Exploring a Theory-Guided Path to the Design of Personal Informatics and Intervention Technologies

Elizabeth Murnane
HCIC 2015 Boaster

At a high level, my goal is to design and deploy behavioral coaching and intervention systems that better complement individual users' unique characteristics. I am motivated by the idea that a person's innate attributes and current contexts interact in ways that produce idiosyncratic behaviors that generic interventions may not suit. Using domain knowledge as a guide, my research develops lightweight profiling methods that leverage people's digital footprints in order to passively sense pertinent psychological, physiological, and behavioral characteristics -- and ultimately tailor support accordingly. I argue that such a theoretically-informed approach that is able to robustly mine, model, and accommodate personal variability will lead to technological solutions that better meet individual needs and produce more positive outcomes.

INTRODUCTION

In this boaster, I survey the current landscape (from an HCI orientation) of personal informatics and behavioral intervention tools grown in both academia and industry. As part of this review, I describe steps that HCI researchers typically take (or fail to take) to personalize systems and how they commonly incorporate (or fail to incorporate) theory as part of persuasive system design. I also identify opportunities to further advance theory-guided personalization -- in particular, by developing computational methods that mine behaviors using approaches aligned with the theoretical underpinnings of those behaviors.

I then present research I have undertaken to pursue this approach: a case study guided by chronobiological theory that mined social sensor data to better understand, model, and predict patterns and events related to sleep behaviors and disruptions.

Behavioral Awareness and Intervention Technology

HCI researchers are increasingly exploring the value of technology for behavioral coaching and intervention, for instance in the form of personal informatics (PI) tools aimed at helping users collect and reflect on personal information (Li et al., 2010) or through persuasive technologies designed to change or maintain attitudes and behaviors (Fogg, 2002)[1]. Compared to traditional behavioral intervention strategies (e.g., in-person therapy), technology-mediated approaches enable broader access, deeper integration into daily life, and more easily repeated administration (Schueller et al., 2013).

Domains of application for research on personal informatics and persuasive design commonly include sustainability, social participation, and health. The lattermost, health -- both physical and mental, is arguably the area of behavioral intervention garnering the most attention from HCI researchers.

Theory-Driven Technology Design

Research advocating for theory-driven approaches to HCI research and BCT development argue that the use of theory[2] can inform the design and evaluation of systems (Hekler et al., 2013) and that the efficacy of technology-mediated interventions depends upon a theoretical understanding of the mechanisms and determinants of behavior and behavior change (Michie, 2008). Prominent examples of frameworks commonly employed to inspire and inform BCT research include goal-setting theory (Locke & Latham, 1990), the Transtheoretical Model (TTM) (Prochaska & Velicer, 1997), and Fogg's behavioral models for persuasive design (Fogg, 2002; Fogg, 2003; Fogg, 2009). Researchers have continued to build on these frameworks; particularly relevant to BCT research is the stage-based model of personal informatics (PI) systems from Li et al. (2010), which adapts TTM concepts to PI processes -- defining the stages in terms of phases of data capture, integration, reflection, and informed action and identifying barriers individuals face within each PI-tuned stage.

Theory-Informed Endeavors

A range of prior work illustrates application of these theories. One of the earliest and perhaps most well-known examples of theory-guided persuasive design is the UbiFit system for supporting individuals in monitoring and maintaining physical activity. UbiFit implements theory-driven goal-setting, rewards, and feedback through a glanceable mobile display, where a garden of flowers represents user-added activities and butterflies represent attained goals (Consolvo et al., 2008a; Consolvo et al., 2009a). Also influenced by the Transtheoretical Model, UbiFit targets the contemplation,

[1] A note on terminology: for the sake of brevity and consistency, I define the term "Behavioral Coaching Technology" (BCT) to refer to this broad class of technology aimed at supporting behavioral awareness and/or Intervention.

[2] Again, a note about terminology: I use "theory" to refer to different forms of domain-knowledge, including conceptual frameworks, theoretical constructs, and experimental findings.

preparation, and action stages of change (Consolvo et al., 2008b, Consolvo et al., 2009b). The system further adopted Goffman's theory of Presentation of Self (Goffman, 1959) as well as Cognitive Dissonance Theory (Festinger, 1957) into the design choices of its garden display. Fish'n'Steps (Lin et al., 2006) also drew inspiration from goal-setting theory as well as the TTM in its similar implementation of a personal display that reflects users' physical activity levels through fish characters, where the size and facial expressions of fish map to goal progress and achievement.

Numerous other examples of technologies designed on premises of goal-setting include (just to give a representative sampling) systems to encourage sustainable behaviors (e.g., Froehlich et al., 2010) such as water conservation (e.g., Laschke et al., 2011) or reduced energy consumption (e.g., He & Greenberg, 2010, Erickson et al., 2013); adoption of learning goals (e.g., Shi & Louvigne, 2014); increased physical activity (e.g., Munson & Consolvo 2012; Munson et al., 2015) including at the group or family-level (e.g., Colineau & Paris, 2011) or in combination with other wellness activities such as diet management and relaxation (e.g., Gasser et al., 2006; Mattila et al., 2008); and adherence to habits generally speaking (Panovich, 2013; Stawarz et al., 2015) for instance by incorporating psychological theory related to habit formation (Skinner, 1953; Wood, 2007) and positive and negative reinforcement (Michael, 1975). Other systems similarly incorporate reward and punishment schemes based on reinforcement theory, notably those utilizing gamification elements to motivate user engagement (Nicholson, 2012). Along the same lines, persuasive technologies guided by Fogg's behavioral theories span multiple domains including sustainability (e.g., Sugarman & Lank, 2015), social activism (e.g., Rege et al., 2006), and physical activity (e.g., Albaina et al., 2009).

Beyond these commonly leveraged frameworks, other theories from psychology and behavioral and cognitive science have also been adopted by HCI researchers. For instance, studies of technology-mediated social participation have a history of drawing upon social psychology to derive motivational strategies (Ling et al., 2005). Motivational frameworks based on Self-Determination Theory (SDT) (Deci & Ryan, 2000) have been incorporated into persuasive technology to encourage exercise (Ferron & Massa, 2013), while the Nudge Framework (Thaler & Sunstein, 2008) and theory of planned behavior (Ajzen, 1991) have been incorporated into personal informatics tools (Cecchinato et al., 2014).

While such work has made swift and impressive advances and imparted legitimate benefit to individuals striving to make positive behavioral changes, *most* current BCT instantiations *do not* make similar efforts to incorporate relevant theory into designs (Abroms et al., 2011; Hekler et al., 2013), even though research shows theoretically-informed interventions are more successful (Michie et al., 2011; Rimer & Glanz, 2005). Further, in many cases theories are only given a cursory mention and leave details ambiguous as to how constructs were actually translated or incorporated into design elements (Orji, 2013). In addition, most BCT take a generic approach to intervention and offer one-size-fits-all strategies, even though individuals respond differently to design cues (Nov et al., 2013), idiosyncratic differences can influence the efficacy of an intervention for a given person (Kaptein, 2012), and overall, successful behavior change depends on more personalized solutions (Busch et al., 2013; Cecchinato et al., 2014).

The Value of Personalizing

Personalization has been defined as, "a process that changes the functionality, interface, information content or distinctiveness of a system to increase its personal relevance to an individual" (Blom 2000) -- that is, to support individual needs and goals. Personalization has been identified as a key factor for behavior change (Berkovsky et al., 2012) given that innate characteristics as well as extrinsic variables can interact to influence behavior (Endler & Parker, 1992; Nov et al., 2013). Research shows that personalizing content and experiences can reduce cognitive load, improve user satisfaction, strengthen the impact of persuasion, and promote adherence (Fan & Poole, 2006; Orji, 2014; Tuzhilin, 2009). Regarding behavioral coaching, a number of personal, behavioral, and environmental factors may be at play that can affect whether a person will perform positively or negatively while pursuing a behavior-change goal (Bandura, 1986; Bandura, 1997; Locke et al., 1981); and focusing even more specifically on technology-mediated behavioral intervention, prior work reveals a number of psychological, cognitive, and contextual traits that can impact intervention efficacy (Kaptein et al., 2010; King et al., 2006).

Commonly explored psychological differences typically revolve around personality since it is a well-established construct with well-validated instruments available to measure it. Studies have found that personality helps determine a person's motivation and engagement (Nov & Arazy, 2013; Nov et al., 2013) as well as reaction to persuasion (Kaptein & Eckles, 2012; Nguyen et al., 2007) including persuasion delivered through healthy behavior change technologies (Halko & Kientz, 2010; Karanam et al., 2014). Research has also found variability in the persuasiveness of behavior change strategies according to individual differences in emotionality and cognitive style (Petty & Brinol, 2015; Tercyak et al., 2009) as well as gender (Orji, 2014).

Turning to external variables that are similarly variable per individual, context-awareness refers to a class of systems capable of determining features of a user's environment such as location, nearby people, and available resources (Schilit et al., 1994). Contextual information additionally includes a user's current situation (Dey 2001); and elements such as time, light, sound, temperature, and movement have also been used to operationalize context (Bettini et al., 2010; Zimmermann et al., 2007). Contextual information may help users increase self-awareness and recognize how certain behaviors (e.g., physical activity) are being impacted by contextual factors (Li, 2012), and context-awareness has been identified as key to mobile-based persuasive technology designs (Bogost, 2007; Fogg, 2002).

User-Driven Personalization

One strategy for determining individual traits on which to personalize BCTs is directly asking people about themselves through self-report scales and surveys. Self-assessment is in fact considered the gold standard for some aspects of behavior and behavior change (Glasgow et al., 2005); and as the movement towards quantifying the self grows and the use of monitoring devices and personal informatics software becomes more widespread, people are increasingly using technology to measure and record behaviors of their own volition. For instance, seven in ten U.S. adults track a health indicator for themselves or for a loved one (Fox & Duggan, 2013a; Fox & Duggan, 2013b). In addition, individuals now combine various types of technology (e.g.. mobile applications, physical devices such as wearable trackers, and so on) to track behaviors through multiple styles (Rooksby et al., 2014).

However, such user-driven data collection is burdensome nonetheless (Connelly et al., 2006), and briefer assessments come with their own limitations related to degraded reliability (Burisch, 1997). In addition, it becomes increasingly infeasible for individuals to capture the array, amount, and granularity of information necessary to produce a comprehensive behavioral profile comprised of multiple personal variables and behavioral determinants (Bentley et al., 2013). Further, an approach dependent on self-report is especially problematic for certain contexts and populations, for instance people with compromised self-assessment abilities, which include individuals with bipolar disorder (Goodwin & Jamison, 2007) or people experiencing sleep deprivation (Dorrian et al., 2003).

Another user-driven form of personalization is customization, which provides users with interactive options to select or tailor features in order to meet their own preferences and needs. Research shows user-driven personalization can be successful, for instance by increasing attachment, appreciation, and satisfaction with services (Fan & Poole, 2006) or by enabling the expression of individuality (Blom & Monk, 2003; Sung et al., 2009). However, there are also disadvantages associated with customization. Ironically, choices can be formulaic, shallow, and inadequate for satisfying individuals' idiosyncratic needs -- needs into which users do not even necessarily have good insight (Simonson, 2005); while at the same time, too many choices can be overwhelming and demotivating (Iyengar & Lepper, 2000).

Thus, interest arises in more system-driven approaches to personalization, for instance by passively measuring and modeling individual characteristics from behavioral trace data to better enable systems to make decisions about what users want and need.

System-Driven Personalization

System-driven approaches to personalization emerged over recent decades and have progressed as computational techniques advance and users' data trails grow.

Much of this effort stemmed from the development of recommender systems, which use strategies such as collaborative filtering or content-based filtering to profile a user's preferences based on information such as her clickstream data or social network activity and connections (Adomavicius & Tuzhilin, 2005). Ever-mounting volumes of user data have allowed system-driven personalization to expand to additional domains including education, entertainment, and health (Mayer-Schonberger & Cukier, 2013).

Indeed, much of what enables system-driven personalization is today's high levels of technology ownership and usage and the wealth of user-specific data being generated. In the US, ownership of technology and levels of usage are at all-time highs with over 78% of people owning desktop or laptop computers, 90% owning cellphones, and 64% owning smartphones specifically (Duggan & Smith, 2014). At the same time, the data-capture features of these technologies are becoming increasingly extensive and sophisticated; for instance, now-standard sensors on most mobile smartphones include GPS, accelerometer, compass, gyroscope, ambient light detection, proximity detection, dual microphones, and dual cameras (Lane et al., 2010). Adoption of sensor-laden wearable devices similarly continues to grow (Chen et al., 2012). Further, web-based digital traces swell as 87% of people in the US use the Internet (compared to 14% in 1995), and 74% use a social networking site of some kind, with 42% using multiple social networking sites (Fox & Rainie, 2014).

Research in persuasive technology is increasingly pursuing system-driven personalizing strategies, with a growing trend toward passively sensing a user's current state; developing profiles about psychological, behavioral, or contextual elements of interest; and tailoring support for behavior-change accordingly (Kaptein et al., 2010; Petkov et al., 2012; Prost et al., 2013).

Systems aimed at inferring user traits include Mood Scope (LiKamWa et al., 2013), which models a user's mood from smartphone patterns (application use, phone calls, text messages, emails, web browsing, and location changes). Similar phone-usage features (including phone call duration, the numbers of received and sent calls and text messages at different times of day, and the tie strength of contacted parties) have similarly been used to predict more stable psychological traits such as personality (De Oliveira et al., 2011). More web-based data such as text (Golbeck, 2011a), preferences (Golbeck, 2011b; Kosinski et al., 2013), and contacts (Quercia et al., 2011) from social media have been used to model personality traits as well. Social-sensing based methods have also used device sensors (e.g., bluetooth proximity) to uncover connections among sleep, mood, and sociability (Moturu et al., 2011).

Other systems are geared around providing feedback to users. The Mobile Health Mashups system (Tollmar, 2012) and follow up Health Mashups system (Bentley et al., 2013) combine automatically-sensed data (digital calendar, location coordinates, weather) with data from wearables and other devices (Fitbit and Withings scale to capture steps, sleep and weight) along with manually-input information (exercise, food, mood, pain) in order to synthesize data and present the results of statistical analyses back to the user as visualizations and feeds of observed correlations expressed in natural language sentences. The BeWell system (Lane et al., 2011; Lin et al., 2012) combines activity recognition (e.g., walking, running, stationary), sleep inference (using noise levels and charging events), and social interaction detection (based on microphone data) to generate well-being scores for these 3 dimensions, each of which are reflected back to the user through different characters in an ambient animated wallpaper display (fish, turtle, and school of fish, respectively). MyBehavior (Rabbi et al., 2015) combines automatic with manual logging to track physical activity and eating behaviors in order to generate healthy lifestyle suggestions personalized to the user's context and previous behaviors.

Closing The Loop: Connecting Theory, Personalization, and Design

However, we again encounter shortcomings -- in this case, of current attempts at system-driven personalization. Namely, their methods are not theoretically informed but rather take more data-driven approaches based mostly on intuition or trial and error (Nov & Arazy, 2013). This is understandable given the aforementioned wealth of available user-specific data and considering the time and effort required to reach out to domain experts or to independently investigate and assimilate relevant theories; understandable, but not desirable.

First, it can be argued that these systems are not true instances of personalization since they essentially treat the user as a metric to be algorithmically optimized, are based on a narrow set of assumptions regarding who the user is, and overall cannot fully accommodate more intricate and multi-faceted user identities (Mayor-Schonberger & Cukier., 2013). This means that such approaches may not be modeling the elements most relevant to the behavioral domain of inquiry or to the idiosyncratic user who will receive the behavioral intervention. Said another way, such systems may simply have a veneer of robustness, where success has been achieved in fitting a model rigorously -- but this achievement holds little meaningful value if the appropriate constructs are not being captured and modeled in the first place.

Similarly, the wealth-of-data justification does not actually hold particularly well since exhaustive approaches become less and less feasible the more comprehensive the user profile and complex the behavioral intervention. I pointed out earlier that it is unrealistic to expect an individual to capture such an expansive amount of data, and it becomes similarly impractical for systems due to limits of computing power and performance. For instance, battery drain can plague even the simplest of continuous sensing toolkits (Lane et al., 2010; Wang et al., 2009). A theory-driven approach that is more targeted towards capturing features likely to hold relevant information and operationalizing them in a way that is theoretically suitable to the task will both help prevent important facets of data from being overlooked as well as avoid unnecessary processing that uses up finite computational resources.

Finally, though these systems may be capable of identifying statistical patterns from sensed behaviors -- e.g., "Do I sleep better on nights after I work out?" (Bentley et al., 2013), without a theoretical foundation, we may easily misinterpret such behaviors or fail to control for confounding factors that actually underlie them (e.g., daylight exposure for this example). Instead, a theory-driven approach can allow us to more confidently interpret observations and go beyond simply describing *what* is observed to get closer to *why*. That is, bringing a consideration of theory to bear enables the potential for research to also speak back to and refine existing theories.

My research aims to close these gaps among theory, personalization, and BCT design. I now focus attention on a domain exemplifying the value in pursuing theory-driven, personalized approaches to the design of personal informatics and intervention technologies: sleep.

CASE STUDY: CHRONOBIOLOGY

For full paper, see: Murnane, E. L., Abdullah S., Matthews, M., Choudhury, T., Gay, G. *To appear in UbiComp 2015*. Social (Media) Jetlag: How Usage of Socio-Computational Technology Modulates and Reflects Circadian Rhythms.

Motivation

Sleep has gained considerable recent interest, leading to the development of tools to help users track sleep patterns and duration, evaluate sleep quality, and adopt healthy sleep and wake schedules. Work is still nascent -- technologies are relatively intrusive or burdensome (e.g. requiring users to wear equipment or manually log sleep and wake events), and sensing is not yet particularly sophisticated (e.g., it cannot handle complex sleep environments for instance with partners or pets). However, the shortcoming on which we focus is the tendency of these systems to present generic recommendations rather than provide support that is personalized and accommodates individual variability, both contextually and biologically speaking; for instance, consider the blanket recommendation, "End caffeine consumption 8-14 hours before bedtime" (Bauer et al., 2012), even though caffeine does not affect everyone equally (Yang et al., 2010). As another example, variants in the per3 clock gene can significantly influence aspects of a person's daily functioning such as the individually-variable response to sleep deprivation (Viola et al., 2007).

This harks back to my central argument: most behavioral coaching technologies do not take individual differences into consideration (in this case, circadian rhythms and chronobiological traits) nor incorporate endogenous and exogenous personal factors such as chronotype, daytime light exposure, and social constraints into their assessments -- in this case, of sleep. Social computing research on associations between technology use and health-related behavior including sleep similarly lack the domain-specific underpinnings necessary to holistically interpret observations in a way that bears in mind latent biological aspects.

Guided by a theoretical understanding of the biology behind sleep and wake behaviors, this case study aimed firstly to better understand the interplay between external factors and internal rhythms and to secondly develop novel sensing techniques that leverage such awareness of circadian factors to more accurately assess neurobehavioral functions and misalignments.

Background

Chronobiology is the field of study concerned with the rhythms that guide biological functioning. The biological cycles of all living organisms, humans included, are coordinated by endogenous body clocks that maintain a circadian period and use external cues (dominantly sunlight) to remain synchronized to periodic environmental changes. Individual differences exist in these functions, are reflected by a person's *chronotype,* and result in individual variations in the preferred timing of sleep and its duration (Kreitzman & Foster, 2011). A common distinction is made between early and late chronotypes – people who prefer to wake earlier or sleep later.

Undergraduate students tend to be on the later end of the chronotype scale and therefore experience the most severe symptoms and consequences of social jet lag: the instability between sleep schedules across days of the week that stems from social schedules interfering with biological sleep preferences (Roenneberg et al., 2007). Studies also show that undergraduates suffer from chronic loss and interruption of sleep, which can lead to poorer academic performance, increased stress, additional mental health problems, and increased consumption of drugs and alcohol (Taylor & Bramoweth, 2010).

Method

To explore undergraduate students' sleep-related behaviors along with how their technology-mediated social interactions not only impact those behaviors but may also enable the computational assessment of circadian patterns and disorders, we captured a combination of quantitative and qualitative data through phone and social media logs (Computer-Mediated Communication data -- "CMC"), survey instruments, sleep diaries, and periodic in-person interviews. Using public mailing lists and snowball sampling, we recruited a sample of 9 participants (7 males, 2 females) aged 19-25 years old. All of our participants had been using smart-phones for at least 6 months prior to the beginning of our study.

Results

Daily Technological and Biological Rhythms

To begin, we analyze phone probes, social media logs, and sleep diaries to gain a sense of typical trends in participants' technology use and sleep-wake behaviors as well as potential links between them. We observe the daily usage trends shown in Figure 2. We see that usage is heaviest in the late evening, until about 11pm. Levels of social media app usage and Facebook posting activity in particular continue slightly later until around 1am. Diaries indicate participants then go to sleep within an average of 49 minutes.

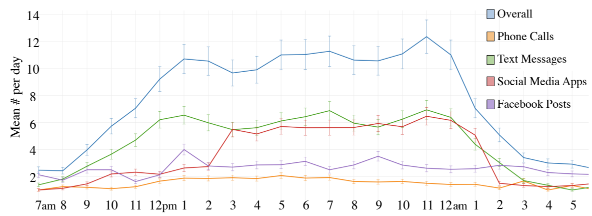

Figure 1: Daily trends in average CMC usage

To support the assumption that our small-scale participant pool is representative of college students more generally, we also compare these observations to those from prior studies, consistently finding close alignment. To go beyond such studies, we are next interested in whether we can use this data to support *circadian-aware sensing* and infer circadian variables: sleep, circadian disruptions, and neurobehavioral functioning.

Leveraging Social Data for Sleep Sensing

To investigate techniques for reliably and unobtrusively detecting sleep behaviors and circadian misalignments, we first attempt to infer sleep events from CMC patterns by implementing the sleep-inference algorithm presented in (Abdullah et al., 2014), which is built on screen on/off patterns. We instantiate our algorithm using our phone probe data, social media app use logs, and Facebook posts to model sleep events according to the longest nightly gaps in CMC usage. Table 1 presents the accuracy of our sleep duration inference compared with the screen on/off approach and with participants' ground truth sleep journals. The results show the reliability of our approach, which achieves an average difference of only 23 minutes between socially-sensed and self-reported sleep duration. This is a more accurate (and ubiquitous, given the data is web-based) prediction than from screen on/off; and we also manage to outperform more complex algorithms based on environmental factors such as light, movement, and sound as well as phone locking and charging events (Chen et al., 2013).

	Social Data	Screen On-Off	Ground Truth Diary
P1	8.44*	8.54*	8.13
P2	7.64*	8.09	7.45
P3	8.21*	8.33*	8.15
P4	7.53*	8.02*	7.25
P5	6.11*	5.44*	6.12
P6	7.15*	7.17*	7.13
P7	7.63	7.16*	7.14
P8	7.38*	7.30*	8.14
P9	7.48	5.42	6.25

Table 1: Average sleep duration according to sensed and self-report data. (* = inference falls within 95% confidence interval based on sleep diaries, p < .01)

Our approach overestimates sleep since the stop and start of CMC use do not precisely ajoin sleep onset and wake. By incorporating an error term to the calculation of sleep duration per participant (based on chronotype plus individual differences in pre-bed and post-wakeup CMC use learned from the first week of study data), we can somewhat correct for this non-usage gap, and more complex learning can further improve accuracy. Conversely, underestimation in sleep duration can occur when we mistake notifications and other forms of incoming media as participant-generated actions or when glitches in phone probing produce errors in sensor logs. We attempt to filter out such data not indicative of user activity according to usage-time thresholds, and more sophisticated instrumentation can further help eliminate such false alarms. Our interview data allows us to uncover other points of failure and opportunities for improvement. P9 described watching movies and using Twitter as typical nightly activities, and she also noted normally checking email and texts upon waking. Similarly, P7 told us that morning phone use involved weather and calendar checking, and he discussed using Facebook and playing video games before bed but explained that he does so on a personal computer rather than the phone. Thus incorporating into our sensing both additional forms of social data (e.g., Twitter, email) as well as broader non-CMC usage data (e.g., app logs, web histories) from across multiple devices would be a straightforward next step towards more precise sleep-event estimations.

Assessing Circadian Disruption

Social constraints can result in later sleep onsets and earlier wake times that oppose our own internal timings. Alarmingly, it is estimated that over 80% of the population suffers from this social jet lag (Roenneberg et al., 2012). Our sensing approach for inferring sleep onset and wake events is next able to detect a related well-known chronobiological phenomenon called the "scissors of sleep" as seen in Figure 2. Here, later chronotypes' sleep is systematically shortened on work days, which causes an accumulation in sleep debt that is then compensated for by sleeping more on the weekend, while weekday schedules fit better with internal timings of early chronotypes (P4) but weekend sleep for them is forced to shift due to social engagements.

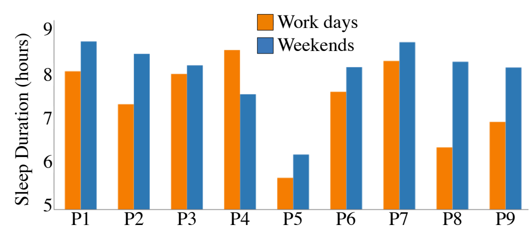

Figure 2: Detecting the "scissors of sleep"

To next quantify social jet lag and assess its severity across our participants, we compute the difference between mid-sleep (the halfway point between sleep onset and waking) on free days (MSF) and on work days (MSW) per Wittman et al. (2006):

$$\Delta MS = |MSF - MSW|$$

Figure 3 shows the results of this calculation according to social-sensed data, presented according to participant chronotype. Specifically, we see increased social jet lag on the extreme ends of the chronotype spectrum, and as expected (Wittman et al., 2006), it is most severe for our later types since their socially-constrained days (work days) outnumber their free days (weekends).

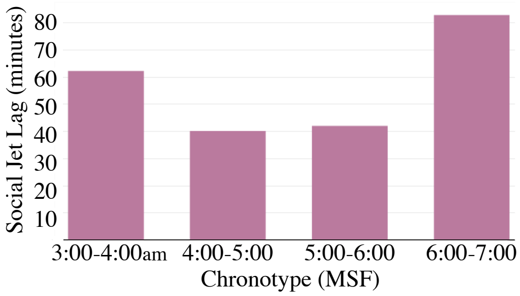

Figure 3: Socially-sensed avg. social jet lag (discrepancy between mid-sleep on free days and work days)

Monitoring Neurobehavioral Functioning

Social jet lag has numerous detrimental consequences, with symptoms manifesting as cognitive difficulties and emotional problems. We explore the impacts of sleep on such neurobehavioral functioning the following day, focusing on attention, cognitive performance, and mood. These characteristics are known to exhibit strong circadian patterns, suffer substantially after sleep loss and interruption, and are considered especially important attributes to evaluate for individuals in our participants' age group (Oginska & Pokorski, 2006).

We explore utilizing a number of socially-sensed variables in order to operationalize activity levels, social interactions, emotions, and cognition, all of which prior research and our own experimentation suggest as highly relevant to performing such circadian assessments. Here we present our analyses that reveal meaningful differences in these variables on days following nights of varying sleep quality. Comparisons are performed on medians using Wilcoxon sign-rank tests. Following established guidelines, we treat sleep durations of 7 - 9 hours as "adequate" and durations outside this range as "inadequate" (Chen et al., 2006), though as with internal timing of sleep, there are individual differences in sleep need as well (Roenneberg et al., 2004).

Attention and Cyberloafing. Cyberloafing is a term used to refer to idling and procrastination behaviors (Lim et al., 2002). Such tendencies to postpone tasks may be explained by a lack of attention and an inability to focus that stem from insufficient self-regulatory resources, which drain over the course of a day and require adequate sleep to become restored (Baumeister et al., 2000). Both sleep quantity and quality are important to this restoration (Hursch et al., 2004), and an individual's failure to obtain both can result in increased levels of cyberloafing (Muraven et al., 1999). To capture cyberloafing behaviors, we therefore compute the following interactivity-based measures:

- Volume: Total number of technology-mediated social interactions a participant performs in a given day between initially waking and finally going to sleep.
- Burstiness: Max number of interactions of a participant in any single hour between wake and sleep.
- Frequency: Number of hours between a participant's successive interactions.

As seen in Table 2, nights of insufficient sleep relate to more CMC-based interactions the following day, and these interactions are more frequent and in tighter temporal bursts. Correlating hours of sleep with the amount of next-day activity shows the same negative relationship ($r = -.52$, $p < .01$). During interviews, participants all mention checking social media when having trouble focusing or concentrating, which they express often happens when tired – e.g., "If I'm more tired, I'm less able to pay attention in class and more likely to use phone to avoid falling asleep or get bored more easily".

	Adequate	Inadequate
Volume **	18	34
Burstiness ***	6.12	9.54
Frequency ***	0.71	0.43

Table 2: Median values of CMC activity after nights of Adequate vs. Inadequate sleep. Sig differences in medians marked on variable (p<.001, ***p<.0001)**

Cognitive Performance. Sleep deprivation leads to impaired academic performance, and sleep loss makes circadian variation in performance most evident. Fatigue coupled with endogenous changes in daily brain function have even been equated to alcohol intoxication. Conversely, adequate sleep duration improves learning and problem solving (Wagner et al., 2012).

Utilizing participants' Facebook posts, we perform standard pre-processing on the text-based content of posts (e.g., removing punctuation and URLs, handling spelling errors, and so on) and then calculate the following cognitive-based measures, which represent the sophistication of a participant's posts and the cognitive complexity the writing required:

- LIX: A readability measure that indicates the difficulty of reading a piece of text (Bjornsson, 1968), computed as the percentage of words of 7 or more letters plus the average number of words per post.
- TReDIX: A LIX-based measure adapted for use with social media content, computed as ratio of total count of words 7 or more letters long appearing in all posts made within a time period over the total number of posts made in that time period (Hutto et al., 2013).

As seen in Table 3, an adequate amount of sleep relates to higher levels of complex thought according to both measures. Linear regression confirms a positive relationship – the fewer hours of sleep, the lower subsequent demonstrated cognitive ability according to social-sensor assessment ($\beta = 2.17$, $r^2 = 0.12$, $p < .001$).

	Adequate	Inadequate
LIX *	0.3592	0.3003
TReDIX **	0.2738	0.2144

Table 3: Median values of cognitive performance after nights of Adequate vs. Inadequate sleep. Sig differences in medians marked on variable (*p<.05, **p<.01)

Mood. Consequences of sleep reduction include tension, negative emotions, and irritability (Oginska & Pokorski, 2006). Conversely, extending sleep improves alertness, reaction time, and mood (Kamdar et al., 2004). To evaluate if social-sensor data can reflect circadian patterns in mood, we again use Facebook posts and this time apply psycholinguistic analysis techniques to compute the following sentiment-based measure:

- Sentiment Intensity Rate: A measure of how intensely positive or negative emotions are, computed as the ratio of the sum of valence intensity of positive or negative language in posts to the total number of posts in a time period (Hutto et al., 2013).

To avoid skewed results from participants with many more Facebook posts, we normalize values of sentiment variables to between 0 and 1 (values closer to 1 indicate levels of the sentiment variable are nearer to the maximum value ever observed for that individual and values closer to 0 indicate levels nearer the minimum). Table 4 shows the differences in positive and negative emotions expressed after adequate and inadequate sleep.

	Adequate	Inadequate
Positive Sent. Intensity ***	0.5373	0.3057
Negative Sent. Intensity ***	0.4176	0.8388

Table 4: Median values of sentiment in Facebook posts after nights of Adequate vs. Inadequate sleep. Sig diff in medians marked on variable (*p<.0001)**

We find positive affect following nights of adequate sleep is 1.75 times higher than nights of inadequate sleep, after which negative sentiment is over twice as high. Figure 4 illustrates differences in sentiment after nights of varying sleep duration. In interviews, participants consistently note usage is higher when energy and mood are lower (e.g., feeling "more down" or "down and frustrated") and also express using social media to "vent" or seek social support when tired and irritated.

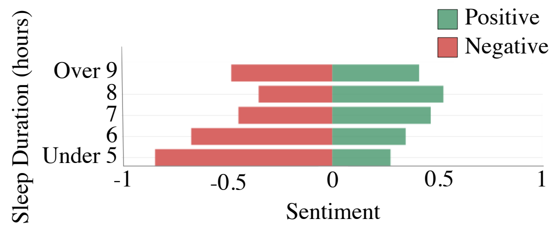

Figure 4: Sleep duration and sentiment the next day

Overall, our findings relate to the larger notion of the complicated interplay between technology use, sleep, and circadian traits. We saw sleep behaviors and social jet lag varying by chronotype; and we observed associated levels of social media use with significantly more lost sleep, which in turn connected to fatigue, cognition, and emotion – however, still unsolved is the chicken-and-egg uncertainty regarding underlying factors behind technology use and other possible explanations for circadian disruption and its neurobehavioral consequences. There is fertile ground for future research to shed light on these complex, multifaceted relationships.

Conclusion

We envision circadian-aware systems that can sense and respond to individual variations in order to more accurately model daily functioning and supply interventions in line with innate biological preferences. A first step is providing feedback to increase users' awareness about how patterns of technology use may act as a gateway to lost sleep, procrastination, or depressed mood. As examples of more direct intervention, systems offering sleep advice could be tailored to each user's genetic and environmental conditions, or a circadian-attuned calendar could provide recommendations for scheduling different types of activities such as studying or exercising based on chronotype.

Overall, this work demonstrates how relationships and phenomenon well known in a particular domain of study can be made apparent through analysis of technology use, which we can leverage as part of more personalized behavioral assessment and intervention appropriate for that context.