



King's Research Portal

DOI:

[10.1016/j.ijhcs.2022.102869](https://doi.org/10.1016/j.ijhcs.2022.102869)

Document Version

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Lyngs, U., Lukoff, K., Csuka, L., Slovák, P., Van Kleek, M., & Shadbolt, N. (2022). The Goldilocks level of support: Using user reviews, ratings, and installation numbers to investigate digital self-control tools.

INTERNATIONAL JOURNAL OF HUMAN COMPUTER STUDIES, 166, [102869].

<https://doi.org/10.1016/j.ijhcs.2022.102869>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



The Goldilocks level of support: Using user reviews, ratings, and installation numbers to investigate digital self-control tools

Ulrik Lyngs^{*,a}, Kai Lukoff^b, Laura Csuka^a, Petr Slovák^c, Max Van Kleek^a, Nigel Shadbolt^a

^a Department of Computer Science, University of Oxford, Oxford, United Kingdom

^b DUB, University of Washington Seattle, Washington, United States

^c Department of Informatics, King's College London, London, United Kingdom

ARTICLE INFO

Keywords:

Attention
Distraction
Digital self-control
Digital wellbeing
Focus
Goldilocks principles
Self-regulation

ABSTRACT

Much effort has been invested in designing digital systems that keep people 'hooked'. By contrast, comparatively little is known about how designers can support people in re-gaining control. Online, however, hundreds of apps and browser extensions promise to help people self-regulate use of digital devices. Reviews and popularity metrics for these digital self-control tools (DSCTs) can indicate which design patterns are useful in the wild. Moreover, they reveal how platforms like Android and iOS differ in the ecosystems they enable for DSCTs, which has important implications for end users. We analysed reviews, installation numbers, and ratings for 334 DSCTs on the Google Play, Chrome Web, and Apple App stores, investigating what user reviews reveal about usage contexts and key design challenges, and how functionality relates to popularity metrics. Our thematic analysis of 1,529 reviews (sampled from a data set of 53,978 distinct reviews scraped in March 2019) found that DSCTs are seen as highly important for focusing on less instantly rewarding tasks when digital distractions are easily available. Users seek DSCTs that adapt to their personal definitions of distraction, and provide support that is sufficient to change behaviour without feeling too coercive. Reviewers suggested combining design patterns to provide a level of support that is 'just right'. This was mirrored in the ratings, where tools that combine different types of design patterns (e.g., website blocking and goal reminders) tended to receive higher ratings than those implementing a single type. We discuss implications for research and design, including how design patterns in DSCTs interact, and how psychological reactance to DSCTs can be reduced.

1. Introduction

The constant connectivity of smartphones and laptops provides innumerable benefits, but also unprecedented opportunities for distraction. Indeed, as the business models of many tech companies have come to rely on nudging people into using digital services extensively (cf. the "attention economy" (Davenport and Beck, 2001; Einstein, 2016; Wu, 2016)), many people routinely struggle to exercise self-control over their device use (Cox et al., 2016; Hiniker et al., 2016; Ko et al., 2015; Marotta and Acquisti, 2017; Tran et al., 2019). In HCI, a growing amount of research attempts to address this challenge by developing and

evaluating design patterns that help users self-regulate (Kim et al., 2019a; Lottridge et al., 2012; Lyngs et al., 2019; Monge Roffarello and De Russis, 2019; Pinder et al., 2019; Tseng et al., 2019a)¹. This work is gradually building an understanding of how interventions ranging from ambient visualisations of time spent (Whittaker et al., 2016) and distraction blocking (Kim et al., 2019a) to goal-setting with social support (Ko et al., 2015) can help people stay in control.

Most of this research has taken an experimental approach and compared user behaviour and perceptions when applying a given intervention to its absence and/or some alternate intervention (Lyngs, 2021a). In the present paper, we take an alternative approach. On online

* Corresponding author.

E-mail addresses: ulrik.lyngs@cs.ox.ac.uk (U. Lyngs), kai1@uw.edu (K. Lukoff), laura.csuka@cs.ox.ac.uk (L. Csuka), petr.slovak@kcl.ac.uk (P. Slovák), max.van.kleek@cs.ox.ac.uk (M. Van Kleek), nigel.shadbolt@cs.ox.ac.uk (N. Shadbolt).

¹ We use the term "design pattern" to refer to specific functions intended to support people in self-regulating digital device use (e.g., hiding newsfeeds, providing goal reminders). We use the term "digital self-control tool" (DSCT) to refer to apps and browser extensions that implement a given design pattern (e.g., a browser extension for hiding Facebooks newsfeed). We occasionally use the term "intervention" to refer to the implementation of a design pattern, in the context of how it changes a users usual digital environment, or when it is being evaluated in a research study.

stores for apps and browser extensions, hundreds of “digital self-control tools” (DSCTs) provide similar interventions that block apps or hide distracting website elements, track and visualise use, provide rewards or punishments for how devices are used, and much more (Lyngs et al., 2019). User reviews and popularity metrics from these tools might provide valuable complementary information about the utility of different design patterns and implementations “in the wild”, and help scope future controlled studies. However, a limited number of studies have investigated what we might learn from this source of data.

Hence, in this paper, we explore how user reviews and other store information can contribute to HCI research on design patterns for digital self-control. Specifically, we analyse user reviews, functionality, installation numbers, and ratings for 334 DSCTs on the Google Play, Chrome Web, and Apple App stores (drawn from Lyngs et al., 2019), investigating user needs, contexts and purposes of use, and key design challenges.

Our analysis of user reviews found that DSCTs are seen as highly important to support focus on less instantly rewarding tasks when digital distractions are easily available, especially in contexts related to productivity (including work, studying, reading, and writing), and going to sleep. DSCTs seemed particularly useful for reviewers describing themselves as “addicts” or otherwise struggling intensely with distraction, for example due to conditions like ADHD.

Reviewers wanted DSCTs to flexibly adapt to their personal definition of distracting use, as seen in common requests for customisation of *what* and *when* particular usage patterns should be blocked or tracked. One important reason was that DSCTs’ interventions, useful as they may be, can elicit negative emotional reactions. This could cause their interventions to be counterproductive, especially if users perceived DSCTs’ framing of distracting use to differ from their own.

Finally, reviewers wanted DSCTs to provide carefully calibrated support that is sufficient to support intended behaviour change without feeling too coercive. This ‘Goldilocks level’² varied between reviews, but most wanted DSCTs to provide support powerful enough to keep the user on track in moments of temptation or absentmindedness. Thus, blocking tools were often criticised if they were too easy to override because they could then fail to hold people accountable in moments of temptation. On the other hand, DSCTs were also criticised if their incentives were *too* strong, for example if blocking tools did not include some way to handle genuine emergency use, or if tools providing reward/punishment made users feel dejected when they failed to use their devices as intended.

One way DSCTs might provide properly calibrated support is by combining different types of design patterns (e.g., combining distraction blocking with goal-reminders). This was sometimes requested in the reviews, and tentatively supported by rating scores, where tools that combine two or more types of design patterns received higher ratings than those implementing a single type.

1.1. Motivation and background

In recent years, HCI researchers have started to build an understanding of how specific design patterns intended to support self-regulation over digital device can influence user behaviour and perceptions (Kim et al., 2019a; 2019b; Kovacs et al., 2018; Whittaker et al., 2016). However, basic questions still remain open. These include *conceptual challenges* (e.g., what outcome measures we should seek to

² After the fairy tale ‘Goldilocks and the Three Bears’, where a young girl tastes three different bowls of porridge and prefers the one that is neither too hot nor too cold, but ‘just right’. This concept of the ‘just right amount’ is often referred to as the ‘Goldilocks principle’. As an example, a widely cited paper on the relationship between screen time and mental well-being found that a balance point in between “too much” and “too little” tech use seemed optimal, which the authors referred to as the ‘digital Goldilocks hypothesis’ (Przybylski and Weinstein, 2017)

optimise, as simply measuring impact on ‘screen time’ is deemed inadequate by most researchers (Cecchinato et al., 2019; Hiniker et al., 2019; Lukoff, 2019)); related gaps in our understanding of *individual variation in user needs* (such as how widely people differ in what they consider ‘distracting’ use (Cecchinato et al., 2019; Lukoff, 2019)); and *practical implementation challenges* (such as how much ‘friction’ tools should apply to hold users accountable for their past preferences (Cox et al., 2016; Kim et al., 2019a; 2019b; Lyngs et al., 2018)).

One approach to addressing these questions is to evaluate new or existing interventions in lab and field studies. While this is appropriate and feasible for evaluating and comparing a small number of design patterns, it is difficult to scale to broadly assess a large numbers of patterns and implementations.

A complementary approach is to investigate the landscape of DSCTs available online. Here, hundreds of tools are exploring a wide range of strategies, from hiding or modifying content on distracting websites (e.g. *Newsfeed Eradicator* (JDev, 2019)) or limiting the amount of functionality available on home screens (e.g. the Android app *LessPhone Launcher* (Mohan, 2019)), to gamifying self-control by tying device use to the wellbeing of virtual creatures (e.g. *Forest* (Seekrtech, 2018)). Some of these tools have millions of users (Lyngs et al., 2019), and they effectively amount to a myriad of natural studies in which individuals self-experiment with tools that represent not only one or more design patterns, but specific ways to implement those patterns (Daskalova, 2018; Lee et al., 2017). Indicators of the outcome of these experiments include user reviews, store ratings, and installation numbers, which can provide valuable information about the usefulness of different design patterns and implementations.

Other areas of HCI research have benefited from such an approach. For example, in mental health research, thematic analysis of user reviews has been used to explore factors important for optimising user experience and support engagement in apps for cognitive behavioural therapy (Bakker et al., 2016; Huguet et al., 2016; Lui et al., 2017; Starwarz et al., 2018; Sucala et al., 2017). However, few studies on digital self-control have taken this route: To the best of our knowledge, four published studies have described and/or categorised functionality in samples of DSCTs available online (Biskjaer et al., 2016; Lyngs et al., 2019; Monge Roffarello and De Russis, 2019; van Velthoven et al., 2018): Velthoven et al. (van Velthoven et al., 2018) presented aims and features in 21 tools identified via the software recommendation platform *alternativeto.net*; Biskjaer et al. (2016) presented a taxonomy of functionality based on 10 tools from the Google Play, Chrome Web, and Apple App stores, and online tech magazines; Monge Roffarello and De Russis (2019) presented a taxonomy of functionality based on 42 mobile apps from the Google Play store; and Lyngs et al. (2019) presented a taxonomy of functionality based on 367 tools from the Google Play, Chrome Web and Apple App stores, which was mapped to psychological mechanisms in a dual systems model.

Only one of these studies analysed user reviews, namely Monge Roffarello and De Russis (2019), which included thematic analysis of 1, 128 reviews from 42 smartphone apps, and none incorporated popularity metrics. Moreover, Monge Roffarello & De Russis’ analysis of user reviews was restricted to 42 apps drawn from the Google Play store and did not specify how the reviews were distributed across apps with different functionality. Thematic analysis is likely to yield different results depending on how reviews are sampled across tools with different functionality, and depending on the analysis strategy (Braun et al., 2018; Chen et al., 2014). Therefore, additional analyses with broader sampling of tools, and transparent sampling of reviews from those tools, might be helpful.

On this background, we extended existing studies of DSCTs by (i) scraping installation numbers, ratings, and user reviews from a broad sample of tools across different stores and types of functionality, (ii) analysing reviews using ‘reflexive’ thematic analysis (Braun et al., 2018), (iii) analysing popularity metrics relationship to design patterns, and (iv) making our materials, analysis scripts and data openly available

in the interest of transparency and reusability (Haroz, 2019).

2. Methods

Data, analysis scripts, and materials for the present paper are available at <https://osf.io/cdqub/>.

We investigated the 367 digital self-control tools reviewed by Lyngs et al. (2019), as this paper provided the largest available review of DSCTs. Lyngs et al. identified these tools by downloading search results on the Chrome Web, Google Play, and Apple App stores in August 2018 for the terms “distraction”, “addiction”, “motivation”, and “self-control”, individually as well as prefixed by “internet”, “smartphone”, and “laptop”. From the initial results (4,890 distinct apps and extensions), they included tools explicitly designed to help people self-regulate digital device use, resulting in 367 tools (222 browser extensions from Chrome Web, 87 apps from Google Play, and 58 apps from the Apple App store).³

We adopted Lyngs et al.’s classification of functionality within these tools, which clustered functionality into four types of design patterns for digital self-control:

- *block/removal* (features for blocking distractions, such as temporarily locking the user out of specific apps, or removing them in the first place, such as by hiding recommended videos on YouTube)
- *self-tracking* (features for tracking and/or visualising device use, such as showing how much time the user has spent in particular apps or on devices overall, or letting the user start a countdown timer when they need to focus)
- *goal-advancement* (features for reminding the user of their usage goals, such as displaying pop-ups when a specific amount of time has been spent, or replacing the content of new browser tabs with to-do lists)
- *reward/punishment* (features that provide incentives for using devices in particular ways, such as rewarding the user with points/streaks when they do not use their phone)

2.1. Collecting tool information and user reviews

In March 2019, we scraped additional information about these tools, including publicly available user reviews, as well as installation numbers, ratings, and store descriptions. For apps on the Google Play store, we used the open-source `google-play-scraper` (Olano, 2018b) to scrape reviews and tool information. For apps on the Apple App store, we used the open-source `app-store-scraper` (Olano, 2018a) to scrape reviews, and wrote our own R script (using the packages `rvest` and `R Selenium`) to scrape tool information. As the `app-store-scraper` sometimes failed to retrieve all reviews visible on tools’ pages on the Apple App store, we wrote an additional R script to scrape reviews when manual inspection found some to be missing. For browser extensions on the Chrome Web store, we wrote our own R script to scrape both tool information and user reviews.

2.2. Data and analysis approach

2.2.1. User numbers and ratings. All stores displayed average ratings, as well as a count of how many ratings a tool had received. For installation numbers, the stores provided different information: the Chrome Web store provided an exact number of users, and the Google Play store provided a “minimum number of installs” (e.g. ‘100,000+’). The Apple App store provided no direct information about user numbers. For processing that ranked tools by installation numbers, we used number of ratings as a proxy for relative numbers of users for tools from the Apple

³ See Lyngs et al. (2019) and their supplementary materials on osf.io/zyj4h/ for further details.

App store.

2.3. Sampling and thematic analysis of user reviews

To get a broad coverage of reviews for tools across stores and types of design patterns, we wrote an R script to systematically sample user reviews for qualitative analysis from the scraped data. To ensure that the reviews analysed came from tools with a minimum level of engagement on the stores, we only sampled reviews from tools on the Google Play and Chrome Web stores that had at least 100 installations/users, and tools from the Apple App Store that had at least 5 ratings. To make the sampled reviews more likely to be informative, we excluded reviews shorter than 10 words. Our script iterated through the four types of design patterns described above, on each iteration picking the tool on each store that included that specific type of design pattern and which had the highest number of users (Google Play and Chrome Web) or ratings (Apple App store). The script then randomly sampled 30 user reviews from each of those tools (if a tool had fewer than 30 reviews, all of its reviews would be included), and started over. The script ran until at least 1,500 reviews were sampled in total. A specific tool could only have its reviews sampled once.

Next, we conducted inductive thematic analysis of the sampled reviews, following the “reflexive” approach described in Braun and Clarke (2006) and Braun et al. (2018). Our work was informed by an earlier analysis of 961 reviews randomly sampled from the same data set by the lead author of this paper (Lyngs, 2021b, ch. 4). However, our analysis in the present paper was conducted from scratch, with a new sample of user reviews drawn in the manner described. After sampling reviews for analysis, the lead author (UL) read through them all and conducted initial coding of recurrent patterns relevant to the research focus. A second author (LC) read through one third of the reviews and independently coded recurrent patterns of meaning. Afterwards, these two authors read through the coded excerpts, recoded them, iteratively sorted the codes into potential themes, and discussed emerging themes.

Thematic coding was conducted using NVivo v1.6.1; all quantitative analyses were conducted in R v4.0.5 (see osf.io/cdqub/ for information on R package versions used).

3. Results

At the time of scraping, 33 tools from Lyngs et al.’s review had become unavailable. Hence, we scraped information for 334 tools: 212 extensions from Chrome Web, 71 apps from Google Play, and 51 apps from the Apple App store.

260 of these tools had received user reviews (160 from Chrome Web, 77 from Google Play, and 23 from the Apple App store). From these, we collected a total of 53,978 distinct reviews (9,070 from Chrome Web, 44,360 from Google Play, and 548 from the Apple App store).

3.1. User reviews

For the tools that had received reviews, the median number of reviews was 6 (min = 1, max = 4,599, IQR = 2 to 48), and the median number of words was 14 (min = 0, max = 645, IQR = 5 to 30; reviews containing only emojis have a word count of 0). Fig. 1 shows the overall frequency of terms used in the reviews, excluding the terms “app/s”, “phone” and “extension” as well as common stop words (‘the’, ‘of’, ‘to’, etc.; excluded using the R package `tidytext` (Silge and Robinson 2016)). Aside from ‘app’ (46% of reviews), the most frequent term was ‘time’, which was used in 18% of all reviews. The terms most commonly following or preceding ‘time’ included ‘limit’, ‘usage’, ‘spent’, ‘screen’, ‘waste’, and ‘management’, suggesting that managing time spent on digital devices was a central topic (which could reflect users’ interests or preferences and/or that currently available tools simply focus on this).

3.1.1. Thematic analysis. We conducted thematic analysis of 1,529 reviews from a total of 61 different tools (Table 1 shows how the reviews

their purpose, and how they might be improved.

As we iteratively worked with the codes and excerpts, we developed three broad themes capturing digital self-control tools' (DSCTs) use contexts and design challenges, assembled from recurrent patterns in the early codes: (i) DSCTs as essential for focusing on less instantly rewarding tasks when digital distractions are easily available, (ii) DSCTs should flexibly adapt to personal definitions of 'distraction', (iii) DSCTs should provide a level of support that is sufficient to change behaviour without feeling too coercive. The NVivo files with our early codes, and their later re-organisation by relation to the three main themes, as well as exports of the code hierarchy to other formats, are available on osf.io/cdquq/.

Theme 1: DSCTs as essential for focusing on less instantly rewarding tasks when digital distractions are easily available A fair number of reviews described reduced time on devices overall or on specific services as the reason why DSCTs were helpful (n = 51, 3.3% of reviews). More commonly than limiting time *per se*, however, DSCTs were described as useful because they helped people focus on a specific task amidst easily available distractions on their devices (n = 212, 14%).

Many reviews described their challenges in this respect as caused by time-wasting patterns of habitual or impulsive use, exacerbated by distracting or 'addictive' design patterns and other adversarial dynamics online. However, DSCTs provided essential countermeasures that reviewers had searched for ("Was looking for a solution for this for over 1 year", R771, Inbox When Ready for Gmail ●▲ C; "Should be on all phones as standard", R982, Mute Screen Time Tracker ▲▲ A), and many reviewers noted that they had experimented with different tools (n = 80, 5%). DSCTs seemed especially helpful when avoiding use of a distracting service was not an option ("This helps me avoid the mind numbing news feed that is of Facebook when work has to be done on Facebook itself", R873, Kill News Feed ● C).

The need for support from DSCTs to stay focused seemed especially acute when tasks were important, but less instantly rewarding than the alternatives. Thus, many reviewers generically stated that DSCTs benefited them in productivity/work contexts (n = 167, 11%; "Not being distracted by unread/misc emails has proven to be of significant benefit to my productivity", R776, Inbox When Ready for Gmail ●▲ C), including studying, writing, or reading (n = 90, 6%). Some reviews expressed this as DSCTs helping overcome 'procrastination' ("seeing [my short-term to-do list] nudges me to get back to work when I am opening a new website to procrastinate", R324, Dayboard ●▲ C).

Another important use context was *sleep*, where some reviewers said their digital devices interfered with going to bed at the intended time. However, DSCTs helped them adjust their habits in favour of the sleep patterns they wanted ("I always used to stay up really late on my phone until I got this app - I have consistently been getting 9 hours sleep and waking up early for the first time in my life!!", R1252, SleepTown ▲★ P; "It really forces me to stay off to the phone, and so it's useful especially when I need to go to bed", R198, BlackOut ● P).

In these situations, DSCTs helped reviewers stay in control by reminding them of what they wanted to do, by reducing exposure to 'temptations' ("Without it, it is too easy for me to be tempted to view and open YouTube sidebar links", R376, DF Tube ● C), or by providing incentives for intended use ("Not only does it help me stay focused, I smile with delight looking at all my trees and flowers at the end of the day", R642, Forest ▲★ C).

Sub-theme: DSCTs as especially useful for people with more acute self-regulation challenges In an extension of the main theme, DSCTs seemed especially useful for individuals who experience self-regulation challenges more intensely (n = 53, 3.5%). This was commonly expressed as a tool being helpful for tackling 'addiction', and/or as reviewers labelling themselves 'addicts', 'procrastinators' or just low in self-control ("This was the perfect solution to my phone addiction. It was consuming my life and my free time", R612, Forest ▲★ P). Occasionally, reviews used clinical labels like ADD or ADHD to explain why tools had been helpful ("As someone with ADHD, the pomodoro

technique has been super effective", R615, Forest ▲★ P).

Theme 2: DSCTs should flexibly adapt to personal definitions of 'distraction'

DSCTs needed to accurately capture users' personal definitions of 'distracting use' to be helpful (n = 243, 16%), a theme often expressed in the form of feature suggestions that would make tools more customisable.

One aspect involved getting tools to accurately capture *what* was considered a distraction. Thus, **self-tracking** tools often received feature requests for the ability to exempt specific functionality from tracking because not all use represented use that reviewers wanted support in managing ("I use a meditation app and can be meditating several times a day, sometimes up to an hour at a time, so as it doesn't distinguish between that and social media it gives a false impression of 'usage'", R989, Mute Screen Time Tracker ▲▲ A).

Similarly, for **block/removal** tools, reviewers wanted to be able to customise which apps or websites were targeted, or target blocking to specific functionality ("I would like to block reddit but still be able to visit productivity related subreddits", R1401, Timewarp ●▲ C). DSCTs on iOS in particular struggled to provide blocking catered to users' needs, because Apple's restrictive developer permissions do not easily allow apps to block other apps. Therefore, iOS tools that provide blocking tended to be limited to blocking websites within the Safari browser. Sometimes, users did not realise this limitation at first, and went on to express their frustration in reviews ("I thought the app would be linked to the apps on my phone I intended to block. So disappointed", R878, Liberate ● A). A workaround provided by one 'blocking' app, was to generate random passwords that would be shown only once (Self-Control to Focus - Lite ●). This app instructed users to go to apps they wanted to block, change their account passwords to the randomly generated one without memorising it, then log out out of their accounts (!). The app would show users the password again on a set date, until which they would be unable to log back into their accounts.

Another aspect involved *when* specific activities should count as distractions. For example, tools for curbing nightly use received feature requests to handle variation in bedtime schedules ("it needs to allow for the real life fact that people can't go to bed and wake up at the same time everyday. If it was done on your amount of sleep you get in hours that would be better", R1250, SleepTown ▲★ A).

Sub-theme: Alignment between user and tool as important for navigating negative emotions Aside from practical reasons that DSCTs should accurately capture behaviour that reviewers wished to change, an important factor seemed to be the negative emotions that interventions could elicit (n = 30, 2%). Thus, whereas the various restrictions and reminders provided by DSCTs were seen as helpful and important, users also needed to carefully navigate the emotional reactions they could create ("I hate it and love it at the same time. 😞", R1299, Stay Focused ●▲ P, "Pisses me off sometimes, but worth it", R357, Detox Procrastination Blocker, ●▲ P).

In order for experiences of "shock", "pressure", etc., to be useful, it was important that the user felt that a DSCT's purpose and framing of usage were closely aligned with their personal perspective. Thus, tools that provided standards for 'good' or 'bad' usage, for example in their presentation of usage statistics or reminder pop-ups, needed to tread carefully. Being reminded that one did not behave in line with one's intentions could be very helpful ("This tells it like I need to be told sometimes... LOVE IT! Simple and effective :)", R690, Go Fucking Work ●★ C; "I like the guilt trip it gives me if I attempt to get into facebook or other social networks", R311, Dayboard, ●▲ C). However, normative messaging could backfire and be profoundly annoying if reviewers did not agree with it ("I was a bit annoyed when Big Brother started getting judgemental on me. Well, first I didnt set a target, and second, I dont think that 54 minutes across a day should make me feel guilty. Especially since I was reading a book during that time" R911, Moment ▲▲ A).

Theme 3: DSCTs should provide a level of support that is sufficient to change behaviour without feeling too coercive

Reviewers searched for DSCTs that provided support which was 'just right' to bring about an intended change in behaviour without feeling too coercive ($n = 271$, 18%). The reviews implied a spectrum of individual differences in where this 'Goldilocks level' was. A few reviewers stated that simply knowing that a tracking tool was monitoring them was sufficient ("I know it's tracking me so I pick up the phone a lot less", R992, Mute Screen Time Tracker ■▲ A). Commonly, however, reviewers wanted DSCTs to provide more strict or persuasive ways to hold themselves accountable, so that they did not succumb to temptations or unwanted habits in moments of weakness or absentmindedness ($n = 132$, 9%). The specifics differed for different design patterns.

For patterns related to distraction **blocking/removal**, it was important that blocking was not too easy to disable in moments of temptation ("This is fantastic but the thing is I can simply remove the extension from chrome which defeats the purpose. Does anyone know how I can stop myself from removing the extension?", R1217, Simple Blocker ● C). On the other hand, if blocking was too difficult to disable, it would be criticised for failing to account for emergencies or otherwise unforeseen circumstances. One developer followed reviewers' suggestion to add a 'pause' button, only to then receive critique that blocking was now too easy to turn off. Reviewers suggested that adding friction to turn off blocking would provide a solution, such as being required to complete a time-consuming or effortful task ("make it necessary to touch the screen for a minute or 2 so that it becomes harder for anyone to stop it", R8, (OffTime) ●▲ P), or type in a password that could be set by a trusted other.

Another aspect was that the amount of effort required to initiate use of blocking tools should be low, because users' initial motivation might be lacking ("Most users have procrastination issues which make them postpone the moment in which they will click on the "Start focusing" button", R545, Focus ●■▲ C). Thus, many feature requests asked for the ability to set a schedule where blocking would turn on automatically ("can I request an automatic start time option? So every M-F, 9-5 the app starts working automatically? I need it to cut me off from whatever dumb thing I'm doing before work starts :)", R238, Block & Focus ●▲ C). An interesting suggestion for tools alternating 'focus' and 'break' sessions, was to flip the order to begin with a break, followed by enforced blocking. This might lower the amount of motivation needed to initiate a focus session ("Hitting the start button the very first time is hard! Starting in a short break would allow me to hit this without a second thought, because I have a few short moments to surf before the focus is forced.", R232, Block & Focus ●▲ C).

For **reward/punishment** design patterns, reviewers sometimes suggested ways to increase the incentives provided to make tools more effective ("If I could also name the trees (so I'll feel even worse about letting them die) it would be even better!", R650, Forest ▲★ C). Here too, however, tools needed to find the right balance. If a punishment was perceived as too harsh, it might make users feel dejected ("I hate the fact that we get a destroyed building (...) I feel like it is too punishing and almost says 'you've failed' ", R1247, SleepTown ▲★ P).

For **goal advancement** patterns, the 'just right' level of support mainly related to the frequency with which users would be reminded of their goals: too infrequent, and users might fail to complete their tasks; too frequent, and users would get annoyed. For example, one browser extension replaced the content of new tabs with to-do lists, and also provided task reminders via pop-ups. This tool was on the one hand criticised for dropping unfinished tasks from the day before on the to-do list, because users would then forget those tasks. On the other hand, its pop-up reminders were criticised for being too frequent ("the todo popup on each page load is quite useful (...) it should be limited (across multiple tabs) to only pop-up every minute or so (...) I find myself snoozing it too often just because the popup bothers me", R339, Dayboard, ●■ C).

For **self-tracking** patterns, reviewers commonly said that DSCTs implementing this helped them understand and assess their usage, which was often described as the first step in changing behaviour. However, many reviews said that it was not in itself sufficient support for change, and that tracking needed to be combined with other

strategies to be effective ("It does not stop me spending 10 hours on Facebook but it does inform me (...) What has worked great is using rescue time with very intentional time blocking", R1157, RescueTime, ▲■●). We elaborate on this in the sub-theme below.

Similarly to the challenges around blocking, self-tracking tools on iOS faced an additional challenge related to the amount of user effort required to track use in the first place. This occurred because Apple's developer permissions do not allow apps to easily track time spent in other apps. As a work-around, some tools required users to take screenshots of their phone's battery page, from which information about time spent in different apps could be extracted. This work-around was, however, seen as too effort-demanding by reviewers ("I should have looked at how you use this. Take a picture everyday???? Really?", R121, App Usage Tracker ▲ A).

Sub-theme: Combining design patterns to find the Goldilocks level of support Rather than simply focusing on the parameters of individual design patterns, some reviews focused on how a combination of different types of patterns could make DSCTs provide the right level of support for behaviour change. For example, some reviews for self-tracking tools suggested that adding a goal-setting or blocking element would make it more effective. Similarly, some reviewers wanted block/removal tools to not simply block distractions, but also actively nudge them in ways that supported intended habits. This might happen by, for example, redirecting blocked websites to productive alternatives, or replacing recommender feeds with content people wanted to engage with more ("if, where the home-screen recommended section is, this could be replaced by a website of our choice (...) This way I could do music theory, or learn German. Things I always intend to do, but don't do enough. But if these things came up every time I log into Youtube I'd definitely use them.", R372, DF Tube ● C). Similarly, some reviews suggested that a particular tool's combination of design patterns was what made it more effective than alternative tools ("This is the first todo list that actually keeps me on task. Normally I go down the rabbit burrow when working and either get way too focused on one task or completely go off track. This tool keeps pulling my attention back to the task list where I can make sure I'm not off track.", R550, Focus ■●▲⁴).

3.2. User numbers and ratings

The distribution of numbers of users/installations and ratings are shown in Fig. 2. For Google Play tools, the median 'minimum number of installs' was 10,000 (min = 5, max = 5,000,000, IQR = 750 to 100,000). For Chrome Web tools, the median number of users was 194.5 (min = 1, max = 1,736,018, IQR = 22 to 2,601). The Apple App store did not display information about number of users (the median number of ratings for DSCTs from the Apple App store was 0; max = 14,900, IQR = 0 to 2.5). The median average rating across all tools was 4.3 (min = 1, max = 5, IQR = 4.0 to 5.0; ratings given on a scale from 1 to 5 stars on all stores).

Fig. 3 shows the top tools, ranked by number of users (Chrome Web and Play) or number of ratings (Apple App store), alongside the design patterns they include. Across all tools, 64% included a single type of design pattern for digital self-control, whereas 32% combined two. Very few combined three (3%, $n = 11$) or all four (1%, $n = 2$). Tools that combined two or more types had slightly higher average ratings than tools which implemented a single type (median average rating with a single type = 4.2, median for tools combining two or more = 4.39, $p = 0.003$ in Wilcoxon's rank-sum test; excluding tools with fewer than 15 ratings; see Fig. 4).

On the Chrome Web and Play stores, Block/removal was the most

⁴ Focus replaces new tab content with a to-do list and activity log, alongside a 'start focusing' button. When the button is clicked, a 25 min work session is started where websites on a blacklist are blocked. If the user tries to access a blocked website, a countdown timer is shown. At the time of writing, this extension is no longer available on the Chrome Web Store.

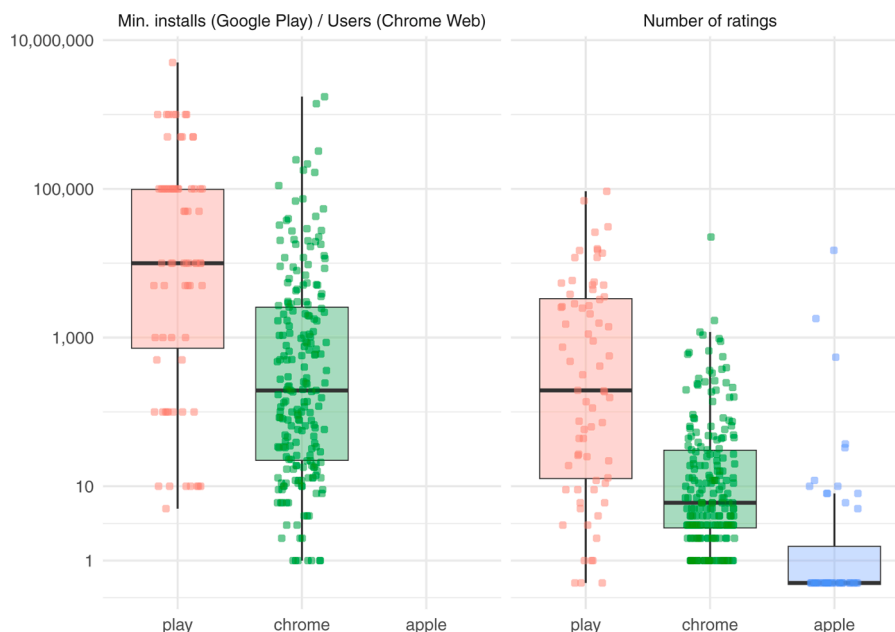


Fig. 2. Distributions of installation numbers and number of ratings for DSCTs on the Google Play ($n=71$), Chrome Web ($n=212$), and Apple App ($n=51$) stores. Values are shown on a log scale — tools with zero ratings are scored as having 0.5 ratings, so that they show in the plot. The Apple App store provided no direct information about installation numbers.

common type of design pattern (included in 87% and 65% of all tools, respectively, compared to 37% on the Apple App store), whereas Reward/punishment was the least common (included in 6% and 15% of tools, respectively, compared to 43% on the Apple App store). Notably, the app ‘Forest’ (which combines reward/punishment and self-tracking in the form of virtual trees that grow while a timer counts down) ranked highest on the Play store and 5th on the Chrome Web store, while a very similar app, ‘Flora’, ranked highest on the Apple App store in terms of number of ratings received.

4. Discussion

To sum up, reviewers described DSCTs as effective means of taking back control and making device use fit better with their lives. A fair number of reviews suggested that limiting time spent on devices, or on specific apps or services, was an important purpose for using DSCTs. More commonly than limiting time *per se*, however (observed four times as frequently in our thematic analysis), DSCTs were specifically used for focusing on less instantly rewarding tasks amidst easily available digital distractions. Specific contexts of use included productivity (including work, studying, reading, and writing), as well as bedtime habits. DSCTs seemed to be especially useful for individuals with higher baseline levels of self-regulation difficulties, as suggested by reviews who framed usefulness in terms of coping with ‘addiction’ or conditions like ADHD.

Reviewers wanted DSCTs to provide interventions that accurately targeted specific patterns of use. This was often expressed in the form of feature requests for, for example, being able to target blocking on specific functionality on a website, or exempt specific apps from usage tracking. A contributing factor in this respect was that DSCTs’ interventions could cause negative emotional reactions in the moment of temptation. If user and tool were not in agreement on their definitions of ‘distracting use’, interventions might simply be considered too annoying.

Reviewers also wanted DSCTs to provide a level of support that was sufficient for effecting their intended change in behaviour, without feeling too coercive. Reviewers differed in the level of support they wanted, but commonly looked for interventions strict or rewarding enough to stay on track when they felt tempted to indulge in digital

distractions or drift into unwanted habitual use. DSCTs needed to strike a delicate balance in this respect, because overly strong incentives might backfire if they made users feel dejected when they failed, or if they did not allow for emergency use or shifting daily schedules.

Some reviewers suggested that combining different types of design patterns was an effective way to provide a level of support that was ‘just right’. This matched the observation that DSCTs which combined two or more types of design patterns received slightly higher overall ratings than tools including only one. A final consideration related to how much effort a tool required to use: whereas high effort to *stop* a tool could help hold users accountable to their goals, the level of effort required to *start* use should be low, to handle potential lack of motivation to initiate blocking on part of the user.

In the following, we discuss implications for design and opportunities for future research.

4.1. Goldilocks principles for DSCTs

Across different design dimensions, users’ desired level of support from DSCTs reflected a Goldilocks principle: ‘too little’ and a DSCT would fail to bring about an intended change in behaviour; ‘too much’ and it could make the user feel frustrated, annoyed, or dejected. The considerations for a ‘just right’ level of support varied for different design dimensions: For example, for distraction blocking/removal the main concern was *how much effort* the user needed to apply to circumvent the blocking. For goal-advancement, it was *how frequently* users were reminded of their goals. In designing for a Goldilocks level of support, developers therefore need to consider more than a single dimension of engineered friction (Cox et al., 2016; Gould et al., 2021). How might we conceptualise Goldilocks principles for different design dimensions and assess the level of support that a DSCT provides?

4.1.1. Using a dual-systems model to frame the Goldilocks level of support.

The dual-systems model commonly used in psychology and HCI provides insight into how DSCTs support or disrupt target behaviors through multiple cognitive pathways. Lyngs et al. (2019) originally used this model (Fig. 5) to analyse how the four main types of design patterns in DSCTs support self-regulation. We suggest that it is also useful for understanding Goldilocks principles for those design patterns.

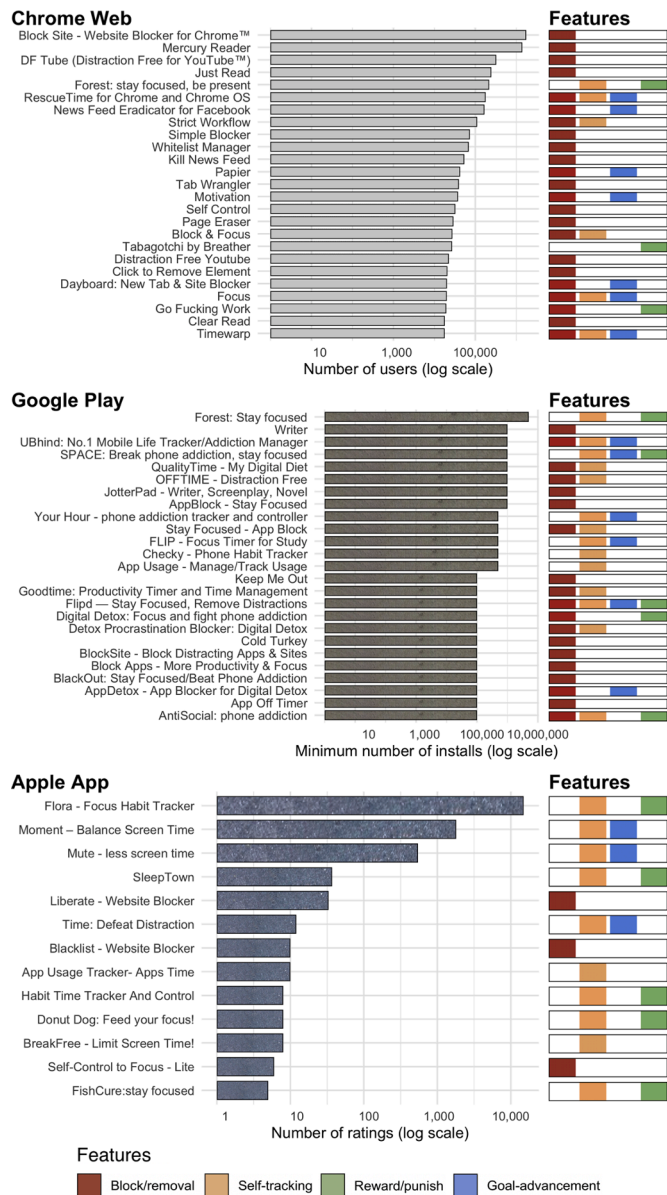


Fig. 3. Top tools in terms of installation numbers (Chrome Web and Play) or number of ratings (Apple App store). Only 13 of the Apple App store tools had received ratings. Differences between the stores in terms of design patterns are readily apparent: Block/removal patterns are particularly common on the Chrome Web store (87% of all tools), where browser extensions can easily modify distracting elements on web pages. Reward/punishment patterns are included in 43% of tools on the Apple App store, compared to, respectively, 6 and 15% on the Chrome Web and Play stores.

According to the dual-systems model, behaviour can be activated via quick, automatic reactions to stimuli in our external and internal environment (“System 1”, e.g. taking out your phone as a habitual response to boredom). It can also be activated via slower, controlled reactions to our conscious goals and standards (“System 2”, e.g. deliberately taking out your phone to text a friend). Potential behaviours compete to get expressed, and what wins out is the action that receives the greatest total activation. The purpose of DSCTs is to influence the competition between System 1 and System 2 responses such that the winning behaviour is more likely to be aligned with our longer-term goals (e.g., focusing on a work task, or reducing time spent on a device or service). Different design patterns achieve this in different ways. Patterns for **blocking/removing** distractions aim to stop unwanted actions from being triggered or expressed via System 1 control, and to decrease the value of

those actions when considered by System 2; **self-tracking** aims to enable conscious System 2 monitoring of how device use compares to the user’s goals and standards; **goal-advancement** aims to bring specific usage goals into consciousness awareness (e.g. via task reminders), thereby enabling System 2 to initiate certain actions or inhibit others; and **reward/punishment** aims to make it more likely that System 2 control succeeds, by increasing or decreasing the expected value of target actions (Lyngs et al., 2019; Shenhav et al., 2017).

What insights can we draw about Goldilocks principles for design patterns in DSCTs from our available data and the dual systems model? In the following, we illustrate this in relation to distraction blocking/removal, and how design patterns may be combined.

4.1.2. Blocking/removal of distractions: effortful to override, effortless to initiate. From the dual systems perspective, ‘too little’ support from a blocking tool means that the effort required to *override blocking* is too low to stop unwanted System 1 habits from being expressed, and to impose costs on unwanted actions that are large enough for System 2 to favour better options (recall the reviewer who was unable to stop themselves from deactivating blocking). Conversely, ‘too much’ effort occurs if blocking interferes with important, but unexpected usage goals.

In providing the ‘just right’ level, the data suggested the following guidelines (which converged with other studies, (Kim et al., 2018; 2019a; 2019b; Tseng et al., 2019b; Zhang et al., 2022)):

1. By default, require the user to invest moderate effort to override blocking (e.g., having to manually type in a long string of random letters), in order to enforce commitment while still allowing emergency use
2. Make the level of required effort adjustable by the user, to account for individual differences
3. Provide ways to initiate blocking with little effort, to counter instances of low motivation (e.g., via automatic blocking according to a schedule)

Thus, as the reviews suggested, a viable solution is to require users to complete an effortful task to override blocking (Kim et al., 2019b; Zhang et al., 2022). It is an open question how much the Goldilocks amount of effort varies between individuals and with emotional state (Kim et al., 2019b). However, a simple design solution found in commercially available tools is to allow users to choose the level of enforcement they find appropriate. For example, Cold Turkey Blocker (<https://getcoldturkey.com>) provides a choice between no enforcement tool (can be turned off at will), having to type in a password or string of random letters, having to restart ones computer, or allowing blocking to be modified only at certain times.

Initiating blocking comes at the short-term cost of excluding ourselves from potentially rewarding actions. This puts System 2, which compares the expected value of different actions (Shenhav et al., 2017), in a conflicted position. At times, manually initiating blocking is therefore ‘too high’ effort, which reviewers expressed as procrastinating on starting a block session. As a solution, blocking could be automatically applied (Kim et al., 2018; Tseng et al., 2019b), or the perceived short-term cost of blocking lowered by e.g. adding a time lag between initiation and application of blocking. Though not directly expressed in our data, initiating distraction blocking with ‘too little’ effort may conversely increase the risk that users apply blocking at the wrong time, or to the wrong content. This risk might be acceptable, however, if blocking is not impossible, but only effortful to override.

Our thematic analysis yielded less specific information about the ‘just right’ level of support for the other three types of design patterns. Overall, however, our data and the dual systems model suggest that **goal-advancement** patterns should find the ‘just right’ frequency with which to bring goals into conscious System 2 awareness: too infrequent, and goals are held in mind too briefly for users to complete their tasks;

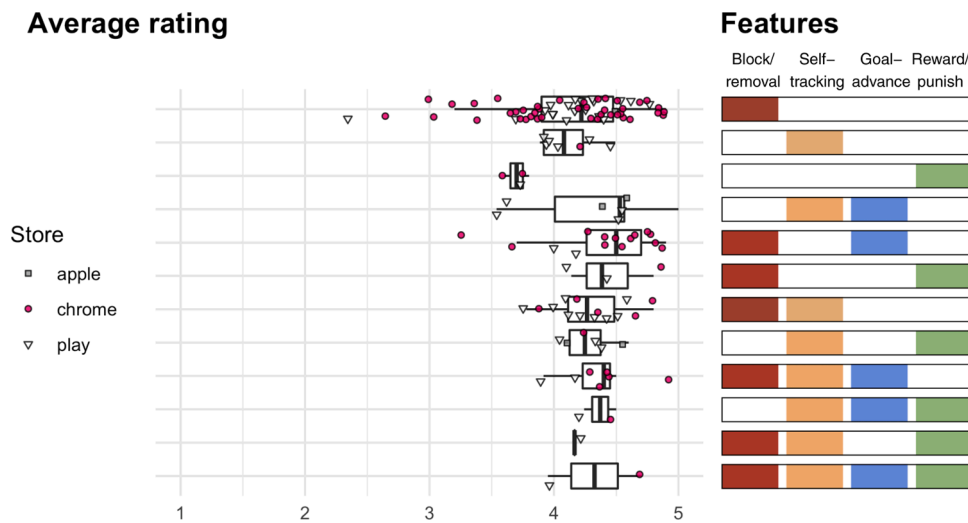


Fig. 4. Average ratings by design patterns in digital self-control tools (for tools with 15 or more ratings). Tools including more than one type of design pattern had higher average ratings (median = 4.39) than tools implementing a single type (median = 4.2, $p = 0.003$, Wilcoxon’s rank-sum test).

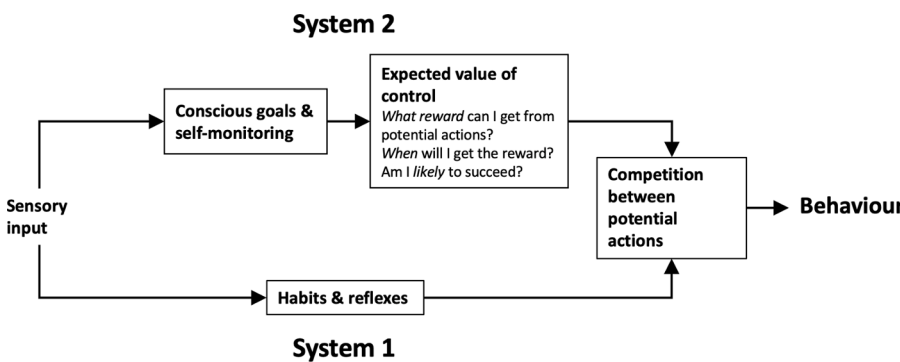


Fig. 5. A simplified dual-systems model of self-regulation, based on Lyngs et al. (2019), Shenhav et al. (2017), Shea et al. (2014), Norman and Shallice (1986). Behaviour can be activated via quick, automatic reactions to stimuli (System 1), as well as via slower, controlled reactions to conscious goals and standards (System 2). The strength of System 2 control is mediated by the expected value of control. DSCTs aim to influence the competition between System 1 and System 2 responses such that the winning behaviour is more likely to be aligned with the user’s longer-term goals (e.g., focusing on a work task, or simply reducing time spent on a device or service).

too frequent and users get annoyed. **Reward/punishment** should provide a ‘just right’ incentive that raises the expected value of target use over that of alternative options: too little, and users will still succumb to unwanted habits; too much, and the incentives can undermine users’ sense of self-efficacy when they fail. **Self-tracking** should present usage information in a form that facilitates goal-setting and activates related behaviour. Too vague a connection to goals, and information merely raises awareness without prompting behaviour; too specific, and users might feel judged by the tool and rebel against it.

4.1.3. *Combining design patterns to provide the Goldilocks level of support.* In this section, we discuss an interesting theme in our data, namely how combining design patterns can provide a Goldilocks level of support. This was often suggested in the reviews, where self-tracking in particular was often described as insufficient support in itself to change behaviour (see also Collins et al., 2014). Indeed, DSCTs that combined two or more design patterns tended to receive higher ratings compared to those that focused on only one (mirroring recent psychological research on self-control strategies in day-to-day life, showing that using multiple strategies is especially effective (Milyavskaya et al., 2021)).

In the dual systems lens, a DSCT’s interventions need to in aggregate provide sufficient influence on the cognitive system to tip peoples digital device use down a path that is aligned with their longer-term goals. This may be achieved via a single design pattern, or by combining multiple patterns that intervene on different psychological mechanisms. The DSCTs and user reviews provide plenty of examples: Browser extensions can place to-do lists on new tabs to remind the user of their tasks, and also provide pomodoro-style blocking of distracting websites. The

immensely popular app *Forest* (Seekrtech, 2018) combines a countdown-timer — which reminds people that they are currently wanting to not use their phone — with an animated growing tree that dies if they fail to adhere to the non-use goal. Reviewers suggested that distracting recommender feeds could be replaced with content they would like their better selves to spend more time engaging with.

Each design pattern’s ‘Goldilocks’ level of support can interact with the other patterns included in a tool. We propose that by combining patterns, each pattern may need to provide only a lower level support to be effective, which can make a tool less likely to be experienced as too coercive. To illustrate, consider a woman who finds it challenging to do what she intends on Facebook (e.g., post an update in the local Taekwondo group, and message a family member) without being side-tracked by irrelevant information on the site. A DSCT might support her via a pop-up that every few minutes reminds her of her intention (goal-advancement). She experiments with the frequency of reminders until she finds the ‘Goldilocks level’: a balance point where her conscious intentions are re-activated sufficiently often for her to stay on task amidst distractions that trigger irrelevant System 1 habits and/or crowds out her intentions from System 2 awareness. Imagine the DSCT now also hides her Facebook newsfeed until she actively chooses to see it (block/removal). The level of potentially distracting on-site information — and the related triggering of irrelevant behaviours — is now lower. Therefore, her Goldilocks frequency for the reminders is also likely to change: she is now able to stay on task (and prefers) less frequent reminders.

We note that the higher ratings we observed for DSCTs that combine multiple design patterns have alternative explanations than such tools

being more likely to provide a Goldilocks level of support: For example, DSCTs that combine multiple design patterns may simply have received more developer attention, and therefore be better designed on average. We encourage future research to explicitly investigate how different design patterns interact in DSCTs, for example via studies with factorial designs similar to those called for in digital behaviour change research (Klasnja et al., 2017).

4.2. Designing to reduce psychological reactance

Another important challenge arising from the user reviews is to design tools that help people navigate negative emotional reactions that DSCTs can cause in the moment of enforcement. The data point to two factors in this respect:

4.2.1. Precisely targeting interventions to users' notion of distracting use.

Future work should focus on how DSCTs' interventions can be precisely tailored to aspects of users' digital environments that cause self-regulation difficulties. The more DSCTs' interventions are off the mark in this respect, the more they may cause reactance.

Most research on DSCTs has focused on innovating 'external mechanisms' that help people manage or monitor their device use, i.e., design patterns like lockout timers or productivity dashboards that apply similarly across many different apps (Lukoff et al., 2021). However, this approach is often too crude, as evidenced by user reviews expressing a desire for control over, for example, what specific website features are blocked by DSCTs, or what apps count toward 'screen time'. Indeed, numerous Chrome Web Store tools included in the present paper instead serve to redesign specific design features experienced as problematic (e.g. *Inbox When Ready for Gmail* hides the inbox on gmail.com until a button is pressed; *Click to Remove Element* lets users remove website elements by clicking on them).

In research so far, less effort has been invested in developing and evaluating such tools that empower users to change 'internal mechanisms' within apps and services. However, recent studies of the effects or directly modifying design elements of Facebook or the mobile Twitter client (e.g., by hiding specific UI elements) provides strong support for the potential of this approach (Lyngs et al., 2020; Zhang et al., 2022). Moreover, new research on the *GreaseVision* platform aims to provide researchers and end users with accessible tools for modifying UI elements in mobile apps that they find distracting or harmful (Datta et al., 2022a; 2022b; Kollnig et al., 2021). This holds great promise for future research on how DSCTs can support self-regulation in a precisely targeted fashion to minimise negative emotional reactions to their interventions.

4.2.2. Developing interventions that support internalisation of motivation for DSCT use.

Another research and design opportunity concerns the ways in which interventions are framed. Moving beyond the dual-systems model, self-determination theory (SDT), an evidence-based theory of human motivation, offers a promising explanation for why DSCTs can trigger frustration and even hostility (Ryan and Deci (2017); recall that reviewers stated that they simultaneously 'loved and hated' their distraction blockers — even though DSCTs were valuable, they were also 'pissed off' by them). According to SDT, people have a basic need for autonomy. When a DSCT applies enforcement mechanisms that hold users accountable to their goals, it can be perceived by users as a form of external control that threatens autonomy (Brehm and Brehm, 2013). This can cause 'reactance', an unpleasant motivational arousal in the user, which may result in behavioural backlash where the users tries to circumvent or intentionally contradict the tool (Fitzsimons and Lehmann, 2004).

However, DSCTs interventions need not be perceived in this way. In SDT, 'extrinsic motivation' refers to activity done to obtain some outcome other than the enjoyment of that activity itself. Usage of DSCTs is typically extrinsically motivated, because people use them to obtain

some change in their use of digital devices. Extrinsic motivation can be experienced as more or less autonomous (see Fig. 6): at one extreme, use of a DSCTs can be experienced as fully controlled by external rewards or punishments (e.g., "I have to use the DSCT"), and therefore be likely to lead to reactance. At the other, their use might be experienced as fully autonomous and directed by the users personal goals (e.g. "I use the DSCT because it helps me achieve my goals and values"), which should cause less reactance.

DSCTs that provide distraction blocking currently tend to simply let users set target apps or websites and then provide generic "You've reached your limit" lock screens at the time of enforcement. Whereas some of the DSCTs we analysed provided creative variations (e.g., *Go Fucking Work* shouts obscenities at its users, and has a dedicated niche following; *RescueTime* shows an animated dog with a wagging tail sending a shameful look to the user), there is much room for researchers and designers to explore ways to create interventions that support the experience of DSCT use as being directed by the users personal goals and values. For example, future research could explore how to support autonomy at the time when the user delegates their goals to the DSCT. Here, DSCTs might borrow from motivational interviewing techniques to elicit and affirm the reasons behind the user's goals (Miller and Rose, 2009). How to support autonomy at the time when the DSCT enforces the user goal should also be explored. Here, DSCTs might, e.g., remind the user of the why behind their use of the tool, using their own words (see Xu et al., 2022 for a recent example), or increase the salience of rewards they value by displaying pictures of loved ones they could spend time with after finishing their tasks (Lukoff et al. 2022).

4.3. Navigating platform challenges for future DSCTs

The experience of reviewers depended heavily on the underlying technical infrastructures, which create different conditions for the emergence of ecosystems of DSCT in the wild. The largest degree of flexibility exists on the web, where browser extensions make it possible for developers — and tech-savvy end-users — to customise the look and functionality of any website via injected stylesheets and scripts. Ecosystems of browser extensions today exist for all major browsers, which all share the cross-browser [Web Extensions API](#).

The situation is very different if a user wishes to experiment with DSCTs on their smartphone or tablet. Unless users jailbreak their devices, no options exist for users to customise the look and function of mobile apps beyond the options that app developers choose to provide. Moreover, the options available for self-tracking or blocking tools differ dramatically between Android and iOS: because of Google's less restrictive approach to developer permissions, a diverse ecosystem exists on Android for apps that provide bespoke usage visualisation or customised app blocking schedules, in addition to what Google provides in their 'Digital Wellbeing' app. On iOS, by contrast, Apple's Screen Time tool was, at the time data for the present paper was scraped, the only option with which users could properly block apps or see their usage⁵ As mentioned in the results section, DSCTs on iOS sometimes resorted to creative workarounds such as 'blocking' apps by asking users to set a new, random password for their social media accounts and then log out, after which they would only see the password again on a set date, or tracking usage by getting users to take screenshots of the battery page in the smartphone settings. Unsurprisingly, such workarounds tended to be seeded by users as too effort-demanding to be workable.

We do not take a strong stance on whether Apple or Google are 'right' in terms of the level of permissions they make available to developers, as greater operating system access also entails greater security risks. However, we do note that such risks are already present in browser

⁵ A Screen Time API, introduced in the summer of '21 after long campaigning by developers (Digital Wellness Warriors, 2018), now allows creation of blocking tools for parental control purposes, but does not enable apps for tracking and visualising screen time.

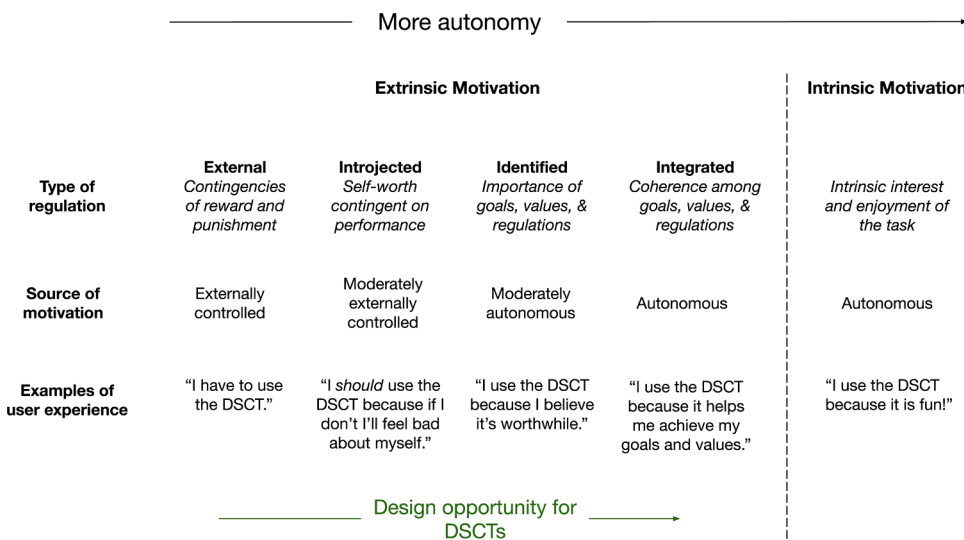


Fig. 6. Self-determination theory distinguishes between different types of motivation, including extrinsic (doing an activity to obtain some separable outcome) and intrinsic motivation (doing an activity because it is inherently interesting or enjoyable). DSCT use tends to be extrinsically motivated, because it serves the separable outcome of changing one’s patterns of device use. DSCTs’ interventions can be experienced as externally controlled (“I have to use the DSCT”) and trigger reactance. Future research should explore how DSCTs might be designed to consistently be experienced as supporting the user’s autonomy (“I use the DSCT because it helps me achieve my goals and values”). Illustration adapted from Lukoff et al. (2022), who adapted it from Peters et al. (2018) and Ryan and Deci (2000).

extensions, which both Apple and Google, alongside other companies, already allow in their web browsers. In relation to the web, such extensions already give users the power to locally block access to specific websites, or provide detailed modifications to their appearance and functionality.

Here, we mainly wish to highlight how Google and Apple’s approaches create distinct roads ahead in relation to providing the support that users seek to stay in control: Google’s less restrictive approach has meant that a diverse range of potent digital self-control tools are available on Android. Therefore, Google face less pressure to ‘get it right’ with their own Digital Wellbeing tools, because less common user needs can be addressed by third-party tools within an ecosystem of DSCTs on their platform. Apple’s approach, however, implies that their own Screen Time app needs to accommodate a very wide range of user needs. This will be challenging, as there can be a trade-off between customisability and effort of use, where providing sufficiently fine-grained controls to account for all users’ needs might make it too complicated to navigate. Researcher-led initiatives such as *GreaseVision* (see Section 4.2.1), may point the way towards a more inclusive future where end-users have the power to modify mobile user interfaces to suit their individual needs (Datta et al., 2022a).

4.4. Limitations and future work

First, the data used in the present paper were scraped in March 2019. Hence, the findings presented may deviate from what would be learned by re-scraping the app and browser extension stores in 2022. However, all of our data and materials are openly available, so our process should be readily repeatable. As such, we believe the present paper provides a useful snapshot of DSCTs across three online stores, which can serve as point of comparison for future research tracking how the landscape of these tools develop.

Second, the content of user reviews is subject to various forms of bias. For example, reviews may be more likely to represent views from users whose experiences with a particular tool are either highly positive or negative. That is, for a user to be sufficiently motivated to bother reviewing a tool in the first place, their level of excitement about a tool might exceed the excitement of the average user (Panichella et al., 2015). Nevertheless, the fact that our findings from reflexive thematic analysis resonate well with findings from studies using other methods such as surveys or interviews (Ko et al., 2015; Tran et al., 2019) suggests that user reviews of DSCTs can provide valid and useful insights.

Finally, for the present paper, we programmatically sampled reviews for thematic analysis from our full data set of reviews with the goal of sampling broadly across stores and types of design patterns. However,

reviews can be sampled in many different ways to provide data tailored to more narrow research questions, e.g. for researchers specifically investigating design parameters relevant for lockout mechanisms (Kim et al., 2019a). We encourage future research on DSCTs to make their data sets and analysis procedures open for others to re-use, to enable synergies between the efforts of different research groups in the field.

5. Conclusion

The landscape of digital self-control tools available online amounts to hundreds of thousands of natural experiments in supporting self-regulation of device use. As such, it provides a running commentary on user needs and struggles in common digital environments, as well as concrete suggestions for solutions. Therefore, studying this landscape can be a valuable complementary approach to evaluating interventions in controlled studies, as the latter is difficult to scale to broad assessment of strategies and implementations across the design space.

In this paper, we investigated what user reviews reveal about digital self-control tools’ contexts of use and design challenges, and how design features relate to ratings and installation numbers, using data from 334 tools on the Google Play, Chrome Web, and Apple App stores. Compared to existing research, our work followed a more transparent procedure for sampling user reviews for thematic analysis from the full data set, with reproducible, programmatic sampling, and open data and materials.

As most daily tasks come to be mediated by digital devices, it is important that users have the tools they need to effectively manage time and attention in their digital environments. As such, discovering effective interventions to support self-regulated digital device use is a practical and urgent research challenge. In this respect, we encourage future work to make further use of the landscape of digital self-control tools on app and browser extension stores as a resource for investigating effective design solutions.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ulrik Lyngs has developed three extensions for the Safari browser that hide distracting UI elements on Facebook, YouTube, and Twitter. These are paid apps (£0.89) available on the Apple App Store. None of these apps were included in the data analysed in the present paper.

Acknowledgments

Funding: Ulrik Lyngs was funded by the UK Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/J017728/2]; and the Carlsberg Foundation [grant number CF20-0678]. The other authors received no specific grant for this research. We thank Claudine Tinsman for proof-reading a late version of the manuscript.

References

- Bakker, D., Kazantzis, N., Rickwood, D., Rickard, N., 2016. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment. Health* 3 (1), e7. <https://doi.org/10.2196/mental.4984>.
- Biskjaer, M.M., Dalsgaard, P., Halskov, K., 2016. Taking action on distraction. *ACM Interact.* 23 (6), 48–53. <https://doi.org/10.1145/3001958>.
- Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. *Qual. Res. Psychol.* 3 (May 2015), 77–101. <https://doi.org/10.1191/1478088706qp0630a>.
- Braun, V., Clarke, V., Hayfield, N., Terry, G., 2018. Thematic Analysis. In: Liamputtong, P. (Ed.), *Handbook of Research Methods in Health Social Sciences*. Springer Singapore, Singapore, pp. 1–18. https://doi.org/10.1007/978-981-10-2779-6_103-1.
- Brehm, S.S., Brehm, J.W., 2013. Psychological reactance: A Theory of freedom and control. *Academic Press*.
- Cecchinato, M.E., Rooksby, J., Hiniker, A., Munson, S., Lukoff, K., Ciolfi, L., Thieme, A., Harrison, D., 2019. Designing for Digital Wellbeing: A Research & Practice Agenda. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. W17:1–W17:8. <https://doi.org/10.1145/3290607.3298998>.
- Chen, N., Lin, J., Hoi, S.C.H., Xiao, X., Zhang, B., 2014. AR-miner: Mining Informative Reviews for Developers from Mobile App Marketplace. Proceedings of the 36th International Conference on Software Engineering. ACM, New York, NY, USA, pp. 767–778. <https://doi.org/10.1145/2568225.2568263>.
- Collins, E.I.M., Cox, A.L., Bird, J., Cornish-Tresstail, C., 2014. Barriers to engagement with a personal informatics productivity tool. Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures the Future of Design - OzCHI '14 370–379. <https://doi.org/10.1145/2686612.2686668>.
- Cox, A.L., Gould, S.J.J., Cecchinato, M.E., Iacovides, I., Renfree, I., 2016. Design Frictions for Mindful Interactions: The Case for Microboundaries. Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 1389–1397. <https://doi.org/10.1145/2851581.2892410>.
- Daskalova, N., 2018. Personalized behavior-powered systems for guiding self-experiments. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. DC06:1–DC06:4. <https://doi.org/10.1145/3170427.3173017>. <http://doi.acm.org/10.1145/3170427.3173017>.
- Datta, S., Kollnig, K., Shadbolt, N., 2022a. GreaseVision: Rewriting the Rules of the Interface. 2204.03731.
- Datta, S., Kollnig, K., Shadbolt, N., 2022. Mind-proofing your phone: Navigating the digital minefield with greaseterminator. 27th International Conference on Intelligent User Interfaces. Association for Computing Machinery, New York, NY, USA, p. 523536. <https://doi.org/10.1145/3490099.3511152>.
- Davenport, T.H., Beck, J.C., 2001. *The attention economy: understanding the new currency of business*. Harvard Business School Press.
- Digital Wellness Warriors, 2018. *Apple Let developers help iphone users with mental wellbeing*.
- Einstein, M., 2016. *Black ops advertising*. OR Books, New York.
- Fitzsimons, G.J., Lehmann, D.R., 2004. Reactance to recommendations: when unsolicited advice yields contrary responses. *Market. Sci.* 23 (1), 82–94. <https://doi.org/10.1287/mksc.1030.0033>.
- Gould, S.J.J., Chuang, L.L., Iacovides, I., Garaialde, D., Cecchinato, M.E., Cowan, B.R., Cox, A.L., 2021. A Special Interest Group on Designed and Engineered Friction in Interaction. Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, Yokohama Japan, pp. 1–4. <https://doi.org/10.1145/3411763.3450404>.
- Haroz, S., 2019. Updates to the CHI Submission and Reviewing Guides Promote Open Research Practice. <https://transparencystatistics.org/2019/08/01/updates-to-chi-submission-and-reviewing-guides/>.
- Hiniker, A., Hong, S.R., Kohno, T., Kientz, J.A., 2016. Mytime: designing and evaluating an intervention for smartphone non-use. *Proc. 2016 CHI Conf. Hum. Factors Comput. Syst.* 4746–4757. <https://doi.org/10.1145/2858036.2858403>.
- Hiniker, A., Radesky, J.S., Livingstone, S., Blum-Ross, A., 2019. Moving Beyond “The Great Screen Time Debate” in the Design of Technology for Children. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. panel01:1–panel01:6. <https://doi.org/10.1145/3290607.3311745>.
- JDev, 2019. News Feed Eradicator for Facebook. <https://chrome.google.com/webstore/detail/news-feed-eradicator/fjclmjmjhkklehbacihajopkjlhlg?hl=en>.
- Kim, I., Lee, U., Cha, N., 2018. Exploring context-aware proactive blocking for distraction management. Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers - UbiComp 18. ACM Press. <https://doi.org/10.1145/3267305.3274108>.
- Huguet, A., Rao, S., McGrath, P.J., Wozney, L., Wheaton, M., Conrod, J., Rozario, S., 2016. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLoS ONE* 11 (5), e0154248. <https://doi.org/10.1371/journal.pone.0154248>.
- Kim, J., Jung, H., Ko, M., Lee, U., 2019. Goalkeeper: exploring interaction lockout mechanisms for regulating smartphone use. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3 (1), 16:1–16:29. <https://doi.org/10.1145/3314403>.
- Kim, J., Park, J., Lee, H., Ko, M., Lee, U., 2019. LocknType: Lockout Task Intervention for Discouraging Smartphone App Use. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 697:1–697:12. <https://doi.org/10.1145/3290605.3300927>.
- Klasnja, P., Hekler, E.B., Korinek, E.V., Harlow, J., Mishra, S.R., 2017. Toward Usable Evidence: Optimizing Knowledge Accumulation in HCI Research on Health Behavior Change. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 3071–3082. <https://doi.org/10.1145/3025453.3026013>.
- Ko, M., Chung, K.-M., Yang, S., Lee, J., Heizmann, C., Jeong, J., Lee, U., Shin, D., Yatani, K., Song, J., 2015. NUGU: A group-based intervention app for improving self-regulation of limiting smartphone use. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15 1235–1245. <https://doi.org/10.1145/2675133.2675244>.
- Kollnig, K., Datta, S., Van Kleek, M., 2021. *I Want My App That Way: Reclaiming Sovereignty Over Personal Devices*. Association for Computing Machinery, New York, NY, USA.
- Kovacs, G., Wu, Z., Bernstein, M.S., 2018. Rotating online behavior change interventions increases effectiveness but also increases attrition. *Proc. ACM Hum.-Comput. Interact.* 2 (CSCW), 95:1–95:25. <https://doi.org/10.1145/3274364>.
- Lee, J., Walker, E., Burleson, W., Kay, M., Buman, M., Hekler, E.B., 2017. Self-experimentation for behavior change: design and formative evaluation of two approaches. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 6837–6849. <https://doi.org/10.1145/3025453.3026038>. <http://doi.acm.org/10.1145/3025453.3026038>.
- Lottridge, D., Marschner, E., Wang, E., Romanovsky, M., Nass, C., 2012. Browser design impacts multitasking. Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting. <https://doi.org/10.1177/1071181312561289>.
- Lui, J.H.L., Marcus, D.K., Barry, C.T., 2017. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. *Prof. Psychol.* 48 (3), 199–210. <https://doi.org/10.1037/pro0000122>.
- Lukoff, K., 2019. Digital wellbeing is way more than just reducing screen time. <https://uxdesign.cc/digital-wellbeing-more-than-just-reducing-screen-time-46223db9f057>.
- Lukoff, K., Lyngs, U., Alberts, L., 2022. Designing to support autonomy and reduce psychological reactance in digital self-control tools. *Self-Determination Theory in HCI: Shaping a Research Agenda*. Workshop at the ACM CHI Conference on Human Factors in Computing Systems (CHI'22) 6.
- Lukoff, K., Lyngs, U., Zade, H., Liao, J.V., Choi, J., Fan, K., Munson, S.A., Hiniker, A., 2021. How the Design of YouTube Influences User Sense of Agency. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, Yokohama Japan, pp. 1–17. <https://doi.org/10.1145/3411764.3445467>.
- Lyngs, U., 2021. *Ulysses in Cyberspace: Examining the Effectiveness of Design Patterns for Digital Self-Control*.
- Lyngs, U., 2021. *Ulysses in Cyberspace: Examining the Effectiveness of Design Patterns for Digital Self-Control*.
- Lyngs, U., Binns, R., Van Kleek, M., Shadbolt, N., 2018. “So, Tell Me What Users Want, What They Really, Really Want!”. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. alt04:1–alt04:10. <https://doi.org/10.1145/3170427.3188397>.
- Lyngs, U., Lukoff, K., Slovak, P., Binns, R., Slack, A., Inzlicht, M., Van Kleek, M., Shadbolt, N., 2019. Self-Control in Cyberspace: Applying Dual Systems Theory to a Review of Digital Self-Control Tools. CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019). ACM, New York, NY, USA. <https://doi.org/10.1145/3290605.3300361>.
- Lyngs, U., Lukoff, K., Slovak, P., Seymour, W., Webb, H., Jirotko, M., Zhao, J., Van Kleek, M., Shadbolt, N., 2020. “I Just Want to Hack Myself to Not Get Distracted”: Evaluating Design Interventions for Self-Control on Facebook. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. ACM, Honolulu HI USA, pp. 1–15. <https://doi.org/10.1145/3313831.3376672>.
- Marotta, V., Acquisti, A., 2017. Online Distractions, Website Blockers, and Economic Productivity: A Randomized Field Experiment.
- Miller, W.R., Rose, G.S., 2009. Toward a theory of motivational interviewing. *Am. Psychol.* 64 (6), 527–537. <https://doi.org/10.1037/a0016830>.
- Milyavskaya, M., Saunders, B., Inzlicht, M., 2021. Self-control in daily life: prevalence and effectiveness of diverse self-control strategies. *J. Pers.* 89 (4), 634–651. <https://doi.org/10.1111/jopy.12604>.
- Mohan, A., 2019. LessPhone Launcher. <https://play.google.com/store/apps/details?id=me.aswinmohan.nophone>.
- Monge Roffarello, A., De Russis, L., 2019. The Race Towards Digital Wellbeing: Issues and Opportunities. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 386:1–386:14. <https://doi.org/10.1145/3290605.3300616>.
- Norman, D., Shallice, T., 1986. Attention to Action: Willed and Automatic Control of Behaviour. In: et al. Davidson, R. (Ed.), *Consciousness and Self-regulation (Advances in Research and Theory, Vol. 4)*. Plenum, pp. 1–18.
- Olano, F., 2018. *App-Store-Scraper*.
- Olano, F., 2018. *Google-Play-Scraper*.
- Panichella, S., Sorbo, A.D., Guzman, E., Visaggio, C.A., Canfora, G., Gall, H.C., 2015. How can i improve my app? Classifying user reviews for software maintenance and

- evolution. 2015 IEEE International Conference on Software Maintenance and Evolution (ICSME), pp. 281–290. <https://doi.org/10.1109/ICSM.2015.7332474>.
- Peters, D., Calvo, R.A., Ryan, R.M., 2018. Designing for motivation, engagement and wellbeing in digital experience. *Front. Psychol.* 9, 797. <https://doi.org/10.3389/fpsyg.2018.00797>.
- Pinder, C., Rocca, J.L., Cowan, B.R., Beale, R., 2019. Push Away the Smartphone: Investigating Methods to Counter Problematic Smartphone Use. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. LBW0254:1–LBW0254:6. <https://doi.org/10.1145/3290607.3313028>.
- Przybylski, A.K., Weinstein, N., 2017. A large-scale test of the goldilocks hypothesis. *Psychol Sci* 28 (2), 204–215. <https://doi.org/10.1177/0956797616678438>.
- Deci, R.M.E.L., 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55 (1), 68.
- Ryan, R.M., Deci, E.L., 2017. Self-Determination theory: basic psychological needs in motivation, development, and wellness. The Guilford Press, New York, NY, US. <https://doi.org/10.1521/978.14625/28806>.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., Frith, C.D., 2014. Supra-personal cognitive control and metacognition. *Trends Cogn. Sci. (Regul. Ed.)* 18 (4), 186–193. <https://doi.org/10.1016/j.tics.2014.01.006>.
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T.L., Cohen, J.D., Botvinick, M. M., 2017. Toward a rational and mechanistic account of mental effort. *Annu. Rev. Neurosci.* 40 (1), 99–124. <https://doi.org/10.1146/annurev-neuro-072116-031526>.
- Silge, J., Robinson, D., 2016. Tidytext: text mining and analysis using tidy data principles in r. *J. Open Source Softw.* 1 (3), 37. <https://doi.org/10.21105/joss.00037>.
- Stawarz, K., Preist, C., Tallon, D., Wiles, N., Coyle, D., 2018. User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. *J. Med. Internet Res.* 20 (6) <https://doi.org/10.2196/10120>.
- Sucala, M., Cuijpers, P., Muench, F., CardoNANA, R., Soflau, R., Dobrean, A., Achimas-Cadariu, P., Davi, D., 2017. Anxiety: there is an app for that. A systematic review of anxiety apps. *Depress. Anxiety* 34 (6), 518–525. <https://doi.org/10.1002/da.22654>.
- Tran, J.A., Yang, K.S., Davis, K., Hiniker, A., 2019. Modeling the Engagement-Disengagement Cycle of Compulsive Phone Use. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 312:1–312:14. <https://doi.org/10.1145/3290605.3300542>.
- Tseng, V.W.-S., Lee, M.L., Denoue, L., Avrahami, D., 2019. Overcoming Distractions During Transitions from Break to Work Using a Conversational Website-Blocking System. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 467:1–467:13. <https://doi.org/10.1145/3290605.3300697>.
- Tseng, V.W.-S., Lee, M.L., Denoue, L., Avrahami, D., 2019. Overcoming distractions during transitions from break to work using a conversational website-blocking system. ACM, New York, NY, USA, pp. 467:1–467:13. <https://doi.org/10.1145/3290605.3300697>. <http://doi.acm.org/10.1145/3290605.3300697>. Event-place: Glasgow, Scotland Uk
- van Velthoven, M.H., Powell, J., Powell, G., 2018. Problematic smartphone use: digital approaches to an emerging public health problem. *Digital Health* 4, 1–9. <https://doi.org/10.1177/2055207618759167>.
- Whittaker, S., Kalnikaite, V., Hollis, V., Guydish, A., 2016. ‘Don’ t Waste My Time’: Use of Time Information Improves Focus. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. ACM, New York, pp. 1729–1738.
- Wu, T., 2016. The attention merchants: the epic scramble to get inside our heads. Knopf Publishing Group.
- Xu, X., Zou, T., Xiao, H., Li, Y., Wang, R., Yuan, T., Wang, Y., Shi, Y., Mankoff, J., Dey, A. K., 2022. TypeOut: Leveraging Just-in-Time Self-Affirmation for Smartphone Overuse Reduction. CHI Conference on Human Factors in Computing Systems. ACM, New Orleans LA USA, pp. 1–17. <https://doi.org/10.1145/3491102.3517476>.
- Zhang, M.R., Lukoff, K., Rao, R., Baughan, A., Hiniker, A., 2022. Monitoring Screen Time or Redesigning It?: Two Approaches to Supporting Intentional Social Media Use. CHI Conference on Human Factors in Computing Systems. ACM, New Orleans LA USA, pp. 1–19. <https://doi.org/10.1145/3491102.3517722>.
- Forest: stay focused, 2018–. (Accessed 13 June 2022).