information that is not only specific to Mary but also to potentially other users in the future.

Scenario C relates to *semi-automated tracking* (*user feedback*), which combines both manual and automated data collection methods [136]. Automated data collection with full reliability is difficult to guarantee, and thus, the smartphone periodically generates a notification consisting of simple, easy-to-understand questions and easy-to-fulfill options (*modality and interface design of interruption*). Through semi-automated tracking, we can confirm various types of current user statuses, including location, activity/action, stress level, etc. (*user feedback: confirmation*). Through this mechanism, the system can correct any incorrect information and better learn about the user and environment.

Scenario D relates to a user's *daily reflection* and *goal setting* (*user preference, feedback, and design methodology*). The system offers a summary of various aspects of smartphone use, including type, length of the apps used, list of the places visited, time remaining at each visited place, daily calorie consumption (number of steps, time for walking, etc.), and others. The summary page should allow Mary to easily reflect on her smartphone use, physical activity, stress, and productivity in various time frames such as daily, weekly, and monthly (*design methodology: interface design*) [105, 137]. It should allow Mary to easily set her new goals based on the summary results, which are new action items that will be considered by the system when generating interruptions. It can be also combined with a social component by providing information on how other people (including their friends) behave, how many points other people have gained, what the popular "de-stress" places are, etc. We expect that this *social component and gamification (design methodology)* will intrinsically or extrinsically motivate users to engage with the system and support their retention [138].

Overall, these scenarios show how the key components of persuasive, positive computing can be articulated through technical, social, and HCI lenses. They help to make design activities more accessible and give direct, clear insights on system development to researchers, designers, and practitioners.

6. Research Directions

6.1. Evidence-based Design of Positive Computing Systems

Driven by the importance of behavior change and the challenge of achieving it, HCI scholars have explored the opportunity of computational interventions to promote positive behavior [139]. The advent of ubiquitous sensing capabilities and context-aware platforms has allowed people to pervasively log various aspects of their lives resulting in self-discovery, to be supported by persistent and unobtrusive feedback as a form of ambient displays, to interact with an intelligent, relational, and persuasive agent, and to leverage social reinforcements.

In this article, we suggest that researchers build positive computing systems by carefully following the guidelines for designing *evidence-based persuasive systems*, which are established based on behavior principles and systematic evaluations. Systems design may contain various elements, ranging from personal factors (i.e., perceived susceptibility, perceived benefit, self-efficacy, incentives, social pressure, personality, learning history, etc.) to environmental factors (e.g., settings, people, system, etc.), grounded by a health belief model, social cognitive theory, theory of reasoned action, behavioral theory, etc. Through quantitative (e.g., surveys) or ethnographic studies (e.g., one-to-one, focus-group interviews, observations), these factors can be summarized and articulated, and be applied to the design of positive computing systems. Once the system is developed, through a series of user studies with varying conditions (e.g., length, user types, etc.), traditional approaches (e.g., RCT) in the context of HCI research can be employed to measure the effect of a positive computing system. In the field of social computing, to make specific claims about design choices (e.g., encouraging contribution and commitment), researchers attempted to find experimental evidence based on relevant theories of motivation and human behavior in social science [140]. Likewise, through iterative testing, both short and long-term effect should be verified, and new design implications (or evidence-bsed design guidelines) identified from the user studies should be summarized and used for design improvement and scenario development.

6.2. Platform Design of Positive Computing System

As reviewed in Section 4.2, handling high-rate sensor data and supporting extensibility are crucial for a general-purposed platform. OpenDataKIT [71], AWARE [72], and mCerebrum [73] all support high-rate sensor data handling and extensibility. In a practical system aspect, a general-purpose platform further requires considerations on efficient data storage management, power usage, network latency, and system robustness across wearables, phones, and the back-end cloud. For example, selective sampling based on needs (including power-aware sensing) and applications, cloud offloading [73] can improve the overall system lifetime.

Differential privacy management is another sound direction for platform design. Recently, Saleheen et al. [141] suggested the importance of differential privacy, which provides anonymity of any user from a multi-user statistical database, especially for physiological data. It is noteworthy that personal activities can be easily inferred by analyzing body-worn sensors (e.g., respiration (RIP), electrocardiogram (ECG), and accelerometer). These include conversation episodes [142] from respiration data, stress level [143, 41] from ECG data, smoking from respiration and wrist-worn sensors [144, 145, 146], and cocaine use from ECG data [147].

Furthermore, while conducting a large scale experiment, a general-purpose platform may support a component of real-time participant monitoring, which guarantees whether all necessary data from users are gathered correctly and timely. It notifies the users (and optionally, the platform operator) if they mistakenly disable sensors or wireless communication medium. This capability will significantly relieve monitoring costs and user's manual endeavors. Moreover, it will maintain the safety and validity of the experimental data.

6.3. Behavioral Marker Detection

Prior studies on behavioral marker have predominantly focused on understanding manifestations of various well-being problems and discovering novel markers. The key concern in these approaches is the lack of generalizability and scalability. These issues are critical because the algorithms must be deployed to a group of heterogeneous individuals in a scalable way. Simple direct behavioral markers (e.g., detecting problematic physical activities, such as a lack of exercise and aggressive driving) are well defined, and it is relatively easy to design robust detection methods using machine learning. Despite existing diagnostic knowledge bases such as DSM-5 [83], however, in many cases there is a lack of our prior knowledge of detailed manifestations of well-being problems. This lack warrants using certain approaches for detecting inferred behavioral markers. Manifestation of well-being problems varies widely across individuals and groups. Even temporal behavioral changes may exist (e.g., due to major life events). Referring to diagnostic knowledge bases as in Wang et al. [84] may help us to narrow down the search space in the plethora of sensor data.

Prior studies lacked systematic considerations of generalizability and scalability because they were mostly developed and validated with limited datasets. Large-scale data collection would easily solve this limitation; however, it is very challenging and expensive. Alternative approaches would include employing user-feedback-based learning methods, such as reinforcement learning [148] and interactive machine learning [149]. While traditional learning models separate model training and model usage, active learning continually updates the model by adaptively asking users to label data items, whereby probing decisions can be made based on various criteria (e.g., measuring the informativeness of unlabeled data points [150] and estimating the value of asking users [104]). In general, this kind of learning can be extended to so called *life-long learning*, where training is ongoing over a prolonged period [104].

It is interesting to note that this kind of active learning requires some level of user interaction. In the case of positive computing, machine learning is applied to everyday well-being problem. It is very important to consider the fact that regular people have limited skills for engagement. For example, it is almost impossible for a lay person to be directly involved in an optimization process of machine learning models. We can thus consider the principle of *interactive machine learning* [149], which is intended to significantly reduce the need for supervision by machine learning experts. This can be achieved by designing user interfaces to help end users to interactively explore the model space and provide intuitive feedback to drive the machine learning system to intended behaviors. Bellotti and Edwards [151], for example, claimed that such intelligent systems should support *intelligibility* features that "must be able to represent to their users what they know, how they know it, and what they are doing about it." For example, Lim and Dey [152] designed a toolkit to support intelligibility in context-aware applications. Thus, positive computing systems should support intelligibility in their core learning algorithms, including feedback-based personalization and model optimization.

6.4. Opportune Moment Detection

Prior studies have focused on context recognition and user feedback to achieve opportune moment detection. One of the major tasks in context recognition is to achieve high granularity in order to ensure that the *interruptible context* is detected. However, achieving this objective using the sensors in mobile, wearable, and IoT devices is challenging, especially since variations always exist among multiple users, and sensors do not always provide accurate results in required granularity. Inaccurate context recognition may result in inaccurate timing of the interruption, which in turn could result in abandoning the program. To compensate for this issue, the most widely used approach is to utilize user feedback. Note that explicit feedback yields promising data about the user's reaction to the interruption; however, repeatedly asking for feedback may cause disruption and irritation. Implicit feedback, on the other hand, minimizes disruption but might lead to misunderstanding of the user's reaction. One promising research direction is to incorporate user feedback with machine learning to continuously adapt to user's preferences, conditions and environment as in behavioral marker detection.

Another research direction is to *proactively* seek for opportune moments by asking people to make micro-spare time. For example, Kang et al. [153] defined micro spare time as tiny fragments of time with low cognitive loads that frequently occur in our daily lives, such as waiting for an elevator, walking to a different building, waiting for public transportation, and so on. We can automatically identify various types of micro spare time for intervention delivery using machine learning, as prior studies did for learning and parenting purposes [154, 153].

6.5. Device and Modality Selection

As discussed earlier, the problem of device and modality selection is to address the following questions: For a given intervention content (instruction and visual aids), (1) how should we render the content for a given set of devices and their modalities?, and (2) to this end, how should we consider device availability, attention performance, user preference, and acceptability constraints?

One of the major research directions is to understand availability and acceptability in multi-device environments. Users carry mobile and wearable devices, and their usage contexts are very diverse. Analyzing mobility of users and device usage patterns under various circumstances will help elucidate the availability and acceptability of devices. In the case of wearable devices, we can perform a log data analysis to find the unique patterns of an individual's wearing behaviors, as in the work of Jeong et al. [112]. Acceptability of devices and their modalities could be inferred by analyzing interaction log data. Content delivery would also consider social acceptability, because it may disturb other people in the shared spaces. While mobile and wearable devices are primarily for personal use, IoT devices are often installed in the shared spaces, as in the smart home/office environments. In this case, we may also consider dealing with conflicts particularly when a device is shared by multiple people [155].

Another direction is to learn user preferences of devices and modalities. What remains challenging is the fact that user preferences are diverse and context dependent. We can address the user diversity by building personalized models and the context dependency by incorporating context-awareness features. However, the problem of such approaches is lack of available user data (or user feedback). In addition, users may show bias such that every sample should not be equally treated. Among various methods of incrementally learning user preferences, we can use reinforcement learning since it considers heterogeneous rewards. In this case, however, defining a user's state and actions would be quite challenging [148].

6.6. Evaluation of Positive Computing Systems

As an agency to change human behaviors, scholars have studied how the use of mobile, wearable, and IoT devices, which are highly accessible, easy to carry, location-free, and cost-effective, supports individuals to overcome psychological barriers to behavior change. Thus, in the context of positive computing, it is equally important to demonstrate the reliability and validity of positive computing systems and their impact on people via a scientifically validated method. American Psychiatric Association (APA)'s mental health app evaluation model clearly states that evidence (i.e., effectiveness) is the key factor for mental health outcomes, besides safety/privacy, ease of use, and interoperability [156].

In the field of psychotherapy, more than three decades have passed since the need for RCT was brought up in early 90's [157]. Thanks to the accumulation of RCT results, evidence-based treatments for many mental disorders became available to the public (e.g., [158]), and people have better chance to get access to the best practices to resolve their own psychological issues. The same logic applies to the development of computational interventions and applications to promote positive behavior. In order to distribute them to the public, empirical evidences via RCT studies should be accompanied. Similar to the field of psychotherapy, long-term accumulation of objective findings via RCT studies would be able to guide which program and device a person should choose depending on their issues or problems. As pointed out earlier, additional efforts should be made to compensate the weaknesses of RCT, including sensitive outcome measures, appropriate sampling and study duration, fidelity and social validity, etc. Over the past few years, it has been observed that HCI research has adopted RCT for measuring the effectiveness of their systems and products developed [60, 33, 61] and is expected that the application of RCT combined with HCI research will be expanded to many domains.

Beyond evaluating effectiveness of positive computing systems on behavior change using RCT, we may need to advance evaluation procedures to gain a deep understanding of how and why the system is used by its target users in order to identify further design opportunities and challenges. To do this, we could tailor measures to gauge perceptions and thoughts of individuals by adapting empirically validated instruments or inventories. For the long-term effect of the positive computing system use, researchers should continue to see whether the users still exhibit the changed behaviors even after the experiments, and regardless of the results, socio-technical and design opportunities/challenges for the effect should be articulated. This is why the evidence-based design of positive computing systems needs cross-disciplinary research.

6.7. Concerns and Negative Aspects

When designing novel positive computing systems, we suggest researchers and practitioners consider possible concerns and negative aspects, such as technology dependence, abandonment, side effects, privacy, and ethical issues.

First of all, positive computing services often contain various reinforcement, gamification, and social engagement components (e.g., badges, points, and social sharing), which may have inducing and reinforcing features that promote addictive tendencies [91]. Some users may focus too much on such mechanisms without concerning about their behavioral changes (e.g., by cheating achievements). Thus, researchers and practitioners should consider addressing possible negative aspects of such components. For example, they can set reasonable limits on daily achievements and employ anti-cheating and reputation mechanisms.

Prior studies investigated various reasons for the abandonment of self-tracking technologies. We expect that users naturally abandon technologies after goal achievements. However, premature technology abandonment may happen due to the cost of data collection and management, discomfort with information and data accuracy concerns [159]. When developing intelligent positive computing systems with mobile and wearable technologies, designers should carefully address such concerns as well as general usability and user experience issues [160].

Sometimes positive computing services may result in unexpected negative consequences. For example, Facebook was originally designed to fulfill the basic human needs for social connection, but recent studies revealed that it may negatively affect human well-being and life satisfaction [161]. Although it is challenging, researchers and practitioners should carefully investigate possible negative ramifications of positive computing systems on well-being and health.

Privacy issues must be carefully considered in the system design, since systems could collect every single detail about individuals. User data handling must be carefully performed, and minimal data should be collected and utilized. For privacy preservation, the mobile platform may consider implementing localized data processing such that private data do not leave a user's mobile device, or at least unlinkable data are only transferred to the mobile cloud. Furthermore, system designers may adopt privacy preserving data mining techniques; e.g., preserving k-anonymity in location data sharing to avoid attackers from reconstructing invasive location information.

Finally, several ethical issues should be well-reflected as Fogg discussed [139]. Positive computing services can possibly manipulate individuals' behaviors, and system designers should implement such manipulative features solely for promoting positive behaviors. Any kinds of unethical use should be avoided; e.g., embedding an implicit persuasion for product sales. Another ethical issue to consider is accountability, because stakeholders and software agents have responsibilities for computing services and their (intended and unintended) outcomes. As discussed in Bellotti and Edwards' work on context-aware systems design [151], positive computing systems design should consider supporting the accountability of interaction and intelligibility of various context-aware features.

7. Conclusion

We proposed a conceptual framework for intelligent positive computing systems research. The core components include design methodologies, mobile platform design, behavior marker detection, opportune moment detection, device and modality selection, and evaluation methodologies. Given that intelligent positive computing spans a wide range of disciplines, this work provided a tutorial about each component and suggested practical guidelines for system design, development, and evaluation. We demonstrated the conceptual framework by proposing and reviewing several practical service scenarios of addressing college students' well-being problems. Research directions on the core components of positive computing systems research were then illustrated, followed by our brief discussion about concerns and challenges such as technology dependence, abandonment, side effects, privacy, and ethical issues.

As new tools for enabling new directions for positive computing, mobile, wearable, and IoT technologies will greatly change the current landscape of well-being and health-care services. We critically synthesized existing literature in diverse domains and provided holistic perspectives on intelligent positive computing systems research. Our work lays foundations for active collaboration among researchers in diverse domains to design, develop, and evaluate novel intelligent positive computing systems.

8. Acknowledgement

Three of the authors had the privilege of having Prof. Mario Gerla as not only an academic advisor, but a life mentor. We dedicate this work to him; we salute his endless scientific endeavor, and more importantly, passion, kindness, and love.

References

[1] R. A. Calvo, D. Peters, Positive Computing: Technology for Wellbeing and Human Potential, MIT Press, 2014.

[2] M. B. Rosson, J. M. Carroll, The Human-computer Interaction Handbook, L. Erlbaum Associates Inc., 2003, Ch. Scenario-based Design, pp. 1032–1050.

[3] T. Kingkade, College Freshmen Are More Depressed And Alone Than Ever, Huffington Post, https://bit.ly/2JeRdsT (2015).

[4] L. Veiber, 82% of Students Suffer from Stress and Anxiety, The National Student, https://bit.ly/2snnG6r (2017).

[5] R. Misra, M. McKean, College Students' Academic Stress and Its Relation to Their Anxiety, Time Management, and Leisure Satisfaction, American Journal of Health Studies 16 (1) (2000) 41.

[6] G. Halkos, D. Bousinakis, The Effect of Stress and Satisfaction on Productivity, International Journal of Productivity and Performance Management 59 (5) (2010) 415–431.

[7] J. Nakazawa, H. Tokuda, T. Yonezawa, Sensorizer: An Architecture for Regenerating Cyber Physical Data Streams from the Web, in: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, UbiComp/ISWC'15 Adjunct, ACM, 2015, pp. 1599–1606.

[8] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, A. T. Campbell, A Survey of Mobile Phone Sensing, IEEE Communications Magazine 48 (9) (2010) 140–150.

[9] B. Cowley, M. Filetti, K. Lukander, J. Torniainen, A. Henelius, L. Ahonen, O. Barral, I. Kosunen, T. Valtonen, M. Huotilainen, N. Ravaja, G. Jacucci, The Psychophysiology Primer: A Guide to Methods and a Broad Review with a Focus on Human-Computer Interaction, Foundations and Trends in Human-Computer Interaction 9 (3-4) (2016) 151–308.

[10] Samsung SmartThings (2017).
URL https://www.smartthings.com

[11] Sen.se Mother (2017).
URL https://sen.se

[12] M. De Choudhury, M. Gamon, S. Counts, E. Horvitz, Predicting Depression via Social Media, in: AAAI ICWSM '13, 2013, pp. 1–10.

[13] L. Canzian, M. Musolesi, Trajectories of Depression: Unobtrusive Monitoring of Depressive States by Means of Smartphone Mobility Traces Analysis, in: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '15, ACM, 2015, pp. 1293–1304.

[14] S. Saeb, M. Zhang, C. J. Karr, S. M. Schueller, M. E. Corden, K. P. Kording, D. C. Mohr, Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-life Behavior: An Exploratory Study, Journal of Medical Internet Research 17 (7) (2015) e175.

[15] P. Klasnja, W. Pratt, Healthcare in the Pocket: Mapping the Space of Mobile-phone Health Interventions, Journal of Biomedical Informatics 45 (1) (2012) 184–198.

[16] M. A. Adams, J. F. Sallis, G. J. Norman, M. F. Hovell, E. B. Hekler, E. Perata, An Adaptive Physical Activity Intervention for Overweight Adults: A Randomized Controlled Trial, PLOS ONE 8 (12) (2013) e82901.

[17] E. S. Poole, HCI and Mobile Health Interventions, Translational Behavioral Medicine 3 (4) (2013) 402–405.

[18] M. Czerwinski, R. Gilad-Bachrach, Designing Interventions for Health and Wellbeing, in: ACM CHI Computing and Mental Health Workshop, 2016.

[19] R. Yun, P. Scupelli, A. Aziz, V. Loftness, Sustainability in the Workplace: Nine Intervention Techniques for Behavior Change, in: International Conference on Persuasive Technology, Springer, 2013, pp. 253–265.

[20] R. W. Proctor, T. Van Zandt, Human Factors in Simple and Complex Systems, CRC Press, 2018.

[21] N. K. Janz, M. H. Becker, The Health Belief Model: A Decade Later, Health Education Quarterly 11 (1) (1984) 1–47.

[22] A. Bandura, Social Cognitive Theory: An Agentic Perspective, Annual Review of Psychology 52 (1) (2001) 1–26.

[23] M. Fishbein, I. Ajzen, Predicting and Changing Behavior: The Reasoned Action Approach, Taylor & Francis, 2011.

[24] G. Doherty, D. Coyle, J. Sharry, Engagement with Online Mental Health Interventions: An Exploratory Clinical Study of a Treatment for Depression, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2012, pp. 1421–1430.

[25] H. Christensen, K. M. Griffiths, A. F. Jorm, Delivering Interventions for Depression by Using the Internet: Randomised Controlled Trial, BMJ 328 (7434) (2004) 265.

[26] D. Coyle, G. Doherty, M. Matthews, J. Sharry, Computers in Talk-based Mental Health Interventions, Interacting with Computers 19 (4) (2007) 545–562.

[27] R. Calvo, K. Dinakar, R. Picard, J. Torous, The 2nd Symposia on Computing and Mental Health, in: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, ACM, 2017, pp. 3073–3080.

[28] E. Driessen, H. L. Van, F. J. Don, J. Peen, S. Kool, D. Westra, M. Hendriksen, R. A. Schoevers, P. Cuijpers, J. W. Twisk, et al., The Efficacy of Cognitive-Behavioral Therapy and Psychodynamic Therapy in the Outpatient Treatment of Major Depression: A Randomized Clinical Trial, American Journal of Psychiatry 170 (9) (2013) 1041–1050.

[29] D. F. Tolin, Is Cognitive-Behavioral Therapy More Effective Than Other Therapies?: A Meta-Analytic Review, Clinical Psychology Review 30 (6) (2010) 710–720.

[30] R. Crawford, You Are Dangerous to Your Health: The Ideology and Politics of Victim Blaming, International Journal of Health Services 7 (4) (1977) 663–680.

[31] J. F. Sallis, N. Owen, E. Fisher, Ecological Models of Health Behavior, Health Behavior: Theory, Research, and Practice 5 (2015) 43–64.

[32] J. Matthews, K. T. Win, H. Oinas-Kukkonen, M. Freeman, Persuasive Technology in Mobile Applications Promoting Physical Activity: A Systematic Review, Journal of Medical Systems 40 (3) (2016) 72.

[33] S. M. Kelders, R. N. Kok, H. C. Ossebaard, J. E. Van Gemert-Pijnen, Persuasive System Design Does Matter: A Systematic Review of Adherence to Web-based Interventions, Journal of Medical Internet Research 14 (6) (2012) e152.

[34] A. J. Lentferink, H. K. Oldenhuis, M. de Groot, L. Polstra, H. Velthuijsen, J. E. van Gemert-Pijnen, Key Components in eHealth Interventions Combining Self-Tracking and Persuasive eCoaching to Promote a Healthier Lifestyle: A Scoping Review, Journal of Medical Internet Research 19 (8) (2017) e277.

[35] N. Lazzaro, K. Keeker, What's My Method?: A Game Show on Games, in: CHI'04 Extended Abstracts on Human Factors in Computing Systems, ACM, 2004, pp. 1093–1094.

[36] N. Lazzaro, Why We Play Games: Four Keys To More Emotion Without Story, in: Game Developers Conference, 2004.

[37] J. Dyck, D. Pinelle, B. A. Brown, C. Gutwin, Learning from Games: HCI Design Innovations in Entertainment Software, in: Proceedings of the Graphics Interface 2003 Conference, 2003, pp. 237–246.

[38] A. Konrad, V. Bellotti, N. Crenshaw, S. Tucker, L. Nelson, H. Du, P. Pirolli, S. Whittaker, Finding the Adaptive Sweet Spot: Balancing Compliance and Achievement in Automated Stress Reduction, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, ACM, 2015, pp. 3829–3838.

[39] R. Gouveia, E. Karapanos, M. Hassenzahl, How Do We Engage with Activity Trackers?: A Longitudinal Study of Habito, in: Proceedings of the 2015 ACM

International Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2015, pp. 1305–1316.

[40] E. Isaacs, A. Konrad, A. Walendowski, T. Lennig, V. Hollis, S. Whittaker, Echoes from the Past: How Technology Mediated Reflection Improves Well-being, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2013, pp. 1071–1080.

[41] K. Hovsepian, M. al'Absi, E. Ertin, T. Kamarck, M. Nakajima, S. Kumar, cStress: Towards a Gold Standard for Continuous Stress Assessment in the Mobile Environment, in: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '15, ACM, 2015, pp. 493–504.

[42] M. E. Morris, Q. Kathawala, T. K. Leen, E. E. Gorenstein, F. Guilak, M. Labhard, W. Deleeuw, Mobile Therapy: Case Study Evaluations of a Cell Phone Application for Emotional Self-awareness, Journal of Medical Internet Research 12 (2) (2010) e10.

[43] E. A. Locke, G. P. Latham, Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-year Odyssey, American Psychologist 57 (9) (2002) 705.

[44] D. C. Mohr, S. M. Schueller, E. Montague, M. N. Burns, P. Rashidi, The Behavioral Intervention Technology Model: An Integrated Conceptual and Technological Framework for eHealth and mHealth Interventions, Journal of Medical Internet Research 16 (6) (2014) e146.

[45] D. L. Schacter, D. T. Gilbert, D. M. Wegner, Introducing Psychology, Macmillan, 2009.

[46] B. F. Skinner, Science and Human Behavior, Simon and Schuster, 1953.

[47] W. T. O'donohue, L. E. Krasner, Theories of Behavior Therapy: Exploring Behavior Change, American Psychological Association, 1995.

[48] E. B. Hekler, P. Klasnja, J. E. Froehlich, M. P. Buman, Mind the Theoretical Gap: Interpreting, Using, and Developing Behavioral Theory in HCI Research, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2013, pp. 3307–3316.

[49] R. Yun, A. Aziz, P. Scupelli, B. Lasternas, C. Zhang, V. Loftness, Beyond Eco-Feedback: Adding Online Manual and Automated Controls to Promote Workplace Sustainability, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, ACM, 2015, pp. 1989–1992.

[50] R. R. McCrae, P. T. Costa Jr, The Stability of Personality: Observations and Evaluations, Current Directions in Psychological Science 3 (6) (1994) 173–175.

[51] D. A. Cobb-Clark, S. Schurer, The Stability of Big-five Personality Traits, Economics Letters 115 (1) (2012) 11–15.

[52] J. G. March, L. S. Sproull, M. Tamuz, Learning from Samples of One or Fewer, Organization Science 2 (1) (1991) 1–13.

[53] S. G. Tibbetts, Perinatal and Developmental Determinants of Early Onset of Offending: A Biosocial Approach for Explaining the Two Peaks of Early Antisocial Behavior, The Development of Persistent Criminality (2009) 179–201.

[54] M. Forsman, Psychopathic Personality in Adolescence: Genetic and Environmental Influences, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 2009.

[55] J. O. Cooper, T. E. Heron, W. L. Heward, Applied Behavior Analysis, Pearson/Merrill-Prentice Hall Upper Saddle River, NJ, 2009.

[56] P. Garneau, Emergence: Making Games Deeper (2003).
URL `https://bit.ly/2LenyNs`

[57] M. J. Apter, J. H. Kerr, Adult Play: A Reversal Theory Approach, Swets & Zeitlinger, 1991.

[58] R. Hunicke, M. LeBlanc, R. Zubek, MDA: A Formal Approach to Game Design and Game Research, in: Proceedings of the AAAI Workshop on Challenges in Game AI, AAAI Press San Jose, CA, 2004, pp. 1–5.

[59] A. Direito, E. Carraça, J. Rawstorn, R. Whittaker, R. Maddison, MHealth Technologies to Influence Physical Activity and Sedentary Behaviors: Behavior Change Techniques, Systematic Review and Meta-analysis of Randomized Controlled Trials, Annals of Behavioral Medicine 51 (2) (2016) 226–239.

[60] J. Stephens, J. Allen, Mobile Phone Interventions to Increase Physical Activity and Reduce Weight: A Systematic Review, The Journal of Cardiovascular Nursing 28 (4) (2013) 320.

[61] F. Bert, M. Giacometti, M. R. Gualano, R. Siliquini, Smartphones and Health Promotion: A Review of the Evidence, Journal of Medical Systems 38 (1) (2014) 9995.

[62] J. Jardine, J. Fisher, B. Carrick, Apple's ResearchKit: Smart Data Collection for the Smartphone Era?, Journal of the Royal Society of Medicine 108 (8) (2015) 294–296.

[63] Apple's ResearchKit (2017).
URL `http://www.researchkit.org/`

[64] Google FIT (2017).
URL `https://www.google.com/fit/`

[65] Micorosft's HealthVault (2017).
URL `https://international.healthvault.com/`

[66] M. Raento, A. Oulasvirta, R. Petit, H. Toivonen, ContextPhone: A Prototyping Platform for Context-aware Mobile Applications, IEEE Pervasive Computing 4 (2) (2005) 51–59.

[67] H. Tangmunarunkit, C. K. Hsieh, B. Longstaff, S. Nolen, J. Jenkins, C. Ketcham, J. Selsky, F. Alquaddoomi, D. George, J. Kang, Z. Khalapyan, J. Ooms, N. Ra-

manathan, D. Estrin, Ohmage: A General and Extensible End-to-End Participatory Sensing Platform, ACM Transactions on Intelligent Systems and Technology (TIST) 6 (3) (2015) 38:1–38:21.

[68] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, A. T. Campbell, Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of the CenceMe Application, in: Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems, SenSys '08, ACM, 2008, pp. 337–350.

[69] S. Carter, J. Mankoff, J. Heer, Momento: Support for Situated Ubicomp Experimentation, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07, ACM, 2007, pp. 125–134.

[70] M. J. van Sinderen, A. T. van Halteren, M. Wegdam, H. B. Meeuwissen, E. H. Eertink, Supporting Context-aware Mobile Applications: An Infrastructure Approach, IEEE Communications Magazine 44 (9) (2006) 96–104.

[71] OpenDataKit (2012).
URL `https://opendatakit.org/`

[72] D. Ferreira, V. Kostakos, A. K. Dey, AWARE: mobile context instrumentation framework, Frontiers in ICT 2 (6) (2015) 1–9.

[73] S. M. Hossain, T. Hnat, N. Saleheen, N. J. Nasrin, J. Noor, B.-J. Ho, T. Condie, M. Srivastava, S. Kumar, N. J. Nas, mCerebrum: A Mobile Sensing Software Platform for Development and Validation of Digital Biomarkers and Interventions, in: Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems, SenSys '17, ACM, Delft, The Netherlands, 2017.

[74] MD2K.org (2012).
URL `https://github.com/MD2Korg/`

[75] S. H. Jain, B. W. Powers, J. B. Hawkins, J. S. Brownstein, The Digital Phenotype, Nature Biotechnology 33 (5) (2015) 462.

[76] G. M. Harari, N. D. Lane, R. Wang, B. S. Crosier, A. T. Campbell, S. D. Gosling, Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges, Perspectives on Psychological Science 11 (6) (2016) 838–854.

[77] T. R. Insel, Digital Phenotyping: Technology for a New Science of Behavior, JAMA 318 (13) (2017) 1215–1216.

[78] D. C. Mohr, M. Zhang, S. M. Schueller, Personal Sensing: Understanding Mental Health using Ubiquitous Sensors and Machine Learning, Annual Review of Clinical Psychology 13 (2017) 23–47.

[79] V. Pejovic, A. Mehrotra, M. Musolesi, Anticipatory Mobile Digital Health: Towards Personalized Proactive Therapies and Prevention Strategies, in: Anticipation and Medicine, Springer, 2017, pp. 253–267.

[80] I. Nahum-Shani, S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari, S. A. Murphy, Just-In-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support, Annals of Behavioral Medicine 52 (6) (2006) 446–462.

[81] A. K. Dey, Understanding and Using Context, Personal and Ubiquitous Computing 5 (1) (2001) 4–7.

[82] A. Schmidt, M. Beigl, H.-W. Gellersen, There is More to Context Than Location, Computers & Graphics 23 (6) (1999) 893–901.

[83] Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (2013).
URL `https://www.psychiatry.org/psychiatrists/practice/dsm`

[84] R. Wang, W. Wang, A. daSilva, J. F. Huckins, W. M. Kelley, T. F. Heatherton, A. T. Campbell, Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2 (1) (2018) 43:1–43:26.

[85] S. Van Dantzig, G. Geleijnse, A. T. van Halteren, Toward a Persuasive Mobile Application to Reduce Sedentary Behavior, Personal and Ubiquitous Computing 17 (6) (2013) 1237–1246.

[86] S. Abdullah, M. Matthews, E. L. Murnane, G. Gay, T. Choudhury, Towards Circadian Computing: Early to Bed and Early to Rise Makes Some of Us Unhealthy and Sleep Deprived, in: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '14, ACM, 2014, pp. 673–684.

[87] E. Thomaz, I. Essa, G. D. Abowd, A Practical Approach for Recognizing Eating Moments with Wrist-mounted Inertial Sensing, in: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '15, ACM, 2015, pp. 1029–1040.

[88] R. Alam, J. Gong, M. Hanson, A. Bankole, M. Anderson, T. Smith-Jackson, J. Lach, Motion Biomarkers for Early Detection of Dementia-related Agitation, in: Proceedings of the 1st Workshop on Digital Biomarkers, ACM, 2017, pp. 15–20.

[89] Y. Huang, J. Gong, M. Rucker, P. Chow, K. Fua, M. S. Gerber, B. Teachman, L. E. Barnes, Discovery of Behavioral Markers of Social Anxiety from Smartphone Sensor Data, in: Proceedings of the 1st Workshop on Digital Biomarkers, ACM, 2017, pp. 9–14.

[90] F. Gravenhorst, A. Muaremi, J. Bardram, A. Grünerbl, O. Mayora, G. Wurzer, M. Frost, V. Osmani, B. Arnrich, P. Lukowicz, G. Tröster, Mobile Phones as Medical Devices in Mental Disorder Treatment: An Overview, Personal and Ubiquitous Computing 19 (2) (2015) 335–353.

[91] U. Lee, J. Lee, M. Ko, C. Lee, Y. Kim, S. Yang, K. Yatani, G. Gweon, K.-M. Chung, J. Song, Hooked on Smartphones: An Exploratory Study on Smartphone Overuse among College Students, in: Proceedings of the 32nd Aannual ACM Conference on Human Factors in Computing Systems, CHI '14, ACM, 2014, pp. 2327–2336.

[92] R. Wang, M. S. H. Aung, S. Abdullah, R. Brian, A. T. Campbell, T. Choudhury, M. Hauser, J. Kane, M. Merrill, E. A. Scherer, V. W. S. Tseng, D. Ben-Zeev, CrossCheck: Toward Passive Sensing and Detection of Mental Health Changes in People with Schizophrenia, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16, ACM, 2016, pp. 886–897.

[93] K. Kroenke, T. W. Strine, R. L. Spitzer, J. B. Williams, J. T. Berry, A. H. Mokdad, The PHQ-8 as a Measure of Current Depression in the General Population, Journal of Affective Disorders 114 (1–3) (2009) 163173.

[94] M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, M. Kumar, Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI '16, ACM, 2016, pp. 2098–2110.

[95] G. Mark, V. M. Gonzalez, J. Harris, No Task Left Behind?: Examining the Nature of Fragmented Work, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05, ACM, 2005, pp. 321–330.

[96] S. Hudson, J. Fogarty, C. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. Lee, J. Yang, Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '03, ACM, 2003, pp. 257–264.

[97] J. Ho, S. S. Intille, 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, CHI '05, ACM, 2005, pp. 909–918.

[98] S. T. Iqbal, B. P. Bailey, Oasis: A Framework for Linking Notification Delivery to the Perceptual Structure of Goal-directed Tasks, ACM Transactions on Computer-Human Interaction (TOCHI) 17 (4) (2010) 15:1–15:28.

[99] G. Mark, D. Gudith, U. Klocke, The Cost of Interrupted Work: More Speed and Stress, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08, ACM, 2008, pp. 107–110.

[100] P. D. Adamczyk, B. P. Bailey, If Not Now, When?: The Effects of Interruption at Different Moments Within Task Execution, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04, ACM, 2004, pp. 271–278.

[101] M. Choy, D. Kim, J.-G. Lee, H. Kim, H. Motoda, Looking Back on the Current Day: Interruptibility Prediction Using Daily Behavioral Features, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16, ACM, 2016, pp. 1004–1015.

[102] S. A. Cambo, D. Avrahami, M. L. Lee, BreakSense: Combining Physiological and Location Sensing to Promote Mobility During Work-Breaks, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17, ACM, 2017, pp. 3595–3607.

[103] V. Pejovic, M. Musolesi, InterruptMe: Designing Intelligent Prompting Mechanisms for Pervasive Applications, in: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '14, ACM, 2014, pp. 897–908.

[104] A. Kapoor, E. Horvitz, Principles of Lifelong Learning for Predictive User Modeling, in: User Modeling, Springer, 2007.

[105] I. Kim, G. Jung, H. Jung, M. Ko, U. Lee, Let's FOCUS: Mitigating Mobile Phone Use in College Classrooms, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1 (3) (2017) 63:1–63:29.

[106] C. Park, J. Lim, J. Kim, S.-J. Lee, D. Lee, Don't Bother Me. I'm Socializing!: A Breakpoint-Based Smartphone Notification System, in: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17, ACM, 2017, pp. 541–554.

[107] A. Ferscha, Attention, Please!, IEEE Pervasive Computing 13 (1) (2014) 48–54.

[108] D. Warnock, M. McGee-Lennon, S. Brewster, The Role of Modality in Notification Performance, in: IFIP Conference on Human-Computer Interaction, Springer, 2011, pp. 572–588.

[109] C. Elting, J. Zwickel, R. Malaka, Device-dependant Modality Selection for User-Interfaces: An Empirical Study, in: Proceedings of the 7th International Conference on Intelligent User Interfaces, IUI '02, ACM, 2002, pp. 55–62.

[110] A. S. Shirazi, N. Henze, Assessment of Notifications on Smartwatches, in: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, MobileHCI '15, ACM, 2015, pp. 1111–1116.

[111] D. Weber, A. Voit, P. Kratzer, N. Henze, In-situ Investigation of Notifications in Multi-device Environments, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16, ACM, 2016, pp. 1259–1264.

[112] H. Jeong, H. Kim, R. Kim, U. Lee, Y. Jeong, Smartwatch Wearing Behavior Analysis: A Longitudinal Study, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1 (3) (2017) 60.

[113] D. P. Siewiorek, A. Smailagic, J. Furukawa, A. Krause, N. Moraveji, K. Reiger, J. Shaffer, F. L. Wong, SenSay: A Context-Aware Mobile Phone, in: Proceedings of the Seventh IEEE International Symposium on Wearable Computers, ISWC '03, 2003, pp. 248–249.

[114] J. Lee, E. Walker, W. Burleson, M. Kay, M. Buman, E. B. Hekler, Self-Experimentation for Behavior Change: Design and Formative Evaluation of Two Approaches, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17, ACM, 2017, pp. 6837–6849.

[115] J. Kim, C. Cho, U. Lee, Technology Supported Behavior Restriction for Mitigating Self-Interruptions in Multi-device Environments, Proceedings of the ACM

on Interactive, Mobile, Wearable and Ubiquitous Technologies 1 (3) (2017) 64:1–64:21.

[116] G. F. Mateo, E. Granado-Font, C. Ferré-Grau, X. Montaña-Carreras, Mobile Phone Apps to Promote Weight Loss and Increase Physical Activity: A Systematic Review and Meta-analysis, Journal of Medical Internet Research 17 (11) (2015) e253.

[117] J. Hamari, J. Koivisto, T. Pakkanen, Do Persuasive Technologies Persuade? - A Review of Empirical Studies, in: International Conference on Persuasive Technology, Springer, 2014, pp. 118–136.

[118] H. Oinas-Kukkonen, A Foundation for the Study of Behavior Change Support Systems, Personal and Ubiquitous Computing 17 (6) (2013) 1223–1235.

[119] M. H. Van den Berg, J. W. Schoones, T. P. V. Vlieland, Internet-based Physical Activity Interventions: A Systematic Review of the Literature, Journal of Medical Internet Research 9 (3) (2007) e26.

[120] M. Rabbi, A. Pfammatter, M. Zhang, B. Spring, T. Choudhury, Automated Personalized Feedback for Physical Activity and Dietary Behavior Change with Mobile Phones: A Randomized Controlled Trial on Adults, JMIR mHealth and uHealth 3 (2) (2015) e42.

[121] M. A. Raskind, Pharmacology of Sleep and PTSD: Prazosin-An Alpha-1 Adrenoreceptor Antagonist Approach to Post-traumatic Stress Disorder Pharmacotherapy, in: Sleep and Combat-Related Post Traumatic Stress Disorder, Springer, 2018, pp. 349–359.

[122] E. J. Aguiar, P. J. Morgan, C. E. Collins, R. C. Plotnikoff, M. D. Young, R. Callister, Efficacy of the Type 2 Diabetes Prevention using Lifestyle Education Program RCT, American Journal of Preventive Medicine 50 (3) (2016) 353–364.

[123] M. Van Reijen, I. Vriend, W. Mechelen, E. Verhagen, Preventing Recurrent Ankle Sprains: Is the Use of an App More Cost-effective Than a Printed Booklet?

Results of a RCT, Scandinavian Journal of Medicine & Science in Sports 28 (2) (2017) 641–648.

[124] P. K. MacDonald, Useful Fiction or Miracle Maker: The Competing Epistemological Foundations of Rational Choice Theory, American Political Science Review 97 (4) (2003) 551–565.

[125] W. Pearce, S. Raman, The New Randomised Controlled Trials (RCT) Movement in Public Policy: Challenges of Epistemic Governance, Policy Sciences 47 (4) (2014) 387–402.

[126] B. A. Morrongiello, S. Hou, M. Bell, K. Walton, A. J. Filion, J. Haines, Supervising for Home Safety Program: A Randomized Controlled Trial (RCT) Testing Community-Based Group Delivery, Journal of Pediatric Psychology 42 (7) (2016) 768–778.

[127] B. S. Fjeldsoe, A. L. Marshall, Y. D. Miller, Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service, American Journal of Preventive Medicine 36 (2) (2009) 165–173.

[128] J. Concato, N. Shah, R. I. Horwitz, Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs, New England Journal of Medicine 342 (25) (2000) 1887–1892.

[129] P. Klasnja, S. Consolvo, W. Pratt, How to Evaluate Technologies for Health Behavior Change in HCI Research, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2011, pp. 3063–3072.

[130] R. H. Howland, Limitations of Evidence in the Practice of Evidence-based Medicine, Journal of Psychosocial Nursing & Mental Health Services 45 (11) (2007) 13.

[131] P. Lally, C. H. M. van Jaarsveld, H. W. W. Potts, J. Wardle, How Are Habits Formed: Modelling Habit Formation in the Real World, European Journal of Social Psychology 40 (6) (2010) 998–1009.

[132] R. Hurling, M. Catt, M. De Boni, B. W. Fairley, T. Hurst, P. Murray, A. Richardson, J. S. Sodhi, Using Internet and Mobile Phone Technology to Deliver an Automated Physical Activity Program: Randomized Controlled Trial, Journal of Medical Internet Research 9 (2) (2007) e7.

[133] Y.-H. Kim, J. H. Jeon, E. K. Choe, B. Lee, K. Kim, J. Seo, TimeAware: Leveraging Framing Effects to Enhance Personal Productivity, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM, 2016, pp. 272–283.

[134] R. Beiter, R. Nash, M. McCrady, D. Rhoades, M. Linscomb, M. Clarahan, S. Sammut, The Prevalence and Correlates of Depression, Anxiety, and Stress in a Sample of College Students, Journal of Affective Disorders 173 (2015) 90–96.

[135] S. Jaschik, The Other Mental Health Crisis, Inside Highered, https://bit.ly/1DhJDjj (2015).

[136] E. K. Choe, S. Abdullah, M. Rabbi, E. Thomaz, D. A. Epstein, F. Cordeiro, M. Kay, G. D. Abowd, T. Choudhury, J. Fogarty, B. Lee, M. Matthews, J. A. Kientz, Semi-Automated Tracking: A Balanced Approach for Self-monitoring Applications, IEEE Pervasive Computing 16 (1) (2017) 74–84.

[137] M. Ko, S. Choi, K. Yatani, U. Lee, Lock N' LoL: Group-based Limiting Assistance App to Mitigate Smartphone Distractions in Group Activities, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI '16, ACM, 2016, pp. 998–1010.

[138] R. M. Ryan, E. L. Deci, Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-being, American Psychologist 55 (1) (2000) 68.

[139] B. J. Fogg, Persuasive Technology: Using Computers to Change What We Think and Do, Ubiquity 2002 (12).

[140] R. E. Kraut, P. Resnick, Building Successful Online Communities: Evidence-Based Social Design, MIT Press, 2012.

[141] N. Saleheen, S. Chakraborty, N. Ali, M. M. Rahman, S. M. Hossain, R. Bari, E. Buder, M. Srivastava, S. Kumar, mSieve: Differential Behavioral Privacy in Time Series of Mobile Sensor Data, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16, ACM, 2016, pp. 706–717.

[142] M. M. Rahman, A. A. Ali, K. Plarre, M. al'Absi, E. Ertin, S. Kumar, mConverse: Inferring Conversation Episodes from Respiratory Measurements Collected in the Field, in: Proceedings of the 2nd Conference on Wireless Health, WH '11, ACM, 2011, pp. 10:1–10:10.

[143] K. Plarre, A. Raij, S. M. Hossain, A. A. Ali, M. Nakajima, M. Al'absi, E. Ertin, T. Kamarck, S. Kumar, M. Scott, D. Siewiorek, A. Smailagic, L. E. Wittmers, Continuous Inference of Psychological Stress from Sensory Measurements Collected in the Natural Environment, in: Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN '11, 2011, pp. 97–108.

[144] A. Parate, M.-C. Chiu, C. Chadowitz, D. Ganesan, E. Kalogerakis, RisQ: Recognizing Smoking Gestures with Inertial Sensors on a Wristband, in: Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys '14, ACM, 2014, pp. 149–161.

[145] A. A. Ali, S. M. Hossain, K. Hovsepian, M. M. Rahman, K. Plarre, S. Kumar, mPuff: Automated Detection of Cigarette Smoking Puffs from Respiration Measurements, in: Proceedings of the 2012 ACM/IEEE 11th International Conference on Information Processing in Sensor Networks, IPSN '12, 2012, pp. 269–280.

[146] N. Saleheen, A. A. Ali, S. M. Hossain, H. Sarker, S. Chatterjee, B. Marlin, E. Ertin, M. al'Absi, S. Kumar, puffMarker: A Multi-sensor Approach for Pinpointing the Timing of First Lapse in Smoking Cessation, in: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '15, ACM, 2015, pp. 999–1010.

[147] S. M. Hossain, A. A. Ali, M. M. Rahman, E. Ertin, D. Epstein, A. Kennedy, K. Preston, A. Umbricht, Y. Chen, S. Kumar, Identifying Drug (Cocaine) Intake Events from Acute Physiological Response in the Presence of Free-living Physical Activity, in: Proceedings of the 13th International Symposium on Information Processing in Sensor Networks, IPSN '14, 2014, pp. 71–82.

[148] M. Hoogendoorn, B. Funk, Machine Learning for the Quantified Self: On the Art of Learning from Sensory Data, Springer, 2017.

[149] S. Amershi, M. Cakmak, W. B. Knox, T. Kulesza, Power to the People: The Role of Humans in Interactive Machine Learning, AI Magazine 35 (4) (2014) 105–120.

[150] N. Lawrence, M. Seeger, R. Herbrich, Fast Sparse Gaussian Process Methods: The Informative Vector Machine, in: Proceedings of the 15th International Conference on Neural Information Processing Systems, NIPS'02, 2002, pp. 625–632.

[151] V. Bellotti, K. Edwards, Intelligibility and Accountability: Human Considerations in Context-aware Systems, Human-Computer Interaction 16 (2) (2001) 193–212.

[152] B. Y. Lim, A. K. Dey, Toolkit to Support Intelligibility in Context-aware Applications, in: Proceedings of the 12th ACM International Conference on Ubiquitous Computing, UbiComp '10, ACM, 2010, pp. 13–22.

[153] B. Kang, C. Min, W. Kim, I. Hwang, C. Park, S. Lee, S.-J. Lee, J. Song, Zaturi: We Put Together the 25th Hour for You. Create a Book for Your Baby, in: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17, ACM, 2017, pp. 1850–1863.

[154] C. J. Cai, P. J. Guo, J. R. Glass, R. C. Miller, Wait-Learning: Leveraging Wait Time for Second Language Education, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, ACM, 2015, pp. 3701–3710.

[155] C. Shin, A. K. Dey, W. Woo, Mixed-initiative Conflict Resolution for Context-aware Applications, in: Proceedings of the 10th International Conference on Ubiquitous Computing, UbiComp '08, ACM, 2008, pp. 262–271.

[156] N. Zagorski, APA to Provide Framework to Evaluate Mobile Health Apps, Clinical and Research News, 2017.

[157] C. R. Blease, S. O. Lilienfeld, J. M. Kelley, Evidence-based Practice and Psychological Treatments: The Imperatives of Informed Consent, Frontiers in Psychology 7 (2016) 1170.

[158] APA Presidential Task Force on Evidence-Based Practice (2006), Evidence-based Practice in Psychology, American Psychologist 61 (4) (2006) 271–285.

[159] D. A. Epstein, M. Caraway, C. Johnston, A. Ping, J. Fogarty, S. A. Munson, Beyond Abandonment to Next Steps: Understanding and Designing for Life After Personal Informatics Tool Use, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI '16, ACM, 2016, pp. 1109–1113.

[160] J. Oh, U. Lee, Exploring UX Issues in Quantified Self Technologies, in: Proceedings of the Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU), 2015, pp. 53–59.

[161] E. Kross, P. Verduyn, E. Demiralp, J. Park, D. S. Lee, N. Lin, H. Shablack, J. Jonides, O. Ybarra, Facebook Use Predicts Declines in Subjective Well-Being in Young Adults, PLOS ONE 8 (8) (2013) 1–6.