Conservation of Procrastination: Do Productivity Interventions Save Time or Just Redistribute It?

Geza Kovacs[1], **Drew Mylander Gregory**[1], **Zilin Ma**[2], **Zhengxuan Wu**[1], **Golrokh Emami**[1], **Jacob Ray**[1], **Michael S. Bernstein**[1]
[1]Stanford University, [2]Bucknell University
{geza,msb}@cs.stanford.edu

ABSTRACT

Productivity behavior change systems help us reduce our time on unproductive activities. However, is that time actually saved, or is it just redirected to other unproductive activities? We report an experiment using HabitLab, a behavior change browser extension and phone application, that manipulated the frequency of interventions on a focal goal and measured the effects on time spent on other applications and platforms. We find that, when intervention frequency increases on the focal goal, time spent on other applications is held constant or even reduced. Likewise, we find that time is not redistributed across platforms from browser to mobile phone or vice versa. These results suggest that any conservation of procrastination effect is minimal, and that behavior change designers may target individual productivity goals without causing substantial negative second-order effects.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**;

KEYWORDS

Behavior change; distractions and interruptions

ACM Reference Format:
Geza Kovacs, Drew Mylander Gregory, Zilin Ma, Zhengxuan Wu, Golrokh Emami, Jacob Ray, Michael S. Bernstein. 2019. Conservation of Procrastination: Do Productivity Interventions Save Time or Just Redistribute It?. *In Proceedings of ACM CHI Conference on Human Factors in Computing Systems (CHI 2019), May 4-9, 2019, Glasgow, Scotland UK.* ACM, New York, NY, USA, Paper 330, 12 pages. https://doi.org/10.1145/3290605.3300560

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300560

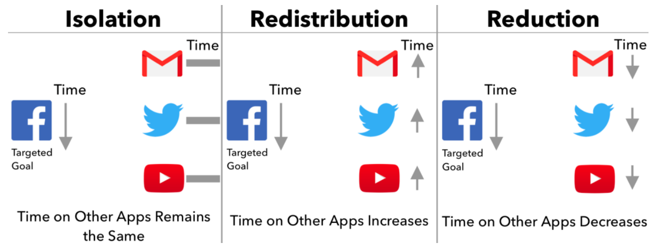

Figure 1: When interventions reduce time on a targeted goal such as Facebook, the time saved may (left) be isolated from effects on other goals, (center) be redistributed to other goals, or (right) decrease time spent on other goals.

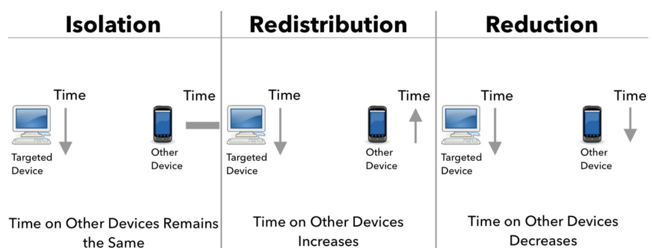

Figure 2: When interventions reduce time on a targeted device e.g. a browser, the time saved may (left) be isolated from effects on other devices, (center) be redistributed to other devices, or (right) decrease time spent on other devices.

1 INTRODUCTION

We use productivity behavior change interventions to try to keep ourselves in focus. But do these systems truly save us time? Or do they just redistribute the time elsewhere? In other behavior change domains, interventions sometimes have effects on behaviors other than the ones they were targeting [17, 55].

One possibility is that interventions narrowly impact just the goal that they target, and have no effect on time spent elsewhere. We will refer to this as the *isolated effects hypothesis.* Taking the relationship between time spent on Facebook and Instagram as an example, the isolated effects hypothesis would predict that an intervention that helps reduce time on Facebook should have no effect on time spent on

Instagram. Persuasive systems often claim to result in the intended behavioral changes without observable consequences elsewhere, lending support for this hypothesis [4, 5, 7, 18, 39]. If the isolated effects hypothesis is true, overall productivity can be boosted through interventions that individually target each goal.

However, people have a limited supply of willpower [8], can maintain focus for only so long [20, 32, 42], and need downtime — so perhaps the time saved is actually just redistributed to other unproductive applications. We will refer to this as the *redistribution hypothesis*: saving time on one unproductive application results in an increase in time spent on other unproductive applications. Returning to our example of a productivity intervention targeting Facebook, redistribution would hypothesize that an intervention that reduces time on Facebook will increase time spent on Instagram. Redistribution may be partial, where the time redistributed is some fraction of what was saved. Or more bleakly, redistribution may be total, where the time redistributed is entirely shifted to other applications and there is no overall improvement in productivity.

A third possibility is that saving time on one application breaks a habit loop [22] and reduces time spent on other applications as well, so the actual net improvement in productivity is even better than just what is saved on the target application. We will refer to this as the *reduction hypothesis*. Returning to our example of a productivity intervention targeting Facebook, this would hypothesize that an intervention that reduces time on Facebook will also reduce time on Instagram. Perhaps once we enter "procrastination mode" and visit one unproductive application, we wind up chaining together visits to another unproductive application, and another—but if a productivity intervention helps us break the chain early on, we will never visit the later unproductive applications.

These three hypotheses lay out the three possibilities of what happens to other goals when we intervene on a focal goal (Figures 1–2): time on those other goals might stay the same (isolated effects), go up (redistribution), or go down (reduction). In this paper, we seek to adjudicate between these hypotheses using HabitLab [36], an in-the-wild productivity experimentation environment that users can voluntarily participate in by installing. Prior work described HabitLab as a Chrome browser extension; in this paper we created and deployed a companion HabitLab Android application, allowing us to study any redistribution of time that might be happening across devices, as when a user avoids Facebook on their browser but ends up checking Facebook on their phone instead.

After installing and agreeing to our experimental protocol, users specify what they wish to reduce time on, which we term *goals*. In the case of the Android version, goals take the form of applications (apps), whereas on the Chrome extension goals are sites. We then deploy interventions to help users reduce their time on these goals, which can appear when the user visits a website (Chrome) or app (Android). To study redistribution, we periodically manipulate the frequency at which interventions appear for each goal — if the goal is in the frequent condition that week, it will appear every time the user visits that application, whereas if the goal is in the infrequent condition that week, it will appear on 20% of visits. This experimental design allows us to observe the effects of a goal being in the *frequent* setting not only on how much time users spend on that focal goal, but also what happens to time on other goals when that focal goal is in the frequent setting.

Our analysis first begins by seeing whether interventions are effective at reducing time on the focal goal, disregarding any possible redistribution effects. We do so by comparing time spent per day on the application on weeks where interventions are shown frequently, vs those weeks where interventions are shown infrequently. We find that they are effective, with time spent on goal sites reduced by 8.0% on the Chrome version, and time spent on goal apps reduced by 37.3% on the Android version.

Next, we investigate whether time is redistributed to other sites/apps on the same platform (browser or mobile) when interventions are frequently shown. We find that giving interventions within the browser produces a reduction effect, with users using sites/apps less when there are more interventions shown on other sites/apps – however, effects of interventions are isolated on mobile.

Finally, we investigate whether time is redistributed across devices. We do not observe any significant time redistribution effects in either direction.

This paper contributes a look into potential unintended side effects of productivity interventions on other sites, apps, and devices. We find that productivity interventions do not appear to have deleterious second-order effects on goals other than the ones they are targeting, and in some cases, may even have beneficial second-order effects by breaking habit loops.

2 BEHAVIOR CHANGE AND MOTIVATION

Persuasive technologies seek to produce behavioral change [26] across goals as diverse as sustainable resource consumption [28], sleep [13, 34], exercise [16], smoking [48], eating habits [23, 46], coping strategies [3, 54] and productivity [35, 36, 61].

They can operate on many different platforms, such as the web or mobile devices. Web-based systems promote a behavior change goals including classroom engagement [4, 5], psychology therapy [7] and healthy habits [18, 39]. In parallel, a number of studies focused on mobile-based interventions [25, 49, 52, 53, 60]. For instance, MyBehavior, a mobile phone app, was built to track physical activities of the users and to provide personalized suggestions that are tailored to the users' historical behavioral data [52]. Similarly, PopTherapy is a mobile phone app that studied micro-interventions for coping with stress [49].

There are a number of theoretical frameworks describing behavior change systems. B=MAT is a popular framework of behavioral change [26], which demonstrates that systems can focus on three elements—motivation, ability, and a trigger (a call to action)—to produce behavior change. The habit loop is another framework for building habits [24], stating that systems can build habits through an iterated process of displaying a trigger, prompting the user to take an action, giving out a reward, and helping the user to invest in the system.

Measuring the effectiveness of a persuasive system remains a major challenge in the design of behavior change systems. While behavior change systems can be effective during experiments [7, 19, 59], many review papers are more restrained in whether behavior change systems remain effective outside studies and bring longitudinal behavioral change [10, 29, 44, 45]. Because behavior changes are long and complex processes, the efficacy of a persuasive system is often difficult to measure [51]. For instance, an intervention promoting healthy habits, which was effective in changing participants' eating habits, might reduce their physical activities, which were not measured in the experiment [17]. Likewise, a system promoting increased physical activity may be unable to observe effects on participants' eating habits [21]. Compared to prior work, our study examines these spillover effects in the context of a more complete ecosystem, including both desktop browsers and mobile devices.

Cyberslacking, referred to as non-work-related computing, is the use of Internet and mobile technology during work hours for personal purposes [33, 38, 50, 58]. One study found that employees spent at least one hour on non-work-related activities during a regular work day [58]. Researchers also reported that non-work-related Internet usage comprises approximately 30%–50% of total usage [1, 31].

Unproductive time begets further unproductive time. For example, increased time spent online can increase sleep debt, which in turn leads to more time spent online [43]. Likewise, the Hook Model claims that many of the most addictive online sites use a cycle of investment techniques to keep users coming back—for example, making a post on Facebook may result in future notifications, which will in turn will get the user to come back and make more posts [24]. Finally, sites such as Facebook, Reddit, Twitter, and Buzzfeed are filled with links to each others' content, so it may be the case that increasing usage of one will increase usage of others. If productivity interventions are able to break this vicious cycle of procrastination for one application, they may actually reduce time spent on other unproductive applications as well.

The importance of understanding the effectiveness of productivity interventions in a complete ecosystem and the rising awareness of unproductive time spent on mobile devices call into focus: would productivity interventions reduce net unproductive time? Or is it a weak palliative with little discernible effect? This led to our research question:

Research Question (RQ). *Do productivity interventions reduce net unproductive time, or just redistribute it to other applications, sites, and devices?*

3 DISTRIBUTION OF UNPRODUCTIVE TIME

In this section, we will examine related studies in behavior change systems to develop testable hypotheses regarding the research question.

Multitasking has become ubiquitous in today's workplaces [6, 12, 40]. Multitasking is both essential and unavoidable in the workplace [27, 41], and it takes 11 minutes on average before people switch to a new task [20].

Studying behavior change effects across multiple devices is important: focusing on a single platform will myopically miss unproductive behaviors on other platforms. Attention is fragmented in both mobile and traditional desktop environments [37, 40]. The time spent on mobile devices has increased more rapidly than time on computers or TVs [9, 15]. On the other hand, mobile applications have been regarded as substitutions of websites in many studies [57]. Large technology companies such as Facebook and Amazon have been focusing on user growth on mobile devices [37].

However, interventions may result in unintended outcomes [29, 30, 56]. Specifically, while some interventions may be highly effective at achieving the measured goal of a behavioral change system, they may reduce desired outcomes elsewhere [29]. In one health-related intervention, while the physical activity of participants increased, calorie intake also increased, working against the goal of promoting a healthy lifestyle [11]. Similarly, using peer pressure to build confidence for students at school would, in turn, lower their self-esteem which actually was opposite to the goal of augmenting confidence [56].

In our system, the time spent on unproductive activities might be decreased in one application yet increased in others. These prompt our hypotheses:

Hypothesis 1 (H1). *Within a single device, productivity interventions will cause the time spent on targeted sites and apps to be redistributed to other sites and apps.*

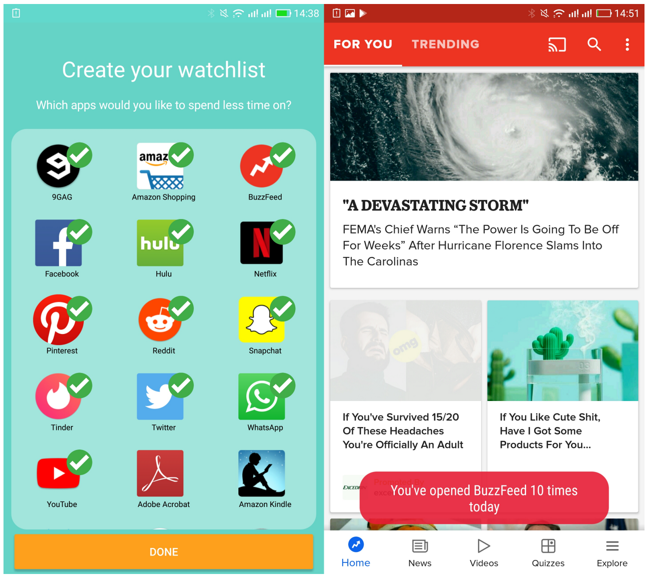

Figure 3: Screenshots from the mobile version of HabitLab. Left: The goal selection screen, where users choose which apps to spend less time on.
Right: An example intervention, which shows the visit count when a user opens a goal app.

Hypothesis 2 (H2). *Between computers and mobile devices, productivity interventions will cause the time spent on one device to be redistributed to other devices.*

4 BEHAVIOR CHANGE PLATFORM: HABITLAB

To gain insight into possible redistribution effects in behavior change, we created and deployed HabitLab [36], an open-source[1] platform which contains a variety of productivity interventions. Our prior work on HabitLab focused only on in-browser interventions, with the goal of studying intervention rotation strategies. With this paper, we track time redistribution and introduce an Android app, allowing us to track redistribution not just within platforms but across platform boundaries as well.

There are two versions of HabitLab: a Chrome extension, and an Android app. Both follow the structure of allowing users to choose what they wish to spend less time on (setting goals), and deploying interventions to meet those goals. On the Chrome version, users choose sites to spend less time on (goal sites – for example, facebook.com), as shown in Figure 4. On Android, users choose particular apps to spend less time on (goal apps – for example, the Facebook Android app), as shown in Figure 3. Interventions are deployed when users visit a goal site on Chrome (Figure 5), and when users open a goal app on Android, as shown in Figure 3.

[1]HabitLab is available at http://habitlab.github.io.

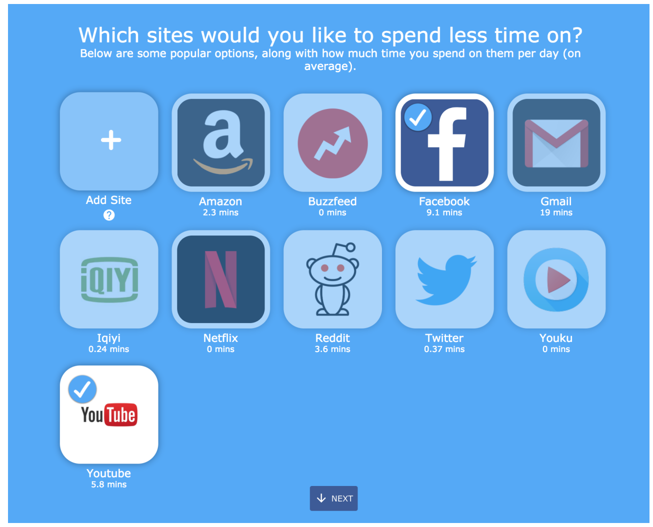

Figure 4: The goal selection screen, where users choose which sites to spend less time on (browser version).

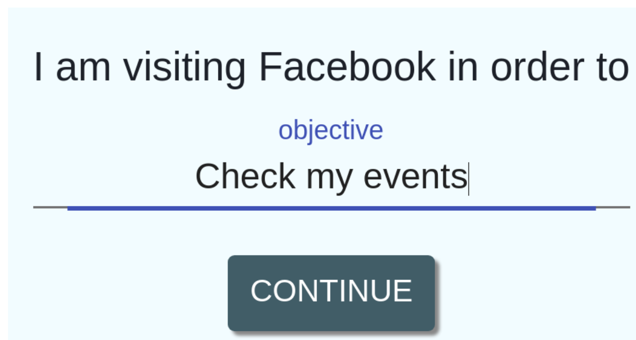

Figure 5: An example intervention, which asks a user to write their objective for visiting a site (browser version).

Mobile and Browser version Differences

The Chrome extension and Android app differ in some minor details. They support different sets of goals: users select apps to reduce time on in the Android version, whereas users choose sites to reduce time on in the Chrome version. Additionally, the specific set of interventions available differs between the platforms to fit the design languages of the browser and the mobile phone. The Chrome version has certain interventions which are site-specific – such as a news feed remover that is specific to Facebook. However, because Android does not allow applications to edit each other's view trees, the Android version's interventions are all glass pane overlays, and thus are general and can be used on any app. The concept of a session is different on the platforms: in the Chrome version, a session is time on a site until that tab is either closed or the user goes to a different domain. Time measured is active time – so if the tab is not focused, or if there is no keyboard or mouse activity for over a minute, the timer is temporarily paused. However, on Android, because there is no concept of a tab, the measurement of a session is different. There, a session is considered the duration over which an app is opened and focused. Closing the app, switching to a different app, or turning off the phone will end the current session.

The design of HabitLab's interventions is based on theories such as Cialdini's factors of influence [14] and the behavior change wheel taxonomy of behavior change interventions [2]. Description of the interventions on the Chrome and Android versions can be found in the Appendix.

As of writing, the Chrome version has over 8000 daily active users, and the Android version has over 500 daily active users. The users were not explicitly recruited, but were rather all organic installs who discovered the extension/app via sources such as the Chrome/Play store, or were referred to it via press coverage in sources such as Wired or the New York Times.

5 STUDY: REDISTRIBUTION OF TIME WITHIN AND ACROSS DEVICES

In this study we aim to analyze whether productivity interventions are reducing or redistributing time. We pursue this through an experiment and three sets of analyses: *(1) Within-device redistribution of time, in the browser.* For example, this would be the effects on time spent on non-Facebook websites, due to interventions that run when visiting the Facebook website. *(2) Within-device redistribution of time, on mobile devices.* For example, this would be the effects of time spent on non-Facebook applications, due to interventions that run when using the Facebook app. *(3) Cross-device redistribution of time.* For example, this would be the effects of time spent on Facebook on the phone, due to interventions that run when visiting the Facebook website.

Participants

Participants in this study consisted of new HabitLab users who installed either the HabitLab Chrome extension or Android app over a period of 132 days (approximately 19 weeks) in July through December 2018. 3747 users installed the HabitLab Chrome version over the course of our experiment and consented to our research protocol. 1483 users did so for the Android version. 298 installed both and signed in with their Google accounts, allowing us to analyze their usage across devices. We discarded participants who were not new users of HabitLab, since some users were re-installs or new devices for existing users. We also discarded participants who did not complete the onboarding process, or who uninstalled the system before they saw their first intervention. This left us with 1790 participants for Chrome, 782 participants for Android, and 82 participants for whom we could analyze usage across both. A summary of our dataset is shown in Table 1.

Table 1: Data Summary. Note that the duration of 132 days are users who kept it installed the longest, but as users can freely install/uninstall we do not have 132 days of data on all users.

	Browser	Android	Synced
Time Duration	132 days	132 days	132 days
No. of Users	1790	782	82
No. of Sessions	4.8 million	11.3 million	3.8 million

Method

In order to observe time redistribution effects between a focal goal and other goals due to interventions, we would ideally randomly turn interventions on and off for goals, then observe the effects on other goals. However, because HabitLab informs users that it will show interventions on goals that they select, there would be negative consequences (e.g., user confusion and dissatisfaction) if interventions for an application disappeared entirely for a week. Therefore, we opt to vary frequency rather than entirely turn off interventions for a goal each week.

So, for each goal on each device, we randomize frequency of interventions each week. On weeks where a goal is set as frequent, an intervention is shown on every visit to the app or site. On weeks where a goal is set as infrequent, an intervention is shown with probability 0.2 on every visit to the app or site. We choose this methodology of varying frequency to approximate the effects of turning interventions entirely on or off.

We analyze the effects interventions have on overall time spent on goals in the browser and mobile environments. We do so with a linear mixed model, which models the relationship between a dependent variable of time spent that day on a goal, an independent variable of goal frequency (frequent or infrequent), and categorical variables for the user and the goal site or app (e.g., Facebook, YouTube, Reddit) as random effects. We run the model separately on both the data from the browser and mobile versions. Our results here can also be replicated with a simpler model of an independent sample t-test modeling the effects of frequency on time spent.

Intensity

Frequency measures how much a user is being nudged in a single goal, but our experiment also needs to measure how much a user is being nudged overall, across all goals on the platform. This allows us to, for example, measure whether mobile device usage increases when browser interventions are overall more frequent, or whether time spent on non-goal sites increases when interventions are more frequent on goal sites. So, we define a measure of *intensity*: the percentage of sessions on any goal that triggered an intervention. For example, if the goal apps are Facebook and YouTube, the user visited Facebook 10 times and saw interventions 2 times, and visited YouTube 3 times and saw interventions 3 times, then the intensity is $\frac{5}{13} = .38$. Intensity will naturally vary over time as goals are re-randomized into *frequent* and *infrequent* conditions, with more frequent goals increasing intensity and more infrequent goals decreasing intensity. This randomization occurs for all goals simultaneously, once a week. We chose this intensity metric for our analysis, as opposed to alternatives such as raw number of times interventions were seen, because: 1) it is independent of the dependent variable, total time spent; 2) it is independent of the number of times the user visits a site/app; 3) it is guaranteed to be between 0 to 1, which is useful for interpretation; and 4) it can be used for both within-device and cross-device analysis.

For each goal, we also define a measure of *intensity of other goals*. This is the intensity measure excluding the current goal. We will use it for analyzing redistribution of time within device: when intensity of other goals varies, what is the effect on time spent on a target goal?

Time Redistribution

Within Device. We analyze the effects of interventions on time redistribution within device. We define *time redistribution within device* as an increase in time spent on the goal on the device, as a result of a change in intensity of other goals. For example, an increase in time spent on YouTube as a result of turning Facebook interventions on would be an example of time redistribution from Facebook to YouTube.

We do so with a linear mixed model, which models the relationship between a dependent variable of time spent that day on all goals, an independent variable of intensity of goals, as well as the user as a random effect. We run the model separately on both the data from the browser and mobile versions. Because our time data is log-normally distributed, we fit our linear mixed models to log time.

Across Device. We analogously define time redistribution between devices as an increase in time spent on the other device, as a result of interventions increasing in frequency in the other device. For example, an increase in time spent on Facebook on the browser, as a result of increasing the frequency of interventions on mobile would be an example of time being redistributed from mobile to browser.

We do so with a linear mixed model, which models the relationship between a dependent variable of time spent that day on all goals on one device, an independent variable of intensity of goals on the other device, and the user as a random effect. We run the model separately on data in both directions: one analyzing the effects of browser intensity on time spent on mobile, and another analyzing the effects of mobile intensity on time spent on the browser. We again log transform our time data for analysis.

Table 2: Browser: Frequent interventions for a goal site cause a reduction of time spent on the site.

	Dependent variable:
	Log daily time on site
Frequent (1=true)	-0.085^{***}
	(0.010)
Baseline	5.904^{***}
	(0.224)
Observations	96,489
Note:	*p<0.1; **p<0.05; ***p<0.01

6 RESULTS

First, we establish that our interventions are effective – that is, increasing the frequency of intervention on a goal app reduces time on that app. Next, we confirm that increasing intensity on a device reduces time on goal apps on that device. Then, we analyze redistribution effects within device – that is, whether increasing intensity effects time on non-goal apps. We also analyze redistribution effects across devices – that is, whether increasing intensity on one device effects time on goal apps on the other device. Finally, we build intuition for the underlying mechanisms by exploring what happens after users visit goal applications.

Are interventions effective?

Browser. **Yes**. We look at the effect of frequency of interventions on time spent on a day on a site, controlling for the user and the goal. We find a significant reduction in time spent on day on an app, when interventions for that goal are frequently shown that day ($p < 0.001$), as shown in Table 2. Estimated log time on a goal when infrequent is 5.747 (313 seconds), while for frequent goals this is reduced to 5.665 (288 seconds). Hence, our methodology of increasing intervention frequency is effective at reducing time on sites.

Mobile. **Yes**. We look at the effect of frequency of interventions on time spent on a day on an app, controlling for the user and the goal. We find a significant reduction in time spent on day on an app, when interventions for that goal are frequently shown that day ($p < 0.001$), as shown in Table 3. Estimated log time on a goal when infrequent is 5.928

Table 3: Mobile: Frequent interventions for a goal app cause a reduction of time spent on the app.

	Dependent variable:
	Log daily time on app
Frequent (1=true)	-0.045^{***}
	(0.011)
Baseline	5.254^{***}
	(0.057)
Observations	96,147
Note:	*p<0.05; **p<0.01; ***p<0.001

(375 seconds), while for frequent goals this is reduced to 5.462 (235 seconds). Hence, our methodology of increasing intervention frequency is effective at reducing time on apps.

Is time spent on goals reduced when there is higher intensity?

In the previous analysis we have shown that increasing frequency of interventions on a single goal allows us to observe reductions in time spent on that goal, on both the browser and mobile platforms. In this section we will show that increasing intensity also allows us to observe reductions in total time spent on all goal apps, on both platforms. This allows us to confirm the validity of our intensity metric, as well as allow us to analyze the aggregate usage of all goal apps on each device. This will be necessary for our later analyses of redistribution effects within device as well as between devices.

Browser. **Yes**. We look at the effect of intensity on total time spent on goal sites each day, controlling for the user. We find a significant reduction in total time spent on goal sites when intensity is higher that day ($p < 0.001$), as shown in Table 4. Estimated log total time on goal sites with low intensity is 6.885 (978 seconds), while with high intensity this is reduced to 6.758 (861 seconds). Hence, when interventions are more frequent in aggregate on the browser (which intensity captures), overall time on goal sites is reduced.

Mobile. **Yes**. Like the browser, we look at the relationship between increasing intensity on one's mobile phone and the total time spent that day on one's goal applications. We find a significant decrease ($p < .05$) in goal time spent, as shown in Table 5. Estimated log total time on goal apps with low intensity is 8.146 (3450 seconds), while with high intensity this is reduced to 8.031 (3075 seconds). Hence, when interventions are more frequent in aggregate on mobile (which intensity captures), overall time on goal apps is reduced.

Table 4: Browser: Increasing intensity results in a reduction of time spent each day on all goal domains

	Dependent variable:
	Log daily time spent on all goal sites
Browser Intensity	-0.187^{***}
	(0.016)
Baseline	6.929^{***}
	(0.033)
Observations	57,204
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5: Mobile: Increasing intensity results in a reduction of time spent each day on all goal apps

	Dependent variable:
	Log daily time spent on all goal apps
Mobile Intensity	-0.049^{*}
	(0.025)
Baseline	8.300^{***}
	(0.042)
Observations	22,970
Note:	*p<0.05; **p<0.01; ***p<0.001

What is the effect of increasing intensity on other, non-goal apps and sites?

Browser. **Reduction**. We look at the effect of intensity on total time spent on non-goal sites each day, controlling for the user. We find a significant reduction in total time spent on non-goal sites when intensity is higher that day ($p < 0.000005$), as shown in Table 6. Estimated log total time on non-goal sites when intensity=0 is 8.207 (3667 seconds), while when intensity=1 this is reduced to 8.038 (3096 seconds). This is the effect predicted by our global reduction hypothesis.

Mobile. **No effect (isolation).** We do not observe a significant effect of Android intensity on time outside of goals, as shown in 7. This suggests that reducing time within Android is an “isolated” behavior. Note there is an insignificant trend towards increasing time on non-goal sites with increasing intensity ($p = 0.07$).

Is time redistributed between devices?

Mobile to Browser. **No effect (isolation)**. We look at the effect of mobile intervention intensity, on total time spent

Table 6: Browser: Increasing intensity results in a reduction of time spent each day on non-goal sites

	Dependent variable:
	Log daily time spent on all non-goal sites
Browser Intensity	-0.169^{***}
	(0.016)
Baseline	8.207^{***}
	(0.028)
Observations	57,204
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 7: Mobile: Increasing intensity has no significant effect of time spent on non-goal apps.

	Dependent variable:
	Log daily time spent on non-goal apps
Mobile Intensity	0.035
	(0.020)
Baseline	9.277^{***}
	(0.044)
Observations	22,970
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 8: Mobile: Varying intervention intensity has no effect on total time spent on browser goal sites

	Dependent variable:
	Log daily time spent on browser goals
Mobile Intensity	0.045
	(0.218)
Baseline	6.736^{***}
	(0.251)
Observations	1,312
Note:	*p<0.05; **p<0.01; ***p<0.001

on browser. We find no significant effect (p>.5), as shown in Table 8.

Browser to Mobile. **No effect (isolation)**. We look at the effect of browser intervention intensity, on total time spent on mobile. We find no significant effect (p>.5), as shown in Table 9.

Table 9: Browser: Varying intervention intensity has no effect on total time spent on mobile goal apps

	Dependent variable:
	Log daily time spent on mobile goals
Browser Intensity	0.064
	(0.068)
Constant	8.219^{***}
	(0.128)
Observations	1,312
Note:	*p<0.05; **p<0.01; ***p<0.001

Destination tracking

Finally, to build intuition as to the mechanism by which the above effects are happening, we analyzed what happens after users their goal applications. We visualized the flow of sessions from the 10 most widely chosen goal apps and sites in our dataset as Sankey diagrams (Figures 6 and 7). On mobile, a majority of sessions end up going to another application, followed by turning off the phone, as shown in Figure 6. On browsers, the majority of sessions went to other sites, as shown in Figure 7. We can also observe differences in goals users choose on mobile as opposed to desktop – on mobile, the most popular apps tend to be messaging apps, whereas on the browser they tend to be content aggregators.

7 LIMITATIONS

Our methodology varied frequency of interventions, instead of comparing having interventions completely on vs completely off. This approach reduces the size of effects we can observe compared to having interventions completely on or completely off. Our approach is also sensitive to variance in the effectiveness levels of the interventions. Some interventions may be more aggressive than others and change users' behavior more drastically even with low frequency. This difference may alter time re-distributions due to varied frequency.

We did not measure time spent on platforms that HabitLab does not support. For instance, HabitLab users may use Facebook on tablet devices, watch TV or engage in other activities that are considered unproductive aside from browsing on a desktop or on an Android phone. These behaviors may potentially change how time redistributed, but we are unable to track it.

Additionally, our study explores time redistribution in the context of productivity. It is possible that this context may not generalize to other behavior change regimes.

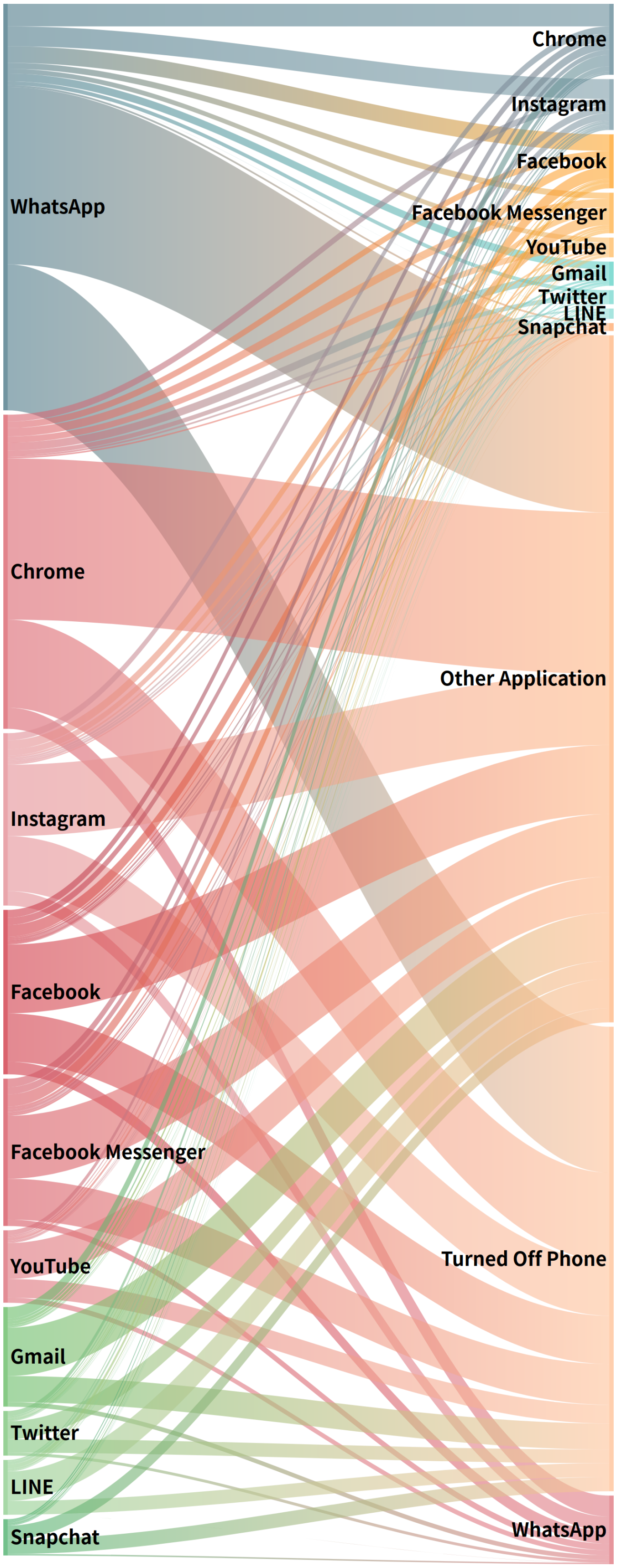

Figure 6: The top 10 goal apps with the most number of sessions on mobile are on the left. On the right is the distribution of where a user ends up immediately after.

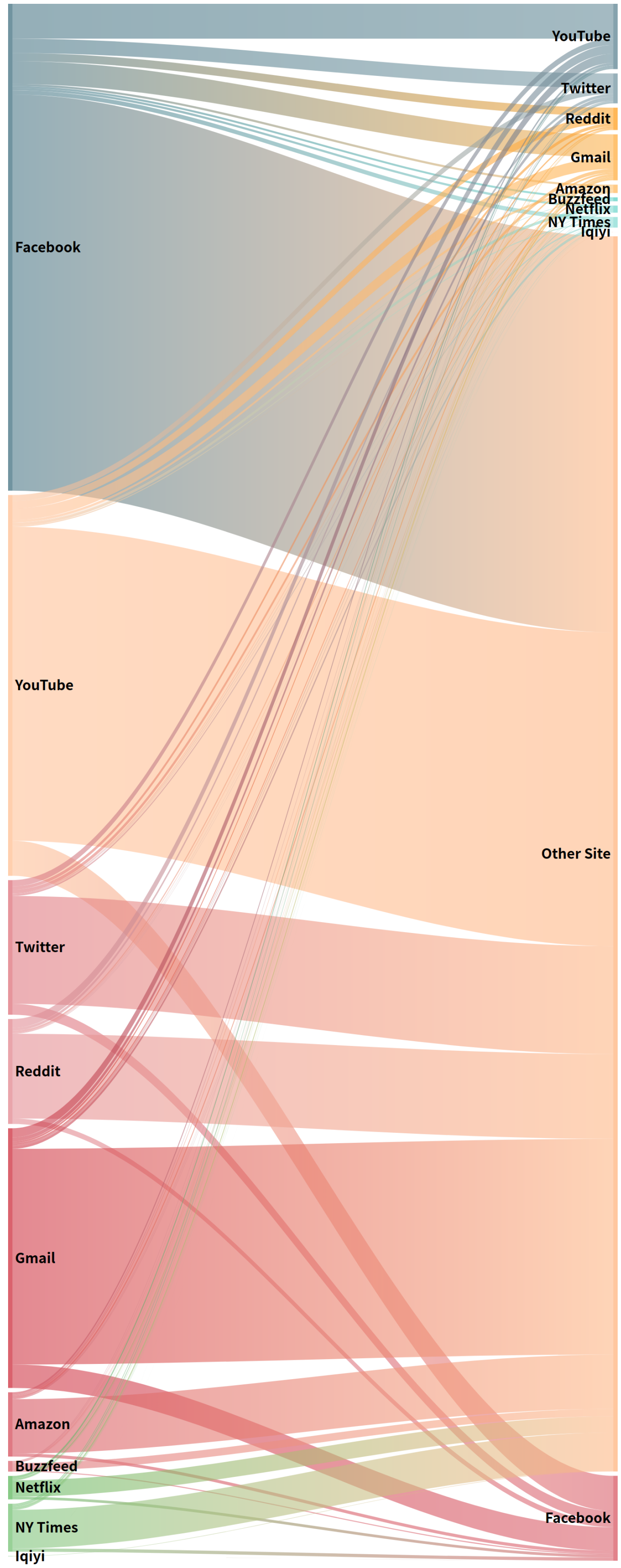

Figure 7: The top 10 goal apps with the most number of sessions on the browser are on the left. On the right is the distribution of where a user ends up immediately after.

8 DISCUSSION

We found that productivity interventions on the browser also reduced time on sites other than the targeted sites, but there was no such effect on mobile or cross-device.

We believe the reason we observed reduction in time on non-goal sites on the browser is several of the most popular goal sites — such as Facebook, Reddit, Twitter — are filled with hyperlinks to other sites, and hence drive traffic to them. For example, if an intervention makes a user spend less on their Facebook feed, they are going to be less likely to stumble upon a New York Times article, hence the Facebook-reducing intervention may also reduce time on New York Times. Part of this may be a difference in how mobile applications work, compared to websites. Several mobile applications embed a web browser so that even if the user clicks a link, it will open within the same app. For example, Facebook is one such app, so if the user clicks on a New York Times link within the Facebook app, it is opened within the Facebook app's built-in browser, so the time they spend reading that article will still be counted towards Facebook app usage.

One possible reason for differences between mobile and web is that the apps users choose to reduce time on in each two platform differ (e.g., messaging apps on Android vs. link aggregators on Chrome). There also exist differences in typical interaction styles (short, notification-driven sessions on Android [47], vs. longer sessions resulting from self-interruption on Chrome). 85% of the apps that Android users frequently chose to reduce time on are for messaging (WhatsApp, Instagram, Facebook Messenger, Twitter, LINE, Snapchat), where a characteristic interaction is receiving a message, unlocking the phone to read it and reply, then turning off the screen (as shown in Figure 6). Thus, users would not be drawn to other apps during this interaction. In contrast, with the Chrome version, the most selected sites are Facebook, YouTube, Reddit, and Twitter, 75% of which are aggregators of links to other sites. The number of daily sessions per app is also greater on Android, though sessions are longer on average on Chrome, and stopping using the browser after a session ends occurs less on Chrome. Thus, the browser-based interactions users were using HabitLab to reduce are not short messaging-driven spurts that end with turning off the screen as on mobile, but rather long sessions of surfing through link aggregators ending with going to another site. So, a proposed mechanism: interventions short-circuit browsing long browser-based sessions, but mobile sessions are already short.

This work brings about implications for designing interventions. Namely, we should consider not only the immediate interaction and its immediately measurable effects, but its longer-term effects in the context of the broader workflow. For example, consider 2 interventions for Facebook: 1) asks users to return to the home screen, vs 2) asks users to turn off the screen. Assuming similar rates of compliance, we would expect that measuring the effects on time spent on Facebook in isolation will show no difference between them. However, if we consider that going to the home screen can lead to users opening other apps, we might predict that a holistic measurement that includes effects on other apps as well will prefer 2) over 1). Or if designing interventions to reduce snacking, should we: a) ask participants to not eat anything until their next meal, or b) give them gum instead? While calorie intake from the immediate interaction would favor a), b) may prevent future snacking down the line. That said, in many cases, interventions are indeed isolated in their effects, and can even have beneficial effects elsewhere.

9 CONCLUSION

In this paper we have compared three hypotheses for how productivity interventions influence time spent on sites, apps, and devices other than the ones they are targeting. Productivity interventions may have no effect on other goals (*isolated effects*), they may cause time to be redistributed to other unproductive goals (*redistribution*), or they may cause a reduction in time spent on other unproductive goals (*reduction*).

We adjudicated between these hypotheses by varying the frequency of productivity interventions on goals that users set in the HabitLab browser extension and mobile app. When interventions were more frequent, users spent less time on their goal sites and apps, showing that the productivity interventions were effective. We also defined a metric of intensity that captures frequency of interventions within device, and investigated the effects of varying intensity of interventions for other apps/sites, on time spent on an app/site. The result differed by device: on the browser we observed a global reduction effect, with time on non-goal sites decreasing with increasing intensity of interventions. However, on mobile we observed no effect. We believe these differences are caused by differing usage patterns and platform differences: websites drive traffic to other websites via hyperlinks, but mobile apps try to keep users remaining on the app.

We have shown that while productivity interventions can sometimes have effects on usage of other, non-targeted sites and apps, they are often isolated in their effects. Hence, when designing for behavior change, while we should be careful about our measurements and the possibility of unintended side effects, in the context of productivity interventions it appears that targeting individual productivity goals does not cause substantial negative second-order effects.

10 ACKNOWLEDGEMENTS

This work was supported by a Stanford Human-Centered Artificial Intelligence seed grant. We thank the many users who have used and contributed ideas and feedback to HabitLab.

REFERENCES

[1] 2011. Yielding to (cyber)-temptation: Exploring the buffering role of self-control in the relationship between organizational justice and cyberloafing behavior in the workplace. *Journal of Research in Personality* 45, 2 (2011), 247 – 251.

[2] Charles Abraham and Susan Michie. 2008. A taxonomy of behavior change techniques used in interventions. *Health psychology* 27, 3 (2008), 379.

[3] Phil Adams, Mashfiqui Rabbi, Tauhidur Rahman, Mark Matthews, Amy Voida, Geri Gay, Tanzeem Choudhury, and Stephen Voida. 2014. Towards personal stress informatics: Comparing minimally invasive techniques for measuring daily stress in the wild. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 72–79.

[4] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2013. Steering user behavior with badges. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 95–106.

[5] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2014. Engaging with massive online courses. In *Proceedings of the 23rd international conference on World wide web*. ACM, 687–698.

[6] Steven H. Appelbaum, Adam Marchionni, and Arturo Fernandez. 2008. The multi-tasking paradox: perceptions, problems and strategies. *Management Decision* 46, 9 (2008), 1313–1325.

[7] Azy Barak, Liat Hen, Meyran Boniel-Nissim, and Na'ama Shapira. 2008. A Comprehensive Review and a Meta-Analysis of the Effectiveness of Internet-Based Psychotherapeutic Interventions. *Journal of Technology in Human Services* 26, 2-4 (2008), 109–160. https://doi.org/10.1080/15228830802094429 arXiv:https://doi.org/10.1080/15228830802094429

[8] Roy E Baumeister, Ellen Bratslavsky, Mark Muraven, and Dianne M Tice. 1998. Ego Depletion: Is the Active Self a Limited Resource? *Journal of Personality and Social Psychology* 74, 5 (1998), 1252–1265.

[9] Chris Beer. 2017. Multi-Device Watching Among TV Viewers. https://blog.globalwebindex.com/chart-of-the-day/multi-device-watching/

[10] Ashly D. Black, Josip Car, Claudia Pagliari, Chantelle Anandan, Kathrin Cresswell, Tomislav Bokun, Brian McKinstry, Rob Procter, Azeem Majeed, and Aziz Sheikh. 2011. The Impact of eHealth on the Quality and Safety of Health Care: A Systematic Overview. *PLOS Medicine* 8, 1 (01 2011), 1–16. https://doi.org/10.1371/journal.pmed.1000387

[11] S N Blair, D R Jacobs, and K E Powell. 1985. Relationships between exercise or physical activity and other health behaviors. 100 (03 1985), 172 – 180.

[12] L. Mark Carrier, Nancy A. Cheever, Larry D. Rosen, Sandra Benitez, and Jennifer Chang. 2009. Multitasking across generations: Multitasking choices and difficulty ratings in three generations of Americans. *Computers in Human Behavior* 25, 2 (2009), 483 – 489. https://doi.org/10.1016/j.chb.2008.10.012 Including the Special Issue: State of the Art Research into Cognitive Load Theory.

[13] Eun Kyoung Choe, Sunny Consolvo, Nathaniel F Watson, and Julie A Kientz. 2011. Opportunities for computing technologies to support healthy sleep behaviors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3053–3062.

[14] Robert B Cialdini and Nathalie Garde. 1987. *Influence*. Vol. 3. A. Michel.

[15] The Nielsen Company. 2009. Television, Internet and mobile usage in the US. Three Screen Report. http://www.nielsen.com/us/en/insights/news/2009/tv-internet-and-mobile-usage-in-us-continues-to-rise.html

[16] Sunny Consolvo, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, Ian Smith, and James A. Landay. 2008. Activity Sensing in the Wild: A Field Trial of Ubifit Garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 1797–1806. https://doi.org/10.1145/1357054.1357335

[17] Alexander P. Cotter, Nefertiti Durant, April A. Agne, and Andrea L. Cherrington. 2014. Internet interventions to support lifestyle modification for diabetes management: A systematic review of the evidence. *Journal of Diabetes and its Complications* 28, 2 (2014), 243 – 251. https://doi.org/10.1016/j.jdiacomp.2013.07.003

[18] Brian Cugelman. 2013. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games* 1, 1 (2013), e3.

[19] Pim Cuijpers, Annemieke van Straten, and Gerhard Andersson. 2008. Internet-administered cognitive behavior therapy for health problems: a systematic review. *Journal of Behavioral Medicine* 31, 2 (01 Apr 2008), 169–177. https://doi.org/10.1007/s10865-007-9144-1

[20] Laura Dabbish, Gloria Mark, and Víctor M González. 2011. Why do i keep interrupting myself?: environment, habit and self-interruption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3127–3130.

[21] Walter Dempsey, Peng Liao, Pedja Klasnja, Inbal Nahum-Shani, and Susan A Murphy. 2015. Randomised trials for the Fitbit generation. *Significance* 12, 6 (2015), 20–23.

[22] Charles Duhigg. 2012. *The power of habit: Why we do what we do in life and business*. Vol. 34. Random House.

[23] Daniel A Epstein, Felicia Cordeiro, James Fogarty, Gary Hsieh, and Sean A Munson. 2016. Crumbs: Lightweight Daily Food Challenges to Promote Engagement and Mindfulness. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*.

[24] Nir Eyal. 2014. *Hooked: How to build habit-forming products*. Penguin Canada.

[25] Brianna S. Fjeldsoe, Alison L. Marshall, and Yvette D. Miller. 2009. Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service. *American Journal of Preventive Medicine* 36, 2 (2009), 165 – 173. https://doi.org/10.1016/j.amepre.2008.09.040

[26] Brian J Fogg. 2002. Persuasive technology: using computers to change what we think and do. *Ubiquity* 2002, December (2002), 5.

[27] David Freedman. 2007. What's Next: Taskus Interruptus. https://www.inc.com/magazine/20070201/column-freedman.html

[28] Jon Froehlich, Tawanna Dillahunt, Predrag Klasnja, Jennifer Mankoff, Sunny Consolvo, Beverly Harrison, and James A Landay. 2009. UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1043–1052.

[29] Juho Hamari, Jonna Koivisto, and Tuomas Pakkanen. 2014. Do Persuasive Technologies Persuade? - A Review of Empirical Studies. In *Persuasive Technology*. Springer International Publishing, Cham, 118–136.

[30] Bernard J Healey and Robert S Zimmerman. 2010. *The new world of health promotion : new program development, implementation, and evaluation*. Jones and Bartlett Publishers.

[31] Hasmida Jamaluddin, Zauwiyah Ahmad, Mazni Alias, and Maimun Simun. 2015. Personal Internet Use: The Use of Personal Mobile Devices at the Workplace. *Procedia - Social and Behavioral Sciences* 172 (2015), 495 – 502. https://doi.org/10.1016/j.sbspro.2015.01.391 Contemporary Issues in Management and Social Science Research.

[32] Jing Jin and Laura A Dabbish. 2009. Self-interruption on the computer: a typology of discretionary task interleaving. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1799–1808.

[33] Pamela R. Johnson and Julie Indvik. 2004. The Organizational Benefits of Reducing Cyberslacking in the Workplace. *Journal of Organizational Culture, Communications and Conflict* 8, 2 (2004). https://www.questia.com/library/journal/1G1-167255389/the-organizational-benefits-of-reducing-cyberslacking
[34] Matthew Kay, Eun Kyoung Choe, Jesse Shepherd, Benjamin Greenstein, Nathaniel Watson, Sunny Consolvo, and Julie A Kientz. 2012. Lullaby: a capture & access system for understanding the sleep environment. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 226–234.
[35] Young-Ho Kim, Jae Ho Jeon, Eun Kyoung Choe, Bongshin Lee, K Kim, and Jinwook Seo. 2016. TimeAware: Leveraging framing effects to enhance personal productivity. In *Proceedings of the SIGCHI conference on human factors in computing systems*.
[36] Geza Kovacs, Zhengxuan Wu, and Michael S Bernstein. 2018. Rotating Online Behavior Change Interventions Increases Effectiveness But Also Increases Attrition. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (Nov. 2018).
[37] Amanda Lenhart, Kristen Purcell, Aaron Smith, and Kathryn Zickuhr. 2009. Social media and mobile Internet use among teens and young adults. *Pew Internet and American Life Project* (2009).
[38] Vivien Lim. 2002. The IT Way of Loafing on the Job: Cyberloafing, Neutralizing and Organizational Justice. 23 (08 2002), 675 – 694.
[39] Elizabeth J Lyons, Zakkoyya H Lewis, Brian G Mayrsohn, and Jennifer L Rowland. 2014. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. *Journal of medical Internet research* 16, 8 (2014), e192.
[40] Gloria Mark. 2015. *Multitasking in the Digital Age.* Morgan and Claypool Publishers.
[41] Gloria Mark, Daniela Gudith, and Ulrich Klocke. 2008. The cost of interrupted work: more speed and stress. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 107–110.
[42] Gloria Mark, Shamsi T Iqbal, Mary Czerwinski, Paul Johns, Akane Sano, and Yuliya Lutchyn. 2016. Email duration, batching and self-interruption: Patterns of email use on productivity and stress. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 1717–1728.
[43] Gloria Mark, Yiran Department of Informatics Wang, Melissa Niiya, and Stephanie Reich. 2016. Sleep Debt in Student Life: Online Attention Focus, Facebook, and Mood. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5517–5528. https://doi.org/10.1145/2858036.2858437
[44] M Neve, Philip J Morgan, PR Jones, and CE Collins. 2010. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. *Obesity reviews* 11, 4 (2010), 306–321.
[45] Gregory J. Norman, Marion F. Zabinski, Marc A. Adams, Dori E. Rosenberg, Amy L. Yaroch, and Audie A. Atienza. 2007. A Review of eHealth Interventions for Physical Activity and Dietary Behavior Change. *American Journal of Preventive Medicine* 33, 4 (2007), 336 – 345.e16. https://doi.org/10.1016/j.amepre.2007.05.007
[46] Jon Noronha, Eric Hysen, Haoqi Zhang, and Krzysztof Z Gajos. 2011. Platemate: crowdsourcing nutritional analysis from food photographs. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 1–12.
[47] Antti Oulasvirta, Sakari Tamminen, Virpi Roto, and Jaana Kuorelahti. 2005. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 919–928.
[48] Jeni Paay, Jesper Kjeldskov, Mikael B Skov, Nirojan Srikandarajah, and Umachanger Brinthaparan. 2015. Personal Counseling on Smart Phones For Smoking Cessation. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 1427–1432.
[49] Pablo Paredes, Ran Gilad-Bachrach, Mary Czerwinski, Asta Roseway, Kael Rowan, and Javier Hernandez. 2014. PopTherapy: Coping with stress through pop-culture. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 109–117.
[50] Loo Geok Pee, Irene M.Y. Woon, and Atreyi Kankanhalli. 2008. Explaining non-work-related computing in the workplace: A comparison of alternative models. *Information and Management* 45, 2 (2008), 120 – 130. https://doi.org/10.1016/j.im.2008.01.004
[51] J O Prochaska and W F Velicer. 1997. The transtheoretical model of health behavior change. *American journal of health promotion : AJHP* 12 (1997), 38–48.
[52] Mashfiqui Rabbi, Angela Pfammatter, Mi Zhang, Bonnie Spring, and Tanzeem Choudhury. 2015. Automated Personalized Feedback for Physical Activity and Dietary Behavior Change With Mobile Phones: A Randomized Controlled Trial on Adults. *JMIR mHealth uHealth* 3, 2 (14 May 2015), e42. https://doi.org/10.2196/mhealth.4160
[53] William T Riley, Katrina J Serrano, Wendy Nilsen, and Audie A Atienza. 2015. Mobile and wireless technologies in health behavior and the potential for intensively adaptive interventions. *Current Opinion in Psychology* 5 (2015), 67 – 71. https://doi.org/10.1016/j.copsyc.2015.03.024 Health behavior.
[54] Pedro Sanches, Kristina Höök, Elsa Vaara, Claus Weymann, Markus Bylund, Pedro Ferreira, Nathalie Peira, and Marie Sjölinder. 2010. Mind the body!: designing a mobile stress management application encouraging personal reflection. In *Proceedings of the 8th ACM conference on designing interactive systems*. ACM, 47–56.
[55] Bryant A Stamford, Sharleen Matter, Ronald D Fell, and Paula Papanek. 1986. Effects of smoking cessation on weight gain, metabolic rate, caloric consumption, and blood lipids. *The American journal of clinical nutrition* 43, 4 (1986), 486–494.
[56] Agnis Stibe and Brian Cugelman. 2016. Persuasive Backfiring: When Behavior Change Interventions Trigger Unintended Negative Outcomes. In *Persuasive Technology*, Alexander Meschtscherjakov, Boris De Ruyter, Verena Fuchsberger, Martin Murer, and Manfred Tscheligi (Eds.). Springer International Publishing, Cham, 65–77.
[57] Alok Tongaonkar, Shuaifu Dai, Antonio Nucci, and Dawn Song. 2013. Understanding Mobile App Usage Patterns Using In-App Advertisements. In *Passive and Active Measurement*, Matthew Roughan and Rocky Chang (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 63–72.
[58] Jessica Vitak, Julia Crouse, and Robert LaRose. 2011. Personal Internet use at work: Understanding cyberslacking. *Computers in Human Behavior* 27, 5 (2011), 1751 – 1759. https://doi.org/10.1016/j.chb.2011.03.002 2009 Fifth International Conference on Intelligent Computing.
[59] L. Thomas Webb, Judith Joseph, Lucy Yardley, and Susan Michie. 2010. Using the Internet to Promote Health Behavior Change: A Systematic Review and Meta-analysis of the Impact of Theoretical Basis, Use of Behavior Change Techniques, and Mode of Delivery on Efficacy. *J Med Internet Res* 12, 1 (17 Feb 2010), e4. https://doi.org/10.2196/jmir.1376
[60] McRobbie H Bullen C Borland R Rodgers A Whittaker, R and Y Gu. 2012. Mobile phone-based interventions for smoking cessation. *Cochrane Database of Systematic Reviews* 11 (2012).
[61] Steve Whittaker, Vaiva Kalnikaite, Victoria Hollis, and Andrew Guydish. 2016. 'Don't Waste My Time': Use of Time Information Improves Focus. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 1729–1738.