Realizing Effective Behavioral Management of Health

By Eric B. Hekler, Predrag Klasnja, Vicente Traver, and Monique Hendriks

The past two centuries have shown radical improvements in health and longevity, with hygiene as the key contributor to this trend in the 19th century and antibiotics and vaccinations in the 20th century. With most infectious diseases largely at bay in the developed world, the greatest contributors to suboptimal health today are largely behavioral. For example, there are three behavioral risk factors—tobacco use, poor diet, and inactivity—that contribute to four chronic diseases: heart disease, type 2 diabetes, lung disease, and some cancers. Together, these behaviors account for more than 50% of preventable deaths (see the Web site 3four50.com). While medical advances will surely continue, it is likely that the next great advancements in health in the 21st century will occur via more effective behavior management.

The study of behavior and, in particular, the development of interventions to promote healthier lifestyles, has traditionally been highly resource constrained compared to the complexity of the problem. The development of behavioral interventions was based on three broad methods: 1) in-lab studies, which are constrained in their ability to glean insights about real-world phenomena; 2) epidemiologic studies, constrained by the use of infrequent, often self-reported measures; and 3) randomized clinical trials (RCTs) for evaluating behavioral interventions, which require extensive resources to conduct effectively, but are limited in their ability to provide idiographic (i.e., case-by-case) information about the utility of an intervention.

The Metamorphosis of Behavioral Science Methods

Just as the microscope opened new opportunities for developments in cell biology, new study methods, sensors, and analytics are opening the door for a revolution in the study of human behavior. We can now measure behavior using sensors in clothes, pots and pans in the kitchen, chips in products people buy in the supermarket, mobile phones, and more. In addition, we can now deliver health interventions not only face to face but also via the embedded technologies in our daily lives. All these sensor streams and interactions can be documented and analyzed to facilitate new insights on behavior.

From deriving idiographic causal conclusions to processing and organizing data for use with new data-analytic techniques to the development and refinement of replicable and potent interventions to change behavior, the classical tool kit of behavioral science is not up to the task. Thankfully, various techniques from engineering, computer science, and other related fields, along with recently rediscovered experimental designs from behavioral science, provide starting points for improving behavioral interventions. In this article, we examine how these new methods can be used to support a more personalized understanding of behavior and behavior change and, by extension, improve behavioral management of health.

Digital Object Identifier 10.1109/MPUL.2013.2271681
Date of publication: 17 September 2013

Studying the Behavioral Change Process

Health behavior interventions are intended to intervene in the behavioral status quo and shift behavior toward more healthful configurations. However, a behavioral intervention, whether technological or otherwise, rarely changes only the target behavior. As the behavior changes—for example, as a person tries to be more physically active—a host of other factors in the person's life change as well. The person might rearrange her daily routine (when she wakes up, how she gets to work, etc.), she might start to feel better and think of herself as a fit person, and she might start to eat differently, even though diet was not the target behavior. In addition, not all of these changes will happen for every person who undergoes the same intervention, and those changes that happen might occur at different rates for each individual. These idiosyncrasies require a more idiographic approach to understanding the behavioral change process.

How these various changes are connected, how they are influenced by the individual differences in people undergoing the intervention, and how these factors influence the target behavior is a process we do not clearly understand. As we collect detailed data about the individual's activities, states, and context over time, we can create dynamic models of the change process itself, enabling us to better understand factors that contribute to successful long-term change. The efforts to create such detailed models of behavioral change are in their infancy, but they promise to significantly improve our understanding of behavioral dynamics.

Another important issue related to the development of behavioral interventions is the following question: How do we know if an intervention is working? This question is more complicated than it first appears. Many behavioral interventions, and the vast majority of behavior change technologies, are highly complex. For example, the Diabetes Prevention Program (DPP) is a well-researched intervention for promoting behavior change and weight loss to prevent diabetes. The DPP incorporates multiple face-to-face sessions that cover a wide range of topics such as goal setting for diet and weight loss, stress reduction, and strategies for counteracting negative thinking. Beyond different behavior change techniques, interventions often use various delivery methods (e.g., text messages, Web sites, smartphone applications, and serious games) that are intended to be used at different frequencies and in different contexts. Furthermore, the same type of functionality can be designed in different ways. For instance, if an intervention involves food tracking, the tool could be a simple paper notebook, a Web site, or a mobile-phone application. In addition, the technological implementations can vary a great deal, with one version using a database of food options to choose from, another using a barcode scanner, and a third a more lightweight surveying technique such as categorizing foods as "healthy," "okay," or "unhealthy." How exactly the intervention is designed can significantly affect how it is perceived and used by people, affecting, in turn, its efficacy. A key methodological challenge for such complex interventions is to understand not only if an intervention is working but also how and why it is or is not working and for whom.

There is increasing interest among those building behavior change technologies to build technologies that support personal agency and avoid unintended coercion.

Recent technological and methodological developments allow us to begin to understand the working of behavior change interventions in several ways. First, especially for theoretically guided interventions, it is important to understand which mechanisms are driving behavior changes if behavior change occurs. For example, an intervention might be intended to increase goal commitment, but depending on how this functionality is designed, the intervention might or might not actually impact goals. Using sensors and techniques such as phone-based ecological momentary assessment is a relatively simple way to gather context-dependent behavioral and self-reported data, which can, for some variables at least, enable researchers to test directly whether the postulated theoretical mechanisms of change are actually being activated. Similarly, well-designed qualitative methods (e.g., interviews) can provide indirect evidence that the intervention is activating the proposed mechanisms of change. Together such methods can help to establish an intervention's theoretical fidelity [1].

Second, no intervention can be effective if it is not used enough. Detailed usage logs and qualitative feedback both during the testing of the intervention and afterwards can help researchers get an accurate picture of how and why the intervention was used or unused. This information can, in turn, help researchers avoid type 3 error, concluding that an intervention is ineffective when in fact its efficacy was never tested because the intervention was not designed well enough to support effective usage and testing [1].

Third, for complex interventions, it is important to understand how different components of the intervention contribute to its overall effectiveness, both individually and in combination. There are a number of challenges that arise from this complexity including appropriate nomenclature for systematic comparison and appropriate experimental designs for supporting component-based versus intervention-package conclusions. Recent methodological and theoretical developments in behavioral science are beginning to address these challenges head on [1].

Finally, evaluations of behavioral interventions are ultimately interested in whether those interventions changed the behavior that they targeted. RCTs have long been the gold standard for answering this question, but the need to more deeply understand how the interventions are working, particularly at an idiographic level, has created the need for other experimental designs. Such alternative experimental designs are increasingly more viable because of the greater ease with which data related to behavior and context can be gathered and aggregated and the ease with which interventions can be delivered in standardized ways via technology.

In summary, behavior change is an interaction of multiple factors and occurs differently from individual to individual. We need new ways to understand the interplay between these factors to determine which factors interact and how and what the effect is on different types of individuals. A key first step

FIGURE 1 The Jawbone UP, a fashionable physical activity and sleep sensor, allows individuals to monitor their habits. (Image courtesy of Jawbone.)

toward studying this complexity is the measurement of factors of interest, which has been enabled by new sensing and measurement techniques.

Sensors and Semiautomated Detection

The proliferation of unobtrusive methods for collecting temporally dense data about people's behavior, states, and contexts in which these behaviors take place is radically reducing the resources required to study behavior in the wild. Rather than just assessing behaviors of interest once or every few months, as is standard in traditional behavioral interventions and RCTs, we can now have a near-real-time picture of what an individual is doing and in what circumstances those actions take place. Such monitoring is made possible by the following four types of data.

Wearable and Integrated Sensors

Inexpensive commercial devices, such as the FitBit, the Phillips DirectLife, and the Jawbone UP (Figure 1), are enabling unobtrusive detection of health-related behaviors, such as physical activity, sedentary behavior, and sleep. Devices such as the BodyMedia allow monitoring of a range of physiological variables, such as heart rate and skin conductance, providing information about an individual's states such as stress, and important biomarkers, such as blood pressure and heart-rate variability [2]. In addition, sensors are increasingly being integrated into the products we use every day, enabling monitoring of highly specific health-related activities. One such device is HapiFork, which is an eating utensil that can track the pace of eating and provide feedback (see http://www.hapilabs.com). Collectively, such sensors can provide us with a rich picture about individuals' health behaviors and the contexts in which they take place.

Sensing devices can collect useful data only if they are being used, however. For this reason, many new devices, such as the Jawbone UP, emphasize not only the quality of the sensors (e.g., accelerometers) and data analytics but also the design. The goal is to create wearable technology that is attractive, personalizable, and can fit into the fabric of a person's everyday life, thus encouraging use and increasing the technology's utility. Furthermore, there is increasingly more work, including research that we and our colleagues are conducting with the Jawbone and other devices, to validate these devices to ensure the highest quality of data that these sensors produce.

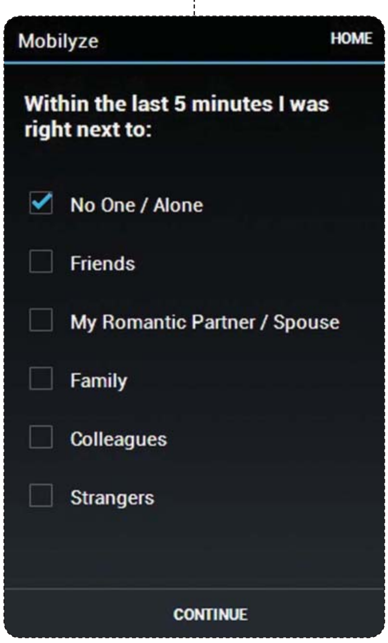

FIGURE 2 The Mobilyze! application is a tool for depression management [5]. (Image courtesy of David Mohr, Northwestern University.)

Mobile-Phone-Based Sensing

In addition to the stand-alone sensors, researchers are increasingly using mobile phones to sense users' states and activities. Modern mobile phones contain many sensors, such as accelerometers, a global positioning system (GPS), gyros, light sensors, and microphones, which can be used to sense both users' activities and the environment in which these activities take place. Beyond providing a secondary means of capturing activities such as physical activity, mobile phones are capable of detecting some unique states and activities as well.

For instance, recent research has shown that it is possible to use audio from phone conversations to detect stress and mood transitions, which could be highly valuable for behavioral problems such as treating bipolar disorder or depression [3]. The phone-based sensors can also be used to trigger differential responses within an intervention. For example, the smartphone application Mobilyze! (Figure 2) has individuals label various locations to enable context-dependent feedback for treating depression (for a review on this and other interventions like it, see [4]).

Similarly, the BeWell app uses sensors in the phone to continuously track three key health behaviors—sleep, physical activity, and social interactions—without requiring any user input. On the basis of sensor data, BeWell provides users feedback about these behaviors in the form of an animated aquatic ecosystem rendered smartphone's wallpaper. In summary, this richness in sensors allow phones to detect where the users are, what they are doing, whether they are alone or with other people (and whom they are with), and even how polluted the air is at the users' location. As new sensors get added to phones, the amount of information that can be detected will only increase.

Digital Footprints

A great deal of information about users' activities and states can be inferred from the various digital traces that individuals

IEEE PULSE | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags THE WORLD'S NEWSSTAND®

leave in their day-to-day lives. Calendars, e-mail, and the use of social-networking applications such as Facebook are all sources of potentially useful information for modeling health-related behavior. Monitoring of such systems can help us determine individuals' social interactions, patterns of free time and business throughout the day, and potentially even individuals' energy level variations.

Low-Effort Self-Reporting

Finally, not all variables that might be influencing an individual's health-related behaviors can be detected completely automatically. For some, such as attitudes, health goals, or self-efficacy, self-reporting is still needed. Here too, new technologies have made frequent collection of data much easier. Self-report tool kits, tools for creation of lightweight questionnaires that can be triggered and answered directly on the phone, are available for most smartphone platforms.

One example is Google's Personal Analytics Companion (Paco, pacoapp.com), a Web and smartphone application that enables users to easily create custom questionnaires to understand their own behavior. Tool kits like this enable questions to be triggered not only at particular times or at random intervals but based on sensor data as well, making it possible, for instance, to trigger a questionnaire when an individual is at a particular location (e.g., a fast-food restaurant) or right after a run. While long-term use of such questionnaires traditionally suffered from large amounts of missing data, these new methods are increasingly aiding in better identifying the best times and contexts to query for a response, improving the overall accuracy and amount of useful data.

The complexity of behavioral interventions poses difficulties for understanding the contributions of the active ingredients within these interventions.

Beyond the RCT: Alternative Experimental Designs

With the availability of these new types of data and the ability to create highly replicable interventions delivered via technologies such as smartphones, there are new opportunities for a much wider range of experimental designs beyond a traditional RCT. We are particularly interested in two broad classes: factorial/fractional factorial study designs and single-case experimental designs (SCEDs).

Factorial/Fractional Factorial Studies

Factorial study designs (see Table 1) are ideal for better understanding how components of interventions individually and interactively impact behavioral and health outcomes. They are central to the multiphase optimization strategy (MOST), a collection of methods developed by Collins et al. [6], which are focused on creating the most resource effective means for forming effective behavioral interventions. Within MOST, a fractional factorial study that uses assumptions about the lack of importance of high-level interactions between factors to reduce the overall number of conditions in an experiment is used as a screening experiment for determining which intervention components should be used or not used within a final behavioral intervention package. With increasing interest in developing the most resource-effective behavioral interventions, this experimental design becomes an increasingly valuable method for understanding active ingredients of behavior change.

Single-Case Experimental Designs

SCEDs provide ideal data for supporting the idiographic study of behavior and behavior change. SCEDs have a long history of use within psychology and related fields and, as a broad class of experimental designs, share the common feature of using each participant as her own control. Furthermore, a central component of almost all SCEDs is the use of repeated measurement of a factor of interest (e.g., behavior, symptoms, health outcome), which is increasingly easier to achieve with new data opportunities described earlier.

Dallery et al. [7] summarize a variety of common SCEDs, including the following.

- *Reversal study design* (often also labeled ABA): an outcome of interest is measured during a baseline phase, followed by the implementation of an intervention, which is then later removed again to reverse back to a baseline-like phase.
- *Alternating treatment design*: following a baseline phase, multiple different interventions are randomly activated and deactivated in an alternating fashion.
- *Multiple-baseline design*: a baseline period is conducted with varying durations across multiple participants and then an intervention is introduced to the participants in a staggered fashion.

TABLE 1. FULL FACTORIAL DESIGN (2 × 2 × 2)

Experimental Condition	Intervention 1 (Setting Goals)	Intervention 2 (Positive Reinforcement)	Intervention 2 (Social Support)
1	Off	Off	Off
2	Off	Off	On
3	Off	On	Off
4	Off	On	On
5	On	Off	Off
6	On	Off	On
7	On	On	Off
8	On	On	On

IEEE PULSE Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

- *Crossover design*: involves a baseline phase, introduction of intervention 1, followed by a switch to intervention 2, with ordering effects controlled across participants.

SCEDs provide much greater precision in supporting modeling of idiographic data compared with RCTs, which use a nomothetic (i.e., group-level) approach for generating insights about population-level causal effects. While there is value in this nomothetic approach, an idiographic approach, which is currently far less common in behavioral health research, will likely lead to more personalized and customized behavioral interventions. This will be accomplished, at least in part, via the development of data-driven case studies, whereby idiographic data will be modeled to delineate exactly how interventions, behaviors, and context interact on a case-by-case basis.

With the collection of enough data-driven case studies, it is likely that patterns across individuals will emerge of different user profiles or characteristics of individuals that predict and inform which intervention components to use when, how, and for whom. These user profiles are akin to tailoring methods from behavioral science, whereby either 1) factors such as demographics (e.g., African American women) was used to target certain message content to a user group, or 2) information about a person's behavior change process was used to tailor the information (e.g., providing the pros and cons for changing a behavior to someone thinking about changing a behavior versus providing actionable steps for someone actively engaged). The user profiles should be able to support a much wider scale of personalization of the content based on a wider realm of information. For example, the PREVE project [8] identified four blocks for doing this: 1) health variables (e.g., health status, personal goals), 2) motivational variables (e.g., personal values and interests), 3) resource variables (e.g., the types of support available), and 4) contextual variables (e.g., information about time, place, mood, activities, weather, presence or not of others) [9].

Health behavior interventions are intended to intervene in the behavioral status quo and shift behavior toward more healthful configurations.

Pivotal Challenges to Overcome

While there are great opportunities, there are also many challenges still ahead, both ethical and methodological. Next, we delineate four important challenges that require much greater focus and research for realizing the potential of these new methods.

Supporting Agency, Not Coercion

From the human–computer interaction (HCI) literature, a classic label for behavior change technologies was persuasive technology. While not intended, persuasive technology implicitly suggests some degree of coercion by persuading individuals to engage in an action that they might not otherwise want to engage in. On the basis of this, many in the HCI community are choosing to use other terms such as behavior change technologies. This shift in terminology reflects a larger desire among those building behavior change technologies to consciously support an individual's personal agency rather than unintentionally coercing them.

The development of behavior change technologies that support agency is still a very difficult problem to understand. Indeed, much previous research from psychology highlights how many decisions, including issues as complicated as morals, are often driven by implicit intuitive processes, rather than conscious rational ones. As such, it is imperative for future developers and researchers to balance the best strategies for promoting healthy behaviors with a person's own agency. We have started to explore some of this work with the Quantified Self, which is a movement of individuals focused on improving their lives through self-tracking (quantifiedself.com). As part of this, we have started to develop a do-it-yourself self-experimentation tool kit, whereby individuals are given the tools and resources, including insights from behavioral science, to design their own home-based solutions for intervening on their own behaviors [10]. This type of tool, whereby individuals are actively involved in not only changing their behaviors but the creation of the intervention itself, is likely to be very important for ensuring that we do no harm with these new technologies.

Privacy

With regard to behavioral big data, while such data collection techniques can greatly enhance behavior change interventions, they do not come without cost. Prime among these is privacy. Although many of the data collection methods we describe above are unobtrusive, they also can be very intrusive, collecting much more information about the person than she might be willing to share. Research is underway to develop methods to minimize such privacy risks. For instance, audio data can be filtered so that conversations cannot be reconstructed, data from social networks can be collected in an anonymized way so that only the frequency of communication is logged without logging exactly with whom the person is communicating, and sensitive data such as GPS coordinates can be kept just long enough to infer higher-level information (e.g., distance traveled) and then can be promptly discarded. Such privacy-preserving techniques of data collection are necessary to mitigate individuals' privacy concerns and to make it more likely that they would be willing to use behavioral interventions that can help them improve their health [11].

From Data to Insight

Beyond ethical concerns, the great variety of new data channels delineated above poses important challenges for appropriately filtering, organizing, and processing these data from their raw form into actionable insights. From signal processing and pattern recognition to fuzzy logic and algorithm development, the task of inferring meaning from sensors and digital footprints is no small task and will require continued exploration from an interdisciplinary perspective. Behavioral scientists can aid in identifying acceptable ground truth metrics or important variables to hone in on, while engineers can build on their wealth of methods to glean insights from raw data. While the challenge is large, there are many examples of data transforming to insight we already discussed, such as the translation of raw accelerometry values into estimates of physical activity and sleep or the detection of mood from audio recordings of speech. Even greater strides in the understanding of behavior will be supported through the continued refinement of our data processing methods.

Solving the Tower of Babel Problem

Building on this, the development of behavioral interventions that are delivered via new technologies increasingly requires interdisciplinary teams of experts and individuals who will use them to create the interventions, specify which data we need to collect, visualize what the devices should look like, build the devices, and write the required software. However, this transdisciplinary work can easily go wrong.

For example, there is a cartoon that depicts different perceptions about the creation of a swing set (see: www.projectcartoon.com/cartoon/2648). One image shows the customer's desire for a comfortable swing hanging from a branch; a second shows an architect's rendering of the swing to include an arm chair; a third image shows the project manager envisioning the swing being tied to two opposing branches on the tree; and the engineer's image shows the swing attached to those two branches but with two supporting beams and a hole cut through the tree to allow it to swing. While easy to see in a cartoon, each player may not see the different interpretations in reality because, even though they may speak to one another, they are often viewing the issue from different perspectives and using different field-specific languages, making communication difficult.

When we start designing behavior change technologies, we need to make sure that all parties involved speak a common language, one that can be understood across disciplinary domains, including those individuals we intend to build for, whenever possible. First steps toward the design of such a language have been taken [12], but much more work is required to fully solve this Tower of Babel problem.

Looking Forward

In this article, we sought to delineate methods for developing the next generation of behavioral interventions for supporting the 21st-century goal of effective behavioral management of health. While there are risks to these efforts, we see far greater potential benefit for helping individuals and the population as a whole to live more healthful, productive lives through the support of technologies that balance the strengths and limitations of humans with the strengths and limitations of technology. As researchers, we look forward to actively engaging in the creation of this new behavioral management-driven model of health.

Eric B. Hekler (ehekler@gmail.com) is with Arizona State University, Phoenix. Predrag Klasnja is with the University of Michigan, Ann Arbor. Vicente Traver is with Valencia Polytechnical University, Spain. Monique Hendriks is with Phillips Research, Eindhoven, The Netherlands.

As researchers, we look forward to actively engaging in the creation of this new behavioral management-driven model of health.

References

[1] E. B. Hekler, P. Klasnja, J. E. Froehlich, and M. P. Buman, "Mind the theoretical gap: Interpreting, using, and developing behavioral theory in HCI research," in *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '13)*, New York, May 2013, pp. 3307–3316.

[2] E. Villalba, D. Salvi, M. Ottaviano, I. Peinado, M. T. Arrendondo, and A. Akay, "Wearable and mobile system to manage remote heart failure," *IEEE Trans. Inform. Technol. Biomed.*, vol. 13, no. 6, pp. 990–996, Nov. 2009.

[3] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury, "StressSense: Detecting stress in unconstrained acoustic environments using smartphones," in *Proc. 2012 ACM Conf. Ubiquitous Computing*, 2012, pp. 351–360. 2009.

[4] S. Schueller, R. Munoz, and D. Mohr, "Realizing the potential of behavioral intervention technologies," *Curr. Directions Psychol. Sci.*, in press.

[5] M. N. Burns, M. Begale, J. Duffecy, D. Gergle, C. J. Karr, E. Giangrande, and D. C. Mohr, "Harnessing context sensing to develop a mobile intervention for depression," *J. Med. Internet Res.*, vol. 13, no. 3, p. e55, 2011.

[6] L. M. Collins, J. J. Dziak, and R. Li, "Design of experiments with multiple independent variables: A resource management perspective on complete and reduced factorial designs," *Psychol. Methods*, vol. 14, no. 3, pp. 202–224, Sept. 2009. doi: 10.1037/a0015826.

[7] J. Dallery, R. N. Cassidy, and B. R. Raiff, "Single-case experimental designs to evaluate novel technology-based health interventions," *J. Med. Internet Res.*, vol. 15, no. 2, p. e22, 2013. URL: http://www.jmir.org/2013/2/e22/ doi: 10.2196/jmir.2227

[8] N. Boye, M. Dahl, H. Hietala, K. Kaipainen, J. Parkka, M. A. Navarro, H. Galan, P. Garcia-Segovia, M. Martinez, J. Martinez, T. Meneu, A. B. Sanchez, V. Traver, A. Sanna, M. Nalin, M. Verga, M. Demeester, and N. Saranummi, "White Paper: ICT research directions in disease prevention," in *PREVE*, N. Saranummi, Ed. Brussels, Belgium: European Commision–Information, Society, and Media, 2011.

[9] A. Honka, K. Kaipainen, H. Hietala, and N. Saranummi, "Rethinking health: ICT-enabled services to empower people to mangage their health," *IEEE Rev. Biomed. Eng.*, vol. 4, pp. 119–139, 2011. doi: 10.1109/RBME.2011.2174217.

[10] E. B. Hekler, W. Burleson, and J. Lee, "A DIY self-experimentation tool kit for behavior change," presented at the Personal Informatics Workshop at CHI'13, Paris, France, 2013.

[11] P. Klasnja, S. Consolvo, T. Choudhury, R. Beckwith, and J. Hightower, "Exploring privacy concerns about personal sensing," in *Proc. Pervasive Computing*, 2009, pp. 176–183.

[12] M. Hendriks, V. Antila, and T. Lavrysen, "Inter-usability and intelligent communication: Usability aspects in a multidevice personal attentive system," in *Workshop on Ambient Intelligence Infrastructures (WAmIi)*, 2012, pp. 15–18.