WORKING PAPER · VERSION 1.5 · OCTOBER 2026

A Behavioural-Science Framework for Problematic Smartphone Use

Theory, Design Principles and Implementation

Farhan Aktar

Creator of Beave · https://beave.app

Abstract. Many people say they use their phones more than they would like to, and a growing body of evidence suggests they are right. Pooled survey data put the prevalence of problematic smartphone use (PSU) at somewhere near one in four, although the figure moves a good deal with the instrument used. Most of the remedies on offer (screen-time dashboards, resolutions, blocking apps, the occasional 'detox') work on one part of the problem at a time. In this paper I bring together findings from habit research, behavioural economics, dual-process psychology and HCI, and argue that compulsive use is held in place by at least seven distinct mechanisms, for each of which there is a counter-measure with experimental support of its own. I organise these counter-measures into a five-layer framework, the Dam Model (Measure, Plan, Pre-commit, Interrupt, Reinforce), bound by seven design principles, and then document how the framework is implemented in Beave, an Android app. The implementation reconstructs usage from the operating system's event log, caps goals by the user's own seven-day baseline, enforces if-then rules from a background service, delays any loosening of those rules, prices the opening of a blocked app at three slow breaths, asks at each unlock what the phone has been picked up for, interrupts long scrolling sittings with a fixed break and verifies focus sessions by having the phone lie face down. Eighteen technical drawings document the design.

Keywords: problematic smartphone use; digital self-control tools; commitment devices; friction; implementation intentions; behaviour change techniques; digital wellbeing.

PLAIN-LANGUAGE SUMMARY

Most people who want to cut down on their phone already know that they use it too much, so a lack of information is rarely what stands in the way. What gets in the way are particular moments: picking the phone up without having decided to, or 'five more minutes' that turns into forty. Research has identified a handful of separate causes behind those moments, and for most of them there is something that helps, such as deciding ahead of time, making that decision hard to reverse on impulse, putting a short pause between the urge and the app, and giving time spent away from the phone a reward that can be seen. Beave stacks these in layers, rather as a beaver's dam slows a river down while still letting water through, so that the phone stays fully usable but gets used on purpose.

1 Introduction

Phones now take up a large, and by most measures still growing, share of waking life. Andrews et al. (2015) logged around 85 phone checks a day among young adults, most of whom guessed

a much lower figure, and in a survey reported by Allcott et al. (2022, Figure 1), smartphone and social media use ranked among the five activities that Americans most often said they did 'too much' or 'too little', next to exercise, retirement saving and healthy eating. On prevalence, Olson et al. (2022) pooled studies that used the Smartphone Addiction Scale in 24 countries ($N = 33{,}831$) and found that scores went up between 2014 and 2020, while a larger meta-analysis put global prevalence at about 27% (Meng et al., 2022), with a similar picture, and consistent links to depression, anxiety and poor sleep, among children and young people (Sohn et al., 2019).[1]

Whether any of this deserves the word 'addiction' is disputed. Smartphone addiction is not a diagnosis in either *DSM-5* or *ICD-11*, where the nearest recognised category is gaming disorder (American Psychiatric Association, 2013; World Health Organization, 2019); several authors think the label runs ahead of the evidence for harm and prefer 'problematic smartphone use' (Panova & Carbonell, 2018; Billieux et al., 2015), and Kardefelt-Winther (2014) has warned that heavy use is sometimes a way of coping with an unrelated problem and should not be read as a disorder in its own right. I follow them and use PSU throughout, meaning use that someone repeatedly intends to cut down and does not. Nothing described here is meant as treatment for a clinical condition.

The evidence on harm is also more mixed than the public debate tends to suggest. Cross-sectional associations are small: in Orben and Przybylski's (2019) specification-curve analysis of three large adolescent datasets, technology use explained at most 0.4% of the variation in well-being, though associations get stronger at the heaviest levels of use (Twenge & Campbell, 2018) and PSU, as opposed to time on screen, correlates fairly consistently with depression and anxiety (Elhai et al., 2017). Since such data cannot tell us which way causation runs (unhappy people may well just use their phones more), experiments carry more weight, and the few that exist lean the same way. Hunt et al. (2018) limited 143 undergraduates to ten minutes per platform per day on Facebook, Instagram and Snapchat for three weeks and found less loneliness and depression than in a control group. In Allcott et al. (2020), 2,743 Facebook users who were paid to deactivate their accounts for the four weeks before the 2018 US midterm elections reported higher subjective well-being, spent more time with friends and family offline, and went on using Facebook less once the experiment was over.

For the purposes of this paper the most useful study is the one by Allcott et al. (2022). Their participants, about 2,000 American adults with Android phones, installed an app built for the experiment, Phone Dashboard, and a randomly chosen group were given the ability to set daily limits on individual apps. Changes to a limit only took effect the next day, and in most cases a limit could not be overridden on the spot, unlike the flexible limits in Android's Digital Wellbeing or iOS Screen Time. Nothing rewarded them for using the feature, yet 78% of those offered it set binding limits and kept them until the end of the study, and their use of the six most tempting apps fell by 22 minutes a day (16%) over twelve weeks. In a second arm, paying people to cut back reduced use by 56 minutes a day while the payments ran and by 19 minutes a day in the three weeks after they stopped, which the authors take as evidence of habit formation; and the control group repeatedly underestimated how much they would use their phones in the following weeks. Fitting a structural model to these results, they estimate that self-control problems account for 31% of use of those apps. People, on this evidence, use their phones more than they themselves would choose to, and most of them will take a credible way of holding themselves to a limit when one is offered.

[1] Prevalence depends heavily on which instrument and cut-off a study uses, so these figures are best read as an indication of scale.

If that is right, the remaining questions are mostly about design. Which parts of a self-control tool actually do the work? How strict can a tool be before people simply remove it? And can such a tool be built without becoming one more product competing for the user's attention? The default answer from the industry, the screen-time dashboard, deals with awareness and little else, and the evidence for awareness on its own is not encouraging. In a two-month experiment, pop-up messages about heavy use did not reduce screen time, phone checking or self-reported PSU (Loid et al., 2020), and consumers say they are happy to monitor their usage but do not expect to reduce it (Zimmermann, 2021). The gap between intending to change and changing is among the best-replicated findings in behavioural science (Sheeran & Webb, 2016), and with phones it is particularly wide, given that the device is picked up dozens of times a day, pays out within seconds, and is checked to a large extent out of habit (Oulasvirta et al., 2012).

The paper makes three contributions. First, it brings together in one place the mechanisms that the psychological, economic and HCI literatures hold responsible for compulsive phone use, and assesses the usual remedies against them (Sections 2 and 3). Second, it proposes a framework, which I call the Dam Model, that pairs each mechanism with a counter-measure backed by independent experimental evidence, sets out seven principles that constrain it, and codes its components against the Behaviour Change Technique Taxonomy (Section 4). Third, it documents an implementation of the framework in Beave, an Android app, down to its parameter values, so that the design can be copied or criticised in detail (Section 5, with the social features in Section 5.7), and specifies how it ought to be evaluated (Section 7). I develop Beave and have a financial interest in it (see Declarations), which is the main reason the paper stops short of any claim about its effectiveness.

2 Mechanisms that sustain compulsive use

Accounts of compulsive phone use differ more in vocabulary than in substance. In the I-PACE model, predisposing variables combine with affective and cognitive responses to cues to produce craving, and over time use drifts from being driven by gratification towards compensation and habit while inhibitory control weakens (Brand et al., 2016, 2019); Tran et al. (2019) describe much the same cycle from the user's side, as a trigger that prompts engagement, a stretch of time that goes unnoticed, and a disengagement that often comes with regret. Below I distinguish seven mechanisms, M1 to M7, which are drawn together in Figure 1. They overlap a good deal, and the reason for keeping them apart is a practical one, namely that each points to a different kind of intervention.

M1. Cue-triggered habit. Habits are learned associations between a context and a response, and once they are established the context sets off the response with little reference to what the person currently wants (Wood & Neal, 2007; Wood & Rünger, 2016). Oulasvirta et al. (2012) called the smartphone version 'checking habits', i.e., brief, repetitive inspections of dynamic content that is always quickly accessible, kept going by small informational rewards and liable to turn into longer sessions, and users themselves rate this sort of use as the least meaningful time they spend on their phones (Lukoff et al., 2018). Because the habit runs largely independently of intention, interventions aimed at intention tend to leave it untouched (Verplanken & Wood, 2006). New habits are also slow to form; in a 12-week study of newly adopted daily behaviours, Lally et al. (2010) estimated a median of 66 days to reach a plateau of automaticity, with individual estimates ranging from 18 to 254 days. The persistence that Allcott et al. (2022) observed after their payments ended suggests that phone habits are formed (and possibly broken) on a time scale of weeks. In practice this leaves two options, which are to break the link between cue and

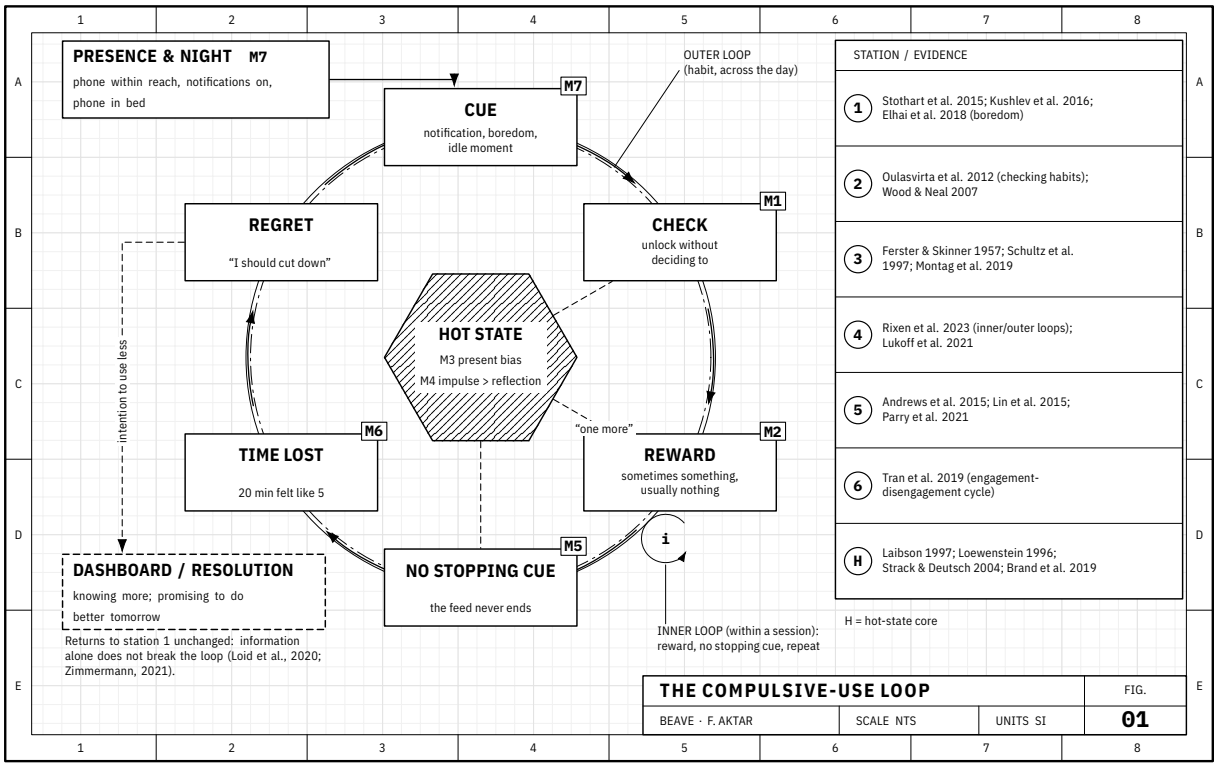


Figure 1. The compulsive-use loop. A cue (a notification, boredom, an idle moment) triggers a habitual check (M1). The app pays out on an unpredictable schedule (M2), in a feed with no stopping point (M5), while elapsed time is underestimated (M6). The decision to carry on is taken in an affectively 'hot' state that discounts later costs (M3, M4), and the device's constant proximity, including in bed, keeps the loop within reach (M7). The table on the sheet lists the principal sources for each station.

response or to change the context in which the cue turns up (Wood et al., 2005).

M2. Variable reward. Feeds, notifications and incoming messages pay out unpredictably. Of the schedules studied in operant research, variable-ratio reinforcement sustained the highest and most persistent rates of responding (Ferster & Skinner, 1957). Dopamine neurons in the midbrain signal the difference between the reward that was predicted and the one received (Schultz et al., 1997), and repeated pairing of cue and reward can sensitise 'wanting' without any matching increase in 'liking' (Robinson & Berridge, 1993). Social media and freemium games are, by design, built around such schedules (Montag et al., 2019).

M3. Present bias. People discount the near future steeply and the far future only gently, so that a plan which looks sensible a week ahead tends to be dropped when the moment comes (Ainslie, 1975; Laibson, 1997). Those who do not recognise this in themselves keep expecting to do better tomorrow (O'Donoghue & Rabin, 1999); the control group in Allcott et al. (2022), who underestimated their future use by about 6 minutes a day (4%) in every survey, are a mild example of exactly this kind of naivety.

M4. Hot-state decisions. Decisions taken while craving something differ systematically from those taken calmly, and people in one state are poor at predicting how they will behave in the other (Loewenstein, 1996). Dual-system models distinguish an impulsive system that responds quickly to cues from a reflective one that depends on spare cognitive capacity (Strack & Deutsch, 2004; Hofmann et al., 2009). Lyngs et al. (2019) used this distinction to classify 367 apps and browser extensions for digital self-control by the system each one targets, and found some design

approaches (habit scaffolding in particular) to be rare among them.

M5. Absent stopping cues. Infinite scroll and autoplay removed the ends of pages and episodes that used to prompt a decision about whether to stop. In a week-long field study of infinite scrolling, Rixen et al. (2023) found loops nested inside loops, one within each session and a habitual one across sessions, and most of the reasons participants gave for stopping had to do with things outside the app. YouTube's interface has been shown to erode users' sense of agency through comparable features (Lukoff et al., 2021), and 'doomscrolling' now has a validated scale of its own (Sharma et al., 2022).

M6. Time distortion. People are poor judges of their own phone use. Besides the underestimates recorded by Andrews et al. (2015), a meta-analysis of 106 effect sizes found self-reports to be only moderately related to logged use, and problematic-use scales even less so (Parry et al., 2021). A distorted sense of elapsed time has been described as characteristic of smartphone addiction specifically (Lin et al., 2015).

M7. Proximity and night-time use. Having the phone nearby is a cue in itself. Receiving a notification hurts performance on an attention task about as much as answering one does (Stothart et al., 2015), and in a field experiment, keeping notifications switched on raised everyday symptoms of inattention and hyperactivity (Kushlev et al., 2016). Some studies report that the mere presence of a phone lowers available cognitive capacity or the quality of a conversation (Ward et al., 2017; Thornton et al., 2014; Przybylski & Weinstein, 2013), though a direct replication of the best known of them did not find the effect (Ruiz Pardo & Minda, 2022). At night, phone use goes with later and poorer sleep (Exelmans & Van den Bulck, 2016; Levenson et al., 2016), and in a controlled crossover study, reading on a light-emitting e-reader before bed (compared with a printed book) delayed circadian timing and left participants less alert the next morning (Chang et al., 2015).

3 Existing remedies and their limits

Table 1 sets the usual remedies against these seven mechanisms. Most of them engage one or two, which goes some way towards explaining why each, taken alone, tends to disappoint.

Dashboards and reports. Usage statistics address M6 and nothing else. Self-monitoring does help people reach goals (across 138 studies, Harkin et al., 2016 found an average effect of $d = 0.40$, larger when progress was physically recorded or reported to other people), but with no goal or consequence attached, the experiments mentioned in Section 1 found no change in use (Loid et al., 2020; Zimmermann, 2021).

Willpower. Relying on resisting each urge as it arises means confronting M4 at its strongest. In an experience-sampling study of 7,827 everyday desires, participants tried to resist a large share of those that conflicted with their other goals, and quite often failed (Hofmann et al., 2012). Choosing or changing one's situation before temptation arrives generally works better, and takes less effort, than suppressing an impulse once it is there (Duckworth et al., 2016).

Hard blocks. Blockers act on M1 and M2 by removing access, and blocking distracting sites at work has been found to improve focus (Mark et al., 2017). But a block with no exception threatens people's sense of autonomy, and threatened freedoms provoke what Brehm (1966) called reactance, in which the forbidden option gains appeal and effort goes into getting it back (Steindl et al., 2015). Reviews of digital self-control tools report that rigid restrictions are often

Table 1. Common remedies and the mechanisms they engage. • direct effect expected; ◦ partial or indirect effect; blank, none expected. The final row anticipates the framework of Section 4.

Remedy	M1 habit	M2 reward	M3 present bias	M4 hot state	M5 no stop cue	M6 time sense	M7 presence, night
Screen-time dashboard						•	
Willpower, resolutions			◦				
Hard app blocker	•	•					◦
Digital detox, deactivation	•	•					•
One-off nudge (greyscale, pop-up)		◦		◦		◦	
Focus timer alone							◦
Dam Model (this paper)	•	•	•	•	•	•	•

found frustrating and are easy to get round (Lyngs et al., 2019; Monge Roffarello & De Russis, 2019), and in the workplace study just cited, participants low in self-control worked longer without breaks and reported more stress once distractions were blocked (Mark et al., 2017).

Detox and deactivation. Deactivation has measurable benefits (Allcott et al., 2020), but a systematic review of digital detox interventions found mixed and often short-lived effects (Radtke et al., 2022), and long abstinence from a device that is also one's wallet, map and office is not realistic for most people anyway.

One-off nudges. Setting the screen to greyscale cuts screen time a little (Holte & Ferraro, 2023). A meta-analysis of choice-architecture interventions found an average effect of $d = 0.43$ (Mertens et al., 2022), which disappeared once Maier et al. (2022) corrected for publication bias, so expectations of any single nudge should probably be modest. The best evidence for a single mechanism in this area concerns friction at the moment an app is opened (Section 5.5), and the framework uses it as one of its five layers.

Existing tools that combine approaches. Several commercial tools already go beyond a single remedy, and they amount, as Lyngs et al. (2019) put it, to a large set of natural experiments in supporting self-control. Forest, which ties abstaining from the phone to the growth of a virtual tree, had over five million Android users at the time of that review; *one sec* adds a short delay and an option to cancel whenever a chosen app is opened (Grüning et al., 2023); Phone Dashboard paired measurement with limits that could not be changed on the same day (Allcott et al., 2022); and some browser extensions make blocking deliberately hard to switch off, which, as Lyngs et al. (2019) note, raises the question of how far a tool should go in holding users to their past preferences. The framework below draws on all of these, and P1 and P3 in Section 4.1 are my answer to that last question.

4 The Dam Model

I use the embankment dam as an organising metaphor, and not only because of the app's name. A dam regulates a river by virtue of its structure, without anyone having to hold it in place once it is built; it is assembled from parts that do different jobs (a foundation, a core, outer shells, outlet works); and it has to let water through, since engineers know that a dam without outlets

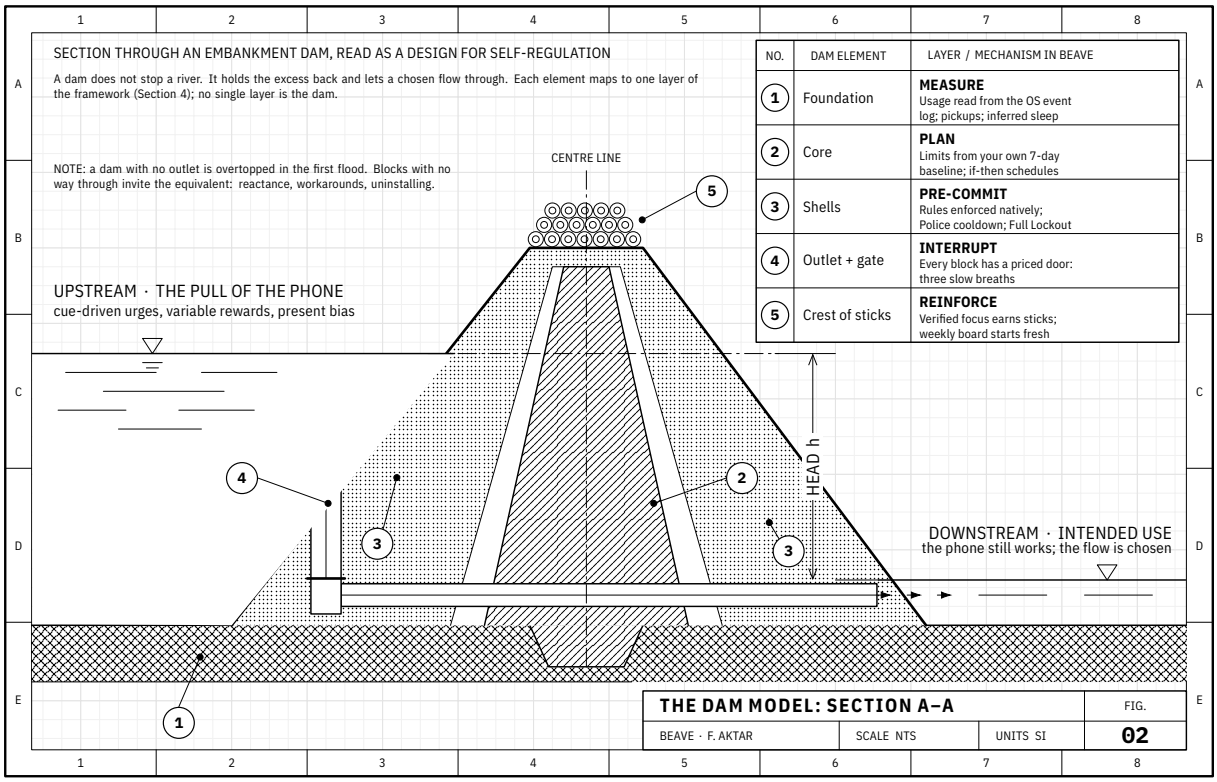


Figure 2. The Dam Model as a section through an embankment dam. The foundation (1) corresponds to measurement and the clay core (2) to the plan. The outer shells (3) are the commitment devices that keep the plan in place without continuous effort from the user. The outlet and its gate (4) are the interruptions, a passage that is always open but costs something to use, and the crest of sticks (5) is reinforcement, raised a little by each verified focus session. The legend names the Beave feature behind each element.

or spillways will eventually be overtopped and then fail all at once. In software the equivalent of overtopping is a blocker that gets uninstalled on the first bad evening.

4.1 Design principles

The framework is defined partly by what it rules out. Seven principles governed the design of Beave, each of them a response either to one of the failure modes described in Section 3 or to a risk peculiar to software that restricts its own user.

P1. Limit rather than forbid. Outside a focus session the user has deliberately started, every block has a way through. Passing costs some effort, but the passage is always there, and the reason is reactance (Brehm, 1966), which in this domain mostly shows up as people uninstalling the tool.

P2. Decide in a cold state, enforce in a hot one. Plans are made while the user is calm and enforced at the moment of temptation, by something the tempted self cannot easily renegotiate (Loewenstein, 1996; Duckworth et al., 2016).

P3. Tighten immediately, loosen with delay. A change that makes a plan stricter takes effect at once, whereas one that relaxes it may be made to wait. This is a form of what Camerer et al. (2003) call asymmetric paternalism, which aims at large benefits for people prone to error at little cost to those who are not.

P4. Price exits in effort, never in money. No way out of a block can be bought. Money-

based commitment contracts can work (smokers in the Philippines who put their own savings at stake, to be forfeited on a failed urine test, were more likely to have quit six months later; Giné et al., 2010), but making users pay to regain control of their own phone would set the app against the very people it is supposed to serve.

P5. Measure accurately and keep data local. Limits are enforced against the operating system's own record of use, and detailed usage data stay on the device.

P6. Reward only what can be verified. Anything that confers standing with other people has to be measured by the device, never self-reported, since a measure that becomes a target soon stops being a good measure (Strathern, 1997).

P7. Keep progress displays away from the problem behaviour. Customers in a café loyalty scheme bought coffee more often the closer they got to a free one (Kivetz et al., 2006), and a running total of minutes scrolled would turn that same goal-gradient effect against the user. The only charts in Beave are therefore ones on which a fall is good news, and the only totals that keep growing measure time away from the phone.

4.2 Five layers

Each layer in Figure 2 answers one or more of the mechanisms in Section 2.

Layer 1: Measure (foundation). Usage, pickups and sleep, as recorded by the device. The layer addresses M6 directly and carries everything else, because a limit enforced against a wrong figure is either useless or unfair.

Layer 2: Plan (core). Goals set from the user's measured baseline and written as if-then rules (if it is after 22:00, Instagram is closed). Plans of this form, which Gollwitzer (1999) called *implementation intentions*, hand the job of starting an action over to a cue in the environment, and across 94 independent tests they had a medium-to-large effect on goal attainment, $d = 0.65$ (Gollwitzer & Sheeran, 2006). Specific goals that are hard but attainable do better than vague ones (Locke & Latham, 2002), and small early successes build the self-efficacy needed for later steps (Bandura, 1977). The layer works on M1, by attaching a new response to an existing cue, and on M3, by moving the decision forward in time.

Layer 3: Pre-commit (shells). Mechanisms that keep the plan in force when the user would rather they did not, namely rules enforced by a background service that does not depend on the interface, a delay on any loosening of those rules, and an optional lockout during focus sessions. Economists call these commitment devices (Bryan et al., 2010; Schelling, 1984), and Ulysses bound to the mast is the usual illustration (Elster, 1979). The layer is aimed at M3 and M4.

Layer 4: Interrupt (outlet and gate). Friction at the moment of impulse, timed so that it can still change what happens. In Beave it takes the form of a short breathing exercise before a blocked app opens, a question at each unlock about what the phone is for, a break inside an endless feed, and a notification at 80% of the daily limit (and not at 100%, by which point it is too late to act on). Interruption breaks the automatic sequence (M1), gives the reflective system a chance to engage (M4) and puts back a stopping cue (M5).

Layer 5: Reinforce (crest). Something else to want: focus sessions that only count while the phone lies face down, sticks that build a visible dam, and a weekly leaderboard of verified

effort. The layer offers a competing reinforcer (M2) and takes the phone physically out of use (M7).

4.3 Mapping to behaviour change techniques

To make the intervention comparable with others, Table 2 codes each feature against the Behaviour Change Technique Taxonomy v1, a consensus classification of 93 techniques in 16 groups (Michie et al., 2013). Nineteen techniques from ten of those groups are represented. The spread is deliberate: a recent systematic review and meta-analysis of digital self-control tools found the evidence for any one mechanism to be limited (Monge Roffarello & De Russis, 2023), which is an argument against relying on just one.

5 Implementation in Beave

Beave is distributed through Google Play for Android, and what follows describes version 3.8.5 as of October 2026. I report parameter values throughout, since a framework cannot be tested without them, and Appendix A lists each one together with the reason for it. Some features belong to the paid tier, Beave Gold; the principles in Section 4.1 apply to paid and free features alike, and in particular no way out of a block is ever put behind a paywall (P4).

5.1 Architecture

An app that enforces a plan only while it happens to be open offers very little in the way of commitment, so Beave keeps planning and enforcement apart (Figure 3). Rules, limits and time windows are written in the interface, compiled into a compact rule set and handed to a native enforcement service, which runs in the background, carries on when the interface is closed and starts again after a reboot. Every 500 ms the service works out which app is in the foreground and checks it against the rules, the time of day and the day's usage so far. Where the interface and the service disagree about whether something should be blocked, the service wins, and the interface only reproduces its logic in order to show the user what is going on.

Three things follow from P5. Usage is computed from the operating system's own event log; detailed usage, sleep and scroll data are kept on the device and never uploaded; and the only data that leave the phone for any of these features are the stick and flip counts behind the leaderboard, which the server credits itself instead of taking the phone's word for them (Section 5.7).

The interface has four tabs (Figure 4). *Block*, the home tab, holds the daily limit, the user's rules, the morning and night windows and four tiles for the scroll limit, offline hours, Unlock check and Beave Police; *Focus* runs focus sessions; *Progress* shows measured usage and the wellness score; and *Dam* shows the sticks collected so far, the map of rivers and the leaderboard.

5.2 Layer 1: Measure

Usage. Screen time is rebuilt by walking through Android's usage event log, merging each app's foreground intervals and clipping every interval to the calendar day. Beave does not use the platform's pre-aggregated daily summaries, because their buckets are not aligned with local midnight, and a request for 'today' can come back with time that was actually spent the evening before (Figure 5, panel C). From the user's side an error of this sort cannot be seen (an app is simply blocked sooner than expected), but it eats away at trust in every block that follows, which is why accurate measurement is treated in Beave as a design requirement and not as an implementation detail.

Some foreground time is not use and is subtracted from screen time, namely time in Beave

Table 2. Mechanisms, layers, Beave features, principal evidence and behaviour change technique codes (BCTTv1; Michie et al., 2013). Section numbers in parentheses refer to the description of each feature.

Mech.	Layer	Beave feature (§)	Principal evidence	BCT codes
M6	Measure	Event-log usage, pickups, inferred sleep (5.2)	Parry et al. (2021); Andrews et al. (2015); Lin et al. (2015)	2.3 Self-monitoring of behaviour
M6	Measure	Digital Wellness Score with benchmark and ranked suggestions (5.2)	Harkin et al. (2016)	2.2 Feedback on behaviour; 2.7 Feedback on outcomes; 6.2 Social comparison
M3, M6	Plan	Goal from 7-day baseline, $\geq$25% cut (5.3)	Locke & Latham (2002); Bandura (1977)	1.1 Goal setting (behaviour); 1.6 Discrepancy; 8.7 Graded tasks
M1, M3	Plan	If-then rules; morning and night windows (5.3)	Gollwitzer & Sheeran (2006); Exelmans & Van den Bulck (2016)	1.4 Action planning; 12.3 Avoidance of cues
M3, M4	Pre-commit	Enforcement independent of the interface (5.4)	Duckworth et al. (2016); Allcott et al. (2022)	12.1 Restructuring the environment
M3, M4	Pre-commit	Beave Police: delayed loosening (5.4)	Camerer et al. (2003); Loewenstein (1996); Bryan et al. (2010); Allcott et al. (2022)	1.9 Commitment
M4, M7	Pre-commit	Full Lockout; offline hours (5.4)	Elster (1979); Mark et al. (2017)	1.9 Commitment; 12.3 Avoidance of cues
M1, M4	Interrupt	Block overlay with breathing gate (5.5)	Grüning et al. (2023); Kim et al. (2019); Zaccaro et al. (2018); Balban et al. (2023)	14.1 Behaviour cost; 11.2 Reduce negative emotions
M1, M6	Interrupt	Unlock check: stated purpose, drift nudge, check-in (5.5)	Gollwitzer (1999); Oulasvirta et al. (2012); Tran et al. (2019)	1.4 Action planning; 7.1 Prompts and cues
M5	Interrupt	Anti-doomscroll break (5.5)	Rixen et al. (2023); Lyngs et al. (2020)	7.1 Prompts and cues; 14.1 Behaviour cost
M6, M3	Interrupt	80% notification; time-used pill; 21:00 recap (5.5)	Okeke et al. (2018); Korotitsch & Nelson-Gray (1999)	7.1 Prompts and cues; 2.2 Feedback on behaviour
M2, M7	Reinforce	Flip-verified focus sessions and Pomodoro runs (5.6)	Stothart et al. (2015); Ward et al. (2017); Ruiz Pardo & Minda (2022)	8.2 Behaviour substitution; 12.3 Avoidance of cues
M2	Reinforce	Sticks, dam, map regions (5.6)	Kivetz et al. (2006); Hamari et al. (2014)	14.5 Rewarding completion; 10.3 Non-specific reward
M2	Reinforce	Weekly leaderboard (5.7)	Patel et al. (2019); Garcia et al. (2013); Dai et al. (2014); Landers et al. (2017)	6.2 Social comparison; 10.4 Social reward
M2	Reinforce	Share cards and badges (5.7)	Aral & Nicolaides (2017); Harkin et al. (2016)	10.4 Social reward; 13.5 Identity associated with changed behaviour

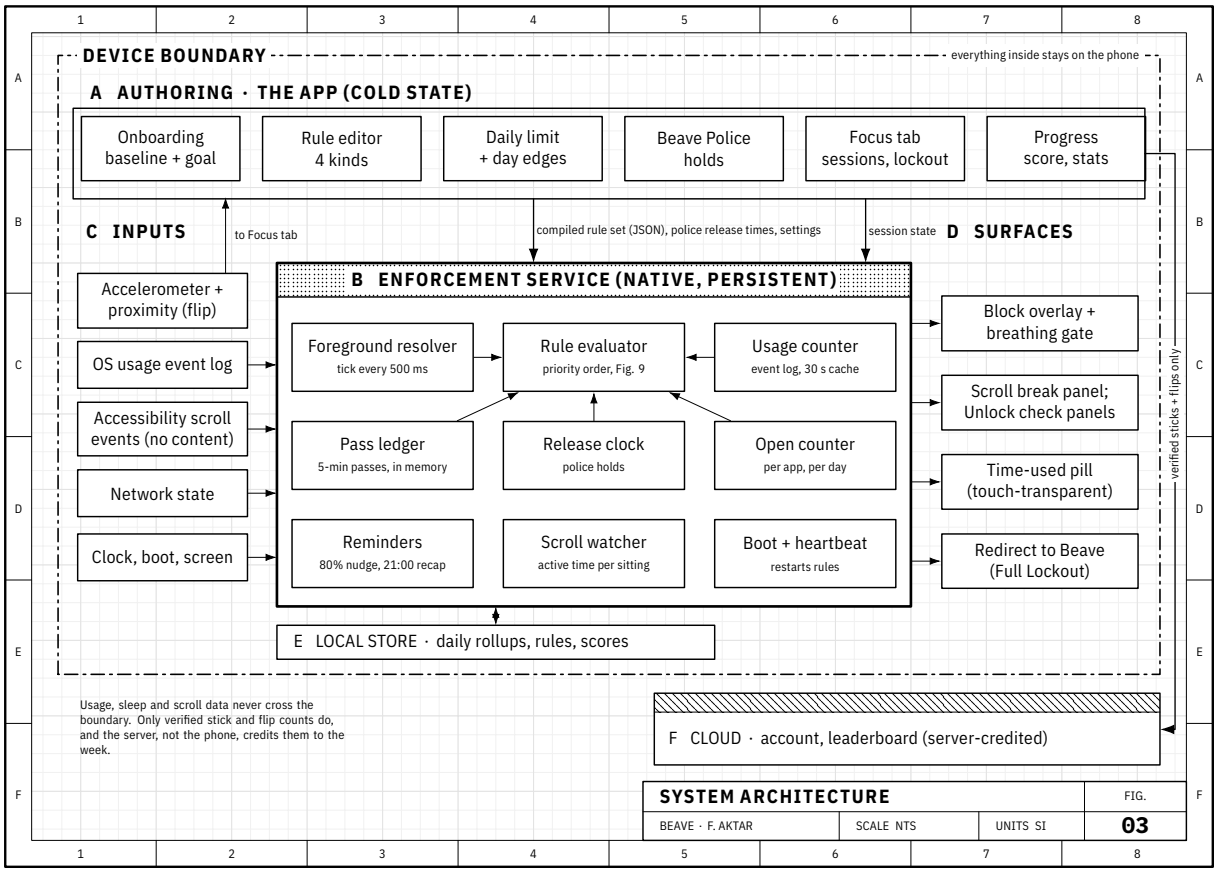


Figure 3. System architecture. Plans are written in the interface (top) and enforced by a native service whether or not the interface is running (centre). Sensors and the operating system's usage log feed the service (left), which reaches the user only through the surfaces on the right. Everything inside the dashed boundary stays on the device; only verified stick and flip counts cross it.

itself, in alarm apps and in the dialler. A ringing alarm keeps the screen on with the clock in front, and during a call the proximity sensor darkens the screen while the system still reports the phone as in use, so a single alarm or call could otherwise account for a large share of a light day. The exclusion is tied to the app and not to a period of time, so another app opened during a call still counts, and pickups and sleep are worked out from the unadjusted record. Calls made inside messaging apps cannot be told apart from messaging and are still counted. Users cannot add exclusions of their own, since that would make the figure easy to game.

Pickups and sleep. A pickup is counted on each unlock, or on each screen-on event for phones that do not report unlocks. Sleep is inferred and never asked for, as the longest overnight stretch without real use, and foreground activity of less than 60 seconds (reading a notification, say) does not break it. A fixed night window would penalise shift workers, so Beave learns each user's typical sleep midpoint as the circular mean of their past overnight gaps[2] and accepts gaps centred within four hours of it; with fewer than five days of history, or a circular standard deviation of more than three hours, it falls back on a fixed window (Figure 5).

The Digital Wellness Score. Raw figures are hard to make sense of, and the Progress tab therefore condenses the last seven days into one score from 0 to 100, built from five factors (Table 3, Figure 6). Each factor is scored from 0 to 1 along a straight line between two thresholds,

[2]The circular mean is $\bar{\theta} = \operatorname{atan2}(\sum_i \sin\theta_i,\ \sum_i \cos\theta_i)$, where the θ_i are clock times expressed as angles on a 24-hour circle. An arithmetic mean has no sensible meaning for times on either side of midnight.

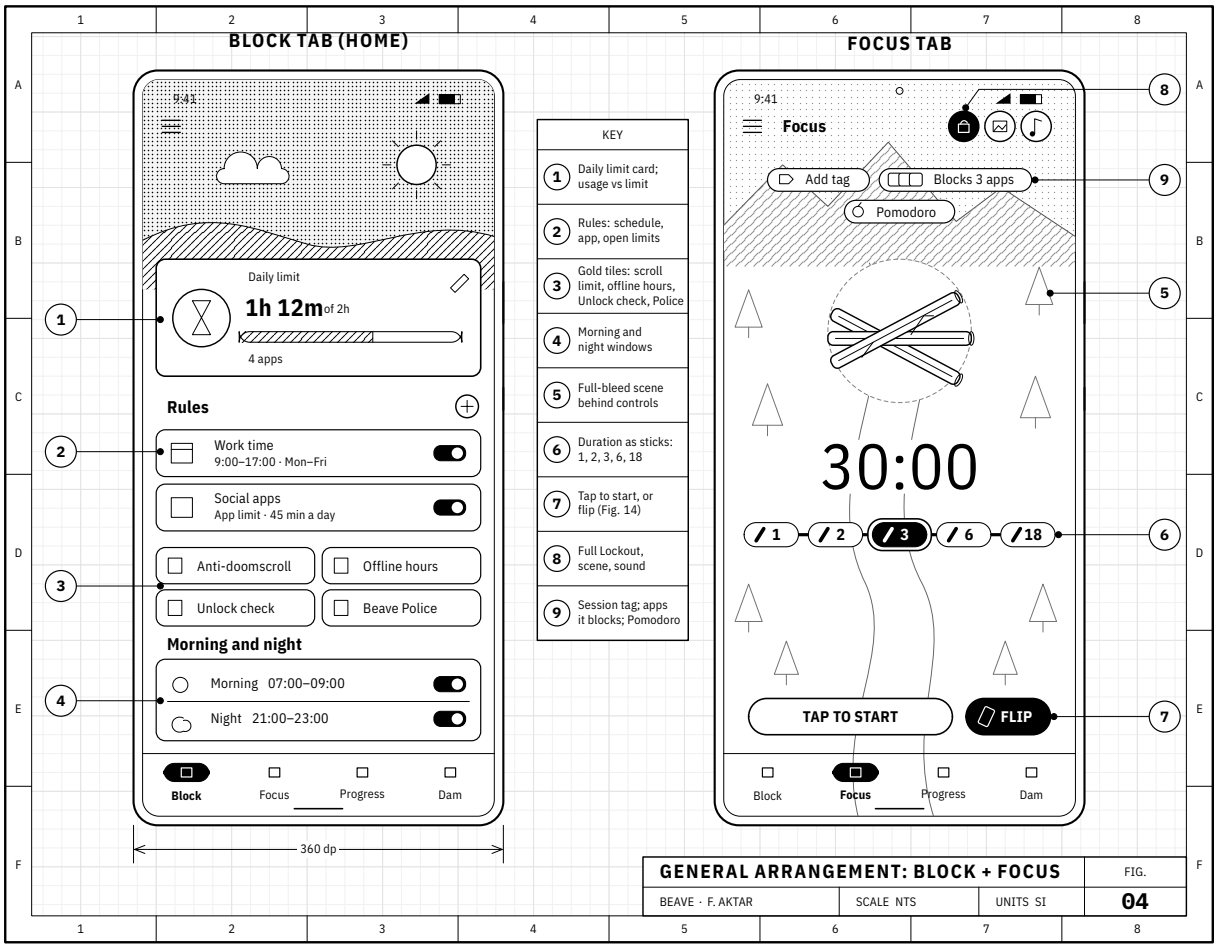


Figure 4. Wireframes of the two main screens. Left, the Block tab: daily-limit card over an animated sky (1), rules (2), tiles for the scroll limit, offline hours, Unlock check and Beave Police (3), and the morning-and-night card (4). Right, the Focus tab: full-screen scene (5); a stepper that sets session length by the sticks it earns, 1, 2, 3, 6 or 18 for 10, 20, 30, 60 or 180 minutes (6); the TAP TO START and FLIP controls (7); the Full Lockout padlock beside the scene and sound controls (8); and chips for the session tag, the apps to be blocked and a Pomodoro run (9). On-screen text is illustrative.

and the weighted sum is scaled to 100 and shown as a whole number. The scale is 100 and not 10 so that a week that improves moves the number visibly; on a scale of 10, a change from 6.9 to 7.1 reads as rounding. Days with less than 12 hours on the clock are left out, which keeps the score from starting every morning close to its maximum, and nothing is shown until at least two usable days exist.

Two decisions about the score need explaining. The thresholds are absolute, not derived from the user's own limit, since otherwise someone who exactly met a six-hour limit would score the same as someone who exactly met a one-hour limit and the scores could not be compared. Time spent in Beave's own focus sessions is also left out, so that the score reflects phone habits and cannot be pushed up just by using Beave. Next to the user's score the Progress tab shows that of a 'typical phone user' put together from published averages (roughly 4 h 40 min of screen time, 6 h 51 min of sleep, 70 pickups and 6 late-night pickups a day; cf. Andrews et al., 2015), a profile that scores 65. This is a fixed point of reference and not a percentile, because a percentile would need other users' scores, and Beave does not collect them. The difference from it is reported in whole points, and the score itself is labelled with one of six fixed bands: Poor below 25, then Low from 25, Fair from 40, Average from 55, Good from 70 and Excellent from 85. The typical

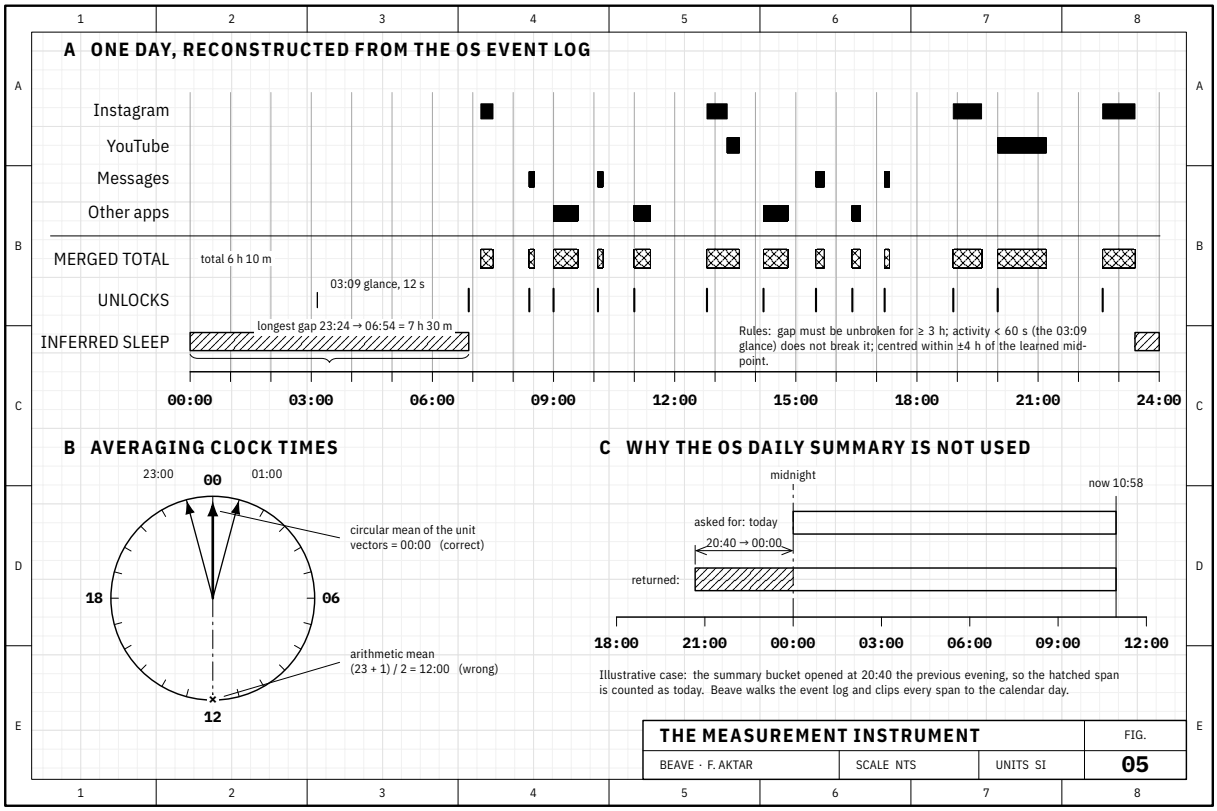


Figure 5. Measurement of a single day. Top: a day rebuilt from the event log, showing each app's foreground intervals (merged into the screen-time total), unlock events (pickups) and the longest overnight gap (sleep). Bottom left: clock times have to be averaged as angles; the arithmetic mean of 23:00 and 01:00 is 12:00, the circular mean 00:00. Bottom right: a daily summary bucket that opened the previous evening, and the overcount of 'today' that results.

Table 3. Components of the Digital Wellness Score. A factor scores 1 at or beyond the full-marks threshold, 0 at or beyond the zero threshold, and linearly in between. Sleep is scored as a band: 7–9 h scores 1, falling to 0 at 4 h and at 11 h, following adult sleep-duration recommendations (Hirshkowitz et al., 2015; Watson et al., 2015).

Factor	Weight	Full marks	Zero	Rationale
Screen time per day	30%	$\leq$ 2 h	$\geq$ 7 h	Primary target of most goals
Sleep	20%	7–9 h	$\leq$ 4 h or $\geq$ 11 h	Displaced by night-time use (M7)
Pickups per day	20%	$\leq$ 30	$\geq$ 120	Checking habits (M1)
Minutes per pickup	15%	$\leq$ 3	$\geq$ 12	Extended sessions (M5)
Pickups 23:00–07:00	15%	0	$\geq$ 12	Sleep disruption (M7)

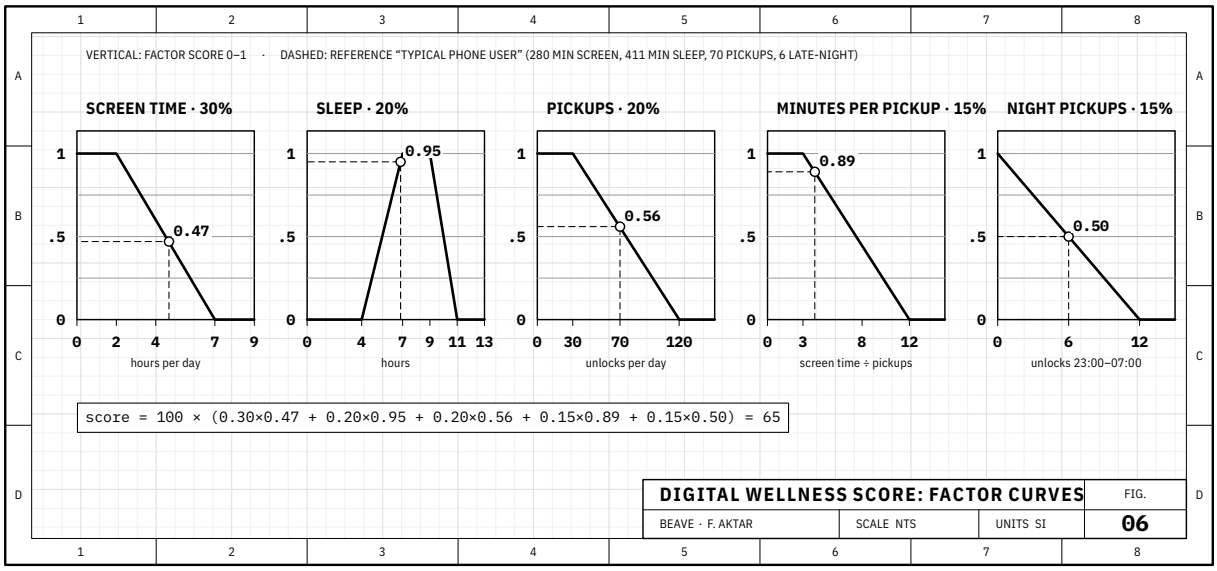


Figure 6. The five scoring functions, drawn to scale. Each panel maps one day's measurement (horizontal axis) to a factor score between 0 and 1 (vertical axis), with the factor's weight in the panel title. Dashed lines mark the typical-user profile, which scores 65 once weighted.

profile falls inside Average, so a user who beats it by a few points is told, accurately, that their week was still average.

Behind the score sits a short list of suggestions. The five factors are ranked by how many points each could still add, that is, its weight multiplied by its shortfall from full marks, and each entry quotes the user's own weekly average against the full-marks threshold (for instance, about 70 pickups a day against 30 or fewer) and links to the feature aimed at that factor: the daily limit for screen time, the morning and night windows for sleep, Unlock check for pickups, Anti-doomscroll for minutes per pickup, and offline hours for late-night pickups. The ranking is arithmetic on the user's week and makes no prediction about how many points any feature would actually recover.

5.3 Layer 2: Plan

A goal anchored to the user's baseline. Onboarding is a nine-step dialogue that ends with a plan already in force (Figure 7). It asks how the phone makes the user feel and when they usually go to sleep and wake up, and then asks for access to usage data, a step that cannot be skipped since everything after it depends on a measured baseline (P5). The baseline is the average of the last seven complete days. Today's partial total is never used, as a baseline taken just after midnight would be close to zero and would go straight into an enforced rule.

The user then picks a daily screen-time goal on a dial whose range is derived from their own usage, running from 15 minutes up to 75% of the baseline and starting at half of it, so that every plan is a cut of at least 25% and no goal can be an increase in disguise. Once the apps the limit should cover have been chosen, the plan is written as an enforced rule and blocking is live from the moment onboarding ends. The plan screen also shows a projection, worked out by re-running the wellness score on the user's own week with screen time and pickups scaled down by the ratio of limit to baseline; it is arithmetic on the user's data and assumes no effect size.

A plan created in the evening raises a further problem. Usage from earlier the same day, before any plan existed, would otherwise count against the new limit and could trigger a block within minutes of installing the app, which is exactly the kind of unearned restriction P1 is

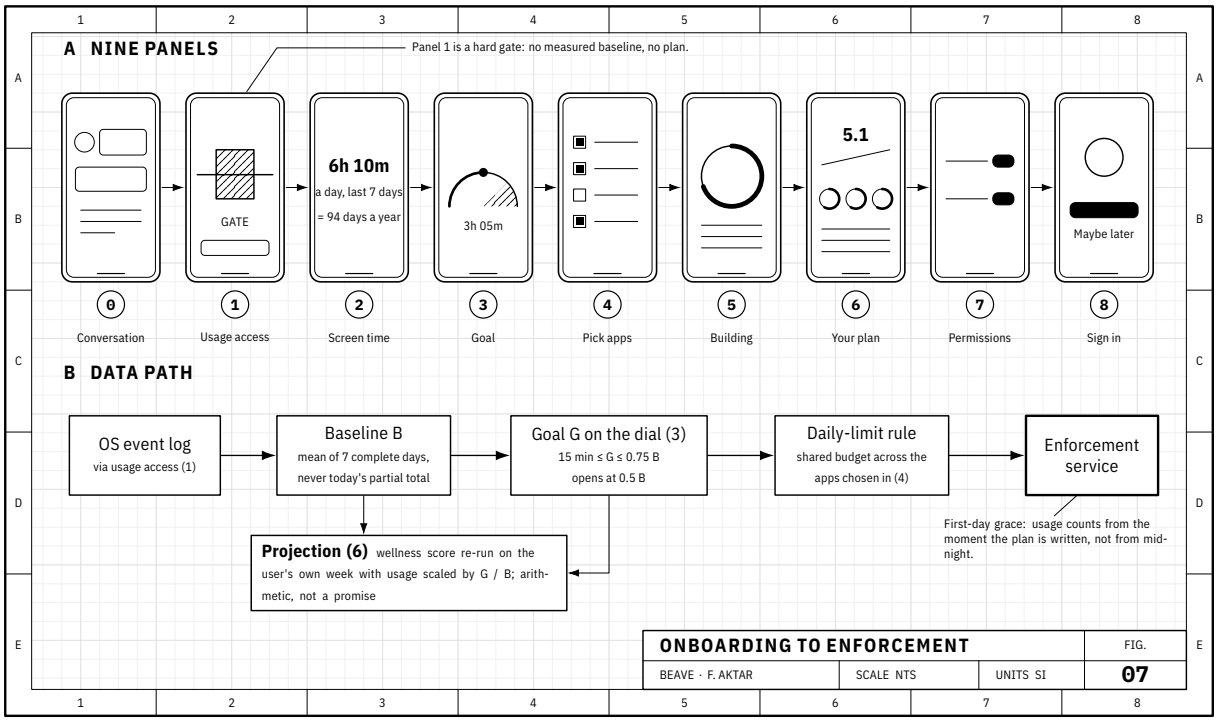


Figure 7. Onboarding, from dialogue to enforcement. Nine panels (top) collect answers and permissions. Below them, the seven-day baseline becomes a bounded goal, the goal and the chosen apps become a daily-limit rule, and the rule reaches the enforcement service before the home screen appears. On the first day, usage counts only from the moment the plan is created.

meant to prevent. On the first day, therefore, usage counts only from the moment the plan is written. This grace is applied once, and later edits to the rule do not reset it, since that would hand out a fresh allowance with every edit.

Rules as implementation intentions. Besides the daily limit, users can write four kinds of rule (Table 4). Each is an implementation intention that the phone carries out on the user's behalf, binding a situation (a time window, a usage threshold, a number of opens, or simply being online) to the response that the chosen apps close. Free users can create one rule in addition to the daily limit, and Gold users as many as they like.

Morning and night windows. Two protected windows frame the day, one after waking and one before bed, during which the chosen apps are closed. They are stored as wake and bed times plus a length (at least one hour, two by default), which is how people tend to think about them, so moving the wake time moves the window along with it. The night window counts back from bedtime, and the apps are already shut by the time the user is trying to fall asleep (Figure 8). Internally both compile down to ordinary schedule rules. The night window responds to the sleep findings under M7 (Exelmans & Van den Bulck, 2016; Chang et al., 2015), and the morning window takes away the first check of the day.

5.4 Layer 3: Pre-commit

Enforcement independent of the interface. Of the commitment devices in Beave, the enforcement service is the one with most consequence. A schedule starting at 09:00 comes into force whether or not the interface has been opened that day, and it survives a reboot. At each evaluation the service consults its sources in a fixed order (Figure 9): an armed Full Lockout first, then the active focus session's block list, then the user's rules, and offline hours last, so that switching

Table 4. Rule types. The enforcement service evaluates every rule twice a second. Windows may cross midnight, and an app covered by several rules is governed by whichever is strictest at that moment.

Type	If (situation)	Then (response), with notes
Schedule	The time lies within $[t_{\text{start}}, t_{\text{end}})$ on a selected day	The selected apps are blocked. Overnight windows are evaluated as 'after the start *or* before the end'.
App limit	The selected apps have together been used for more than L minutes today	All of them are blocked for the rest of the day; they draw on one shared budget.
Open limit	An app has been opened more than N times today, or the current open has lasted M minutes	That app is blocked. Counts are kept per app, so one app cannot use up another's allowance.
Offline hours	Within the window, with a working internet connection	Everything except essentials and up to ten apps the user keeps is blocked until Wi-Fi and mobile data are switched off (Gold).

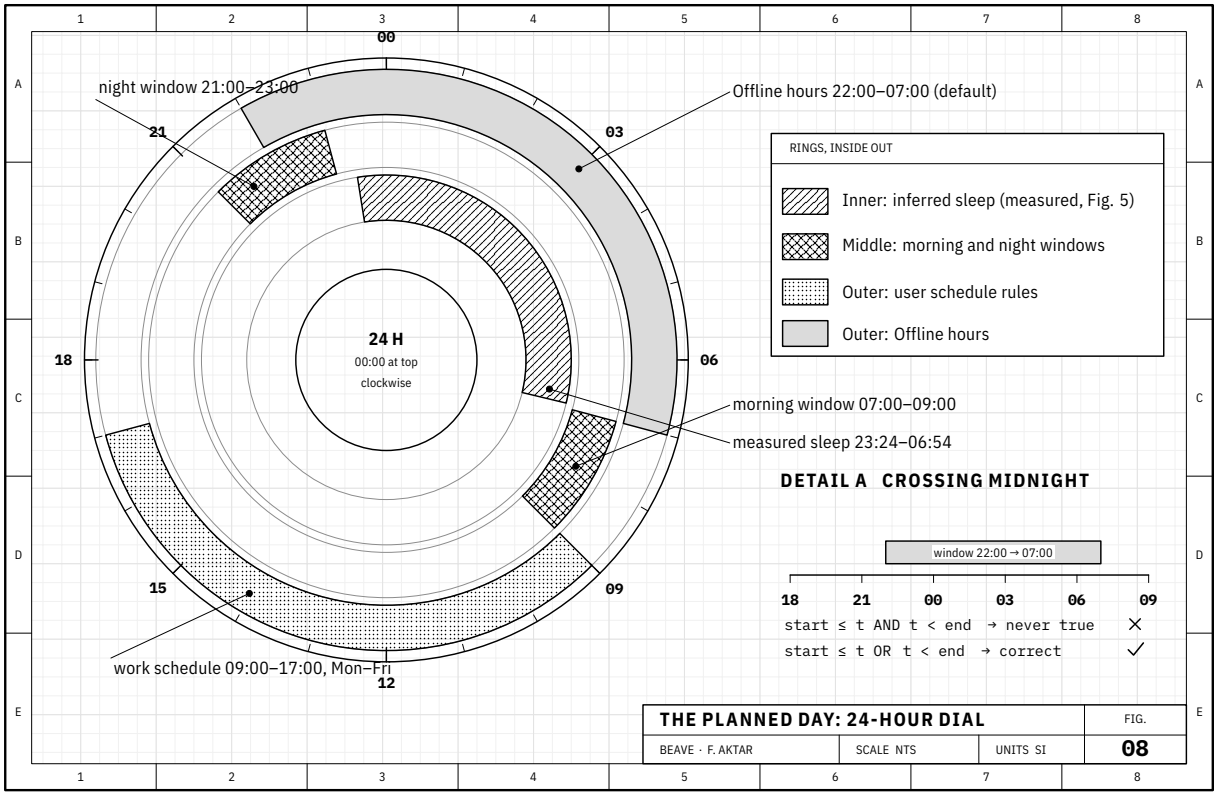


Figure 8. Time-based plans on a 24-hour dial. Inner ring: inferred sleep (from Figure 5). Middle ring: the morning window after waking and the night window before bed. Outer ring: a weekday work schedule and the default offline hours, 22:00 to 07:00. The detail shows why a window that crosses midnight must be evaluated with OR in place of AND.

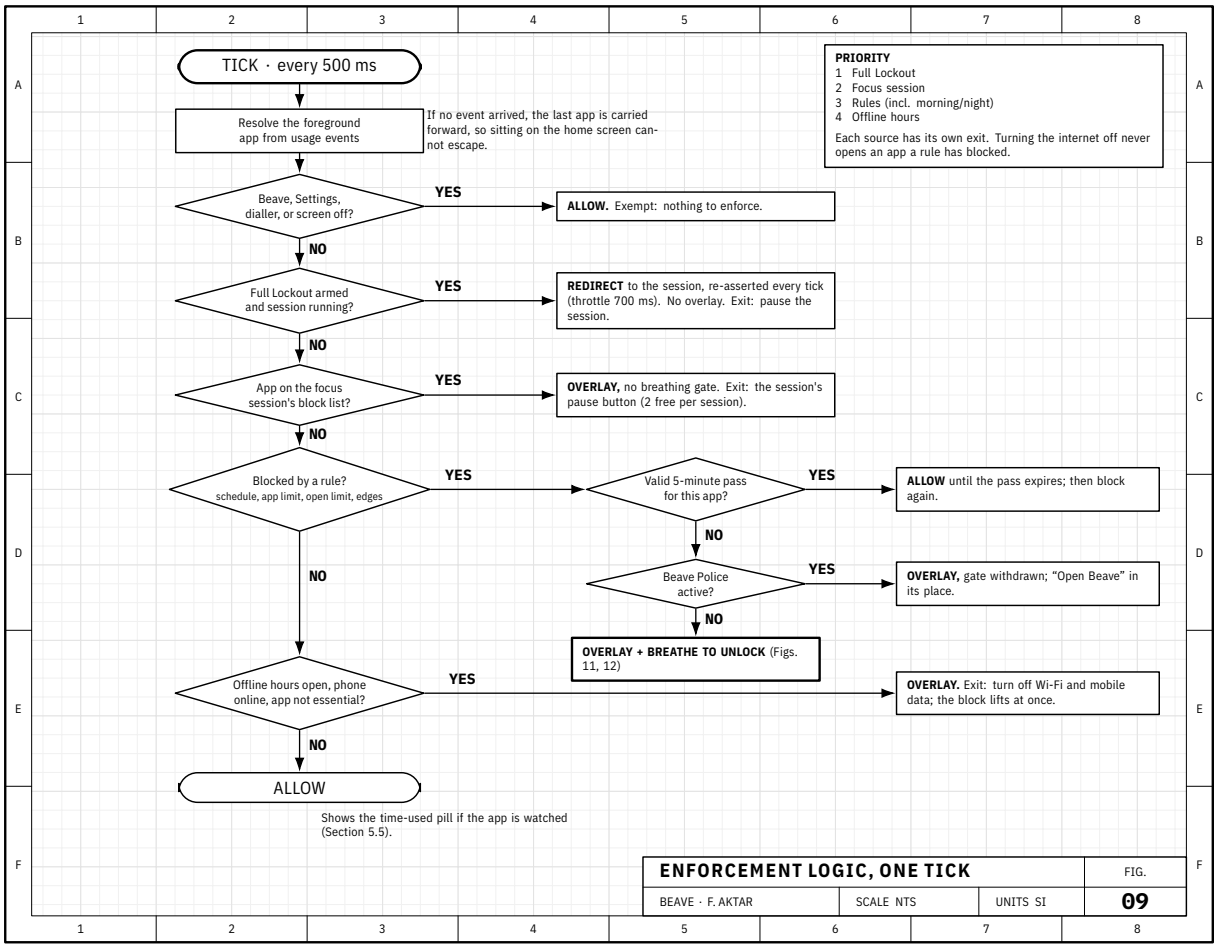


Figure 9. The service's evaluation cycle, repeated every 500 ms. The foreground app is identified and tested against four sources in order, each of which has its own exit: pausing the session lifts Full Lockout and session blocks, the breathing gate opens a rule block for five minutes (unless Beave Police is active), and going offline lifts offline hours.

off the internet can never release an app that a rule has closed. Usage figures are cached for 30 seconds, since the evaluation runs twice a second but the figures themselves only change from minute to minute.

Beave Police: immediate tightening, delayed loosening. Beave Police (Gold) is the most direct implementation of P3 (Figure 10). While it is switched on, any change that would loosen blocking is held for a cooling-off period chosen when Police is turned on, one of one hour, three hours, one day or three days, with one hour as the default. Loosening here covers disabling or deleting a rule, raising a limit, shortening a window and removing an app or a day. A held change is shown as pending together with its release time and is applied automatically when the period runs out, whether Beave is open or not, while changes that tighten blocking apply immediately. Turning Police off is itself treated as loosening and held in the same way, and for as long as Police is on, the breathing gate is withdrawn from rule blocks.

What counts as 'looser' is decided conservatively. A change is held unless it is clearly at least as strict as the rule in force, so a different rule type, fewer apps or days, a window that is narrower at any minute, or a higher limit will all be held. A second edit to the same rule replaces the pending one; it keeps the original release time if it is no looser than the current rule and restarts the clock if it is.

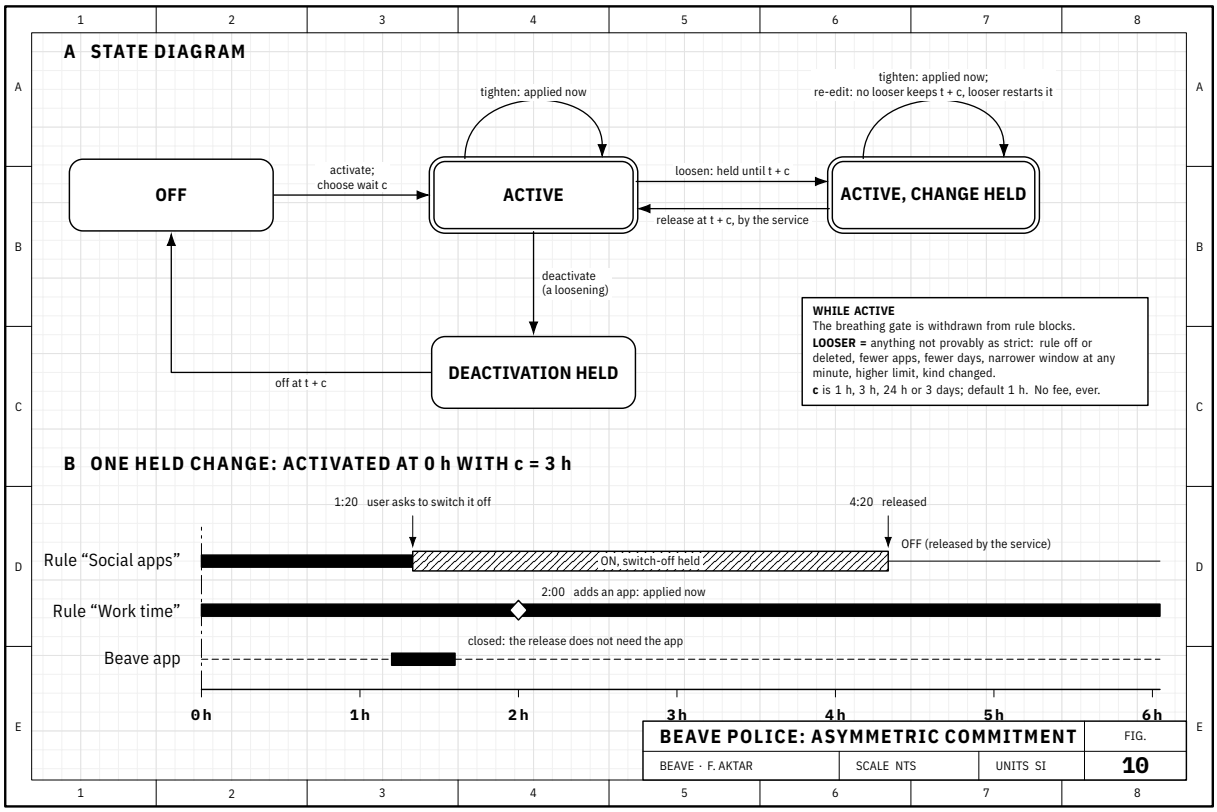


Figure 10. Beave Police. Top: state diagram. Tightening goes through immediately in every state; loosening becomes a hold that is released automatically when its delay runs out, and turning Police off is itself a held change. Bottom: one held change with a three-hour delay. The release is carried out by the background service, so the interface need not be open.

The delay is motivated by what Loewenstein (1996) called the hot–cold empathy gap. Requests to loosen a rule are most likely at precisely the moment the user wants the blocked app, and a delay means the change only happens if it is still wanted once that state has passed. Cooling-off periods of this kind are the standard example of asymmetric paternalism (Camerer et al., 2003), and the next-day rule for changing limits in Phone Dashboard (Allcott et al., 2022) is a close relative. Take-up of commitment devices has generally been low in the economics literature (Bryan et al., 2010), although the 78% of participants who set binding limits in Allcott et al. (2022) suggests that demand for commitment over phone use may be unusually high. Police is nevertheless opt-in, its shortest delay is short, and no delay can be cut short by paying (P4).

Full Lockout and offline hours. Two further commitments apply to the phone as a whole. *Full Lockout* (Gold) is armed at the start of a focus session, and until the session ends, opening anything other than Beave, the dialler or Settings (the home screen included) sends the user back to the session within about a second; the notification shade stays reachable as a safety valve. Once armed it cannot be disarmed mid-session. Pausing the session does lift it, however, because a pause button that stopped the timer but went on refusing every app would contradict its own label. Free users get two pauses per session.

Offline hours (Gold) define a window, 22:00 to 07:00 by default, during which the phone can only be used offline. Beave does not cut the connection itself; if the phone is online during the window, every app apart from a short list of essentials (dialler, messages, alarms, Settings and the launcher) and up to ten apps the user chooses to keep, such as a course app that is needed

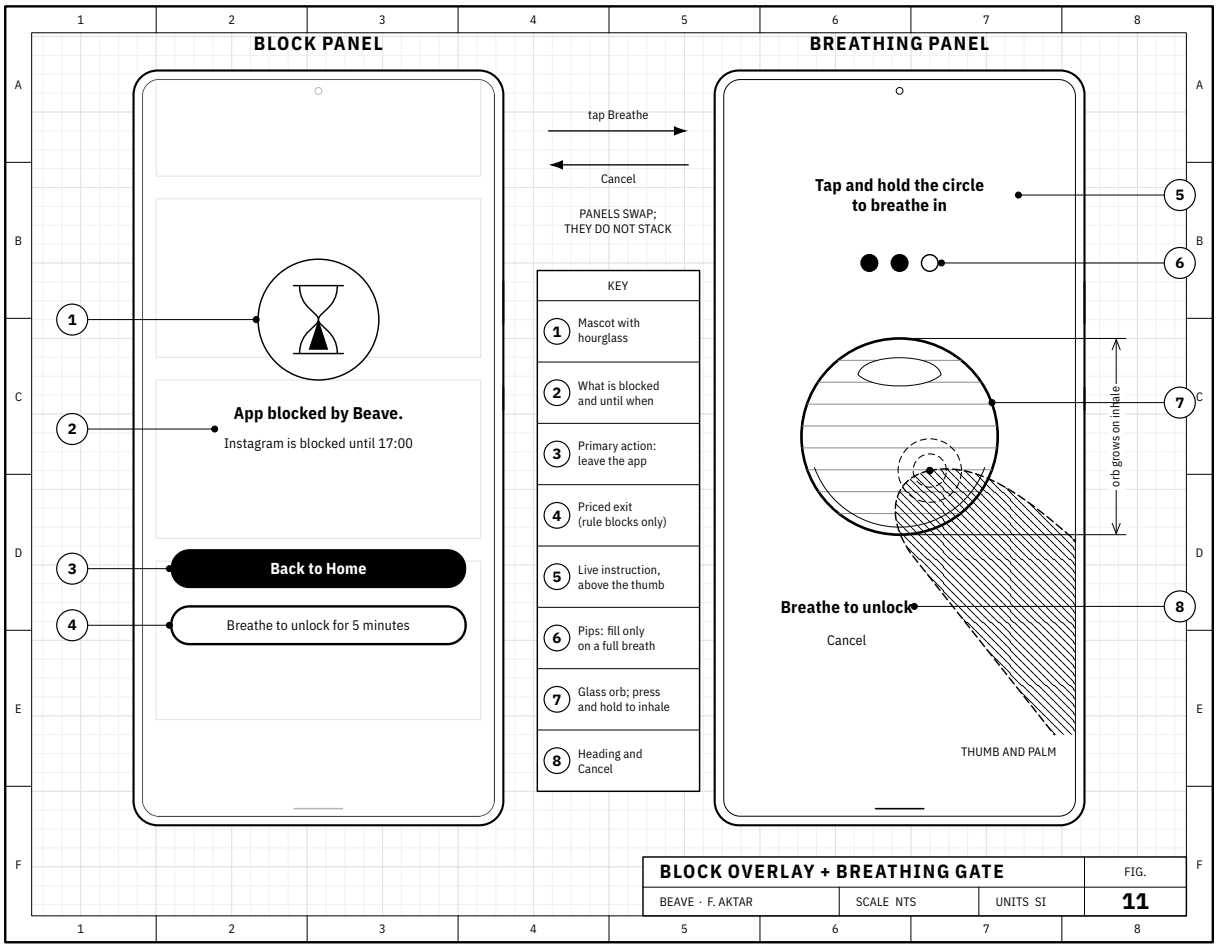


Figure 11. The block overlay and breathing panel, with principal dimensions. Left: the block panel, with mascot, message, primary action (back to the home screen) and the effortful exit. Right: the breathing panel. The instruction and the three progress pips sit above the orb because the thumb and palm hide whatever lies below it. The two panels replace one another and are never stacked.

online, is blocked, and the block lifts as soon as Wi-Fi and mobile data are off. Under Police, adding an app to that list counts as loosening. The user thus makes a single decision in advance, to accept a phone without internet at night, in place of a fresh decision to resist it every evening. Time spent offline inside the window is added up and shown on the Progress tab as a lifetime total, which, like the dam (Section 5.6), grows only with time away from the phone.

5.5 Layer 4: Interrupt

The block overlay and the breathing gate. When a blocked app comes to the foreground, the service covers it with a full-screen overlay (Figure 11). Outside focus sessions this offers two actions, going back to the home screen or *Breathe to unlock for 5 minutes*, and the second opens a breathing panel in which the user presses and holds an orb to breathe in. After four seconds a cue tells them to let go, and the orb shrinks over a six-second breath out; three completed cycles unlock the app for five minutes, after which it is blocked again (Figure 12).

The gate is built so that it cannot be rushed. A hold of under three seconds is rejected with a prompt to hold a little longer, a new breath cannot start until the previous one has been fully let out, and a pip only fills once a cycle is complete, so a pass costs at least $3 \times (3 + 6) = 27$ seconds, or about 30 for someone who follows the cue. The button reads ‘Breathe to unlock’ and never just ‘Unlock’, since an exit that looks free soon stops working as friction.

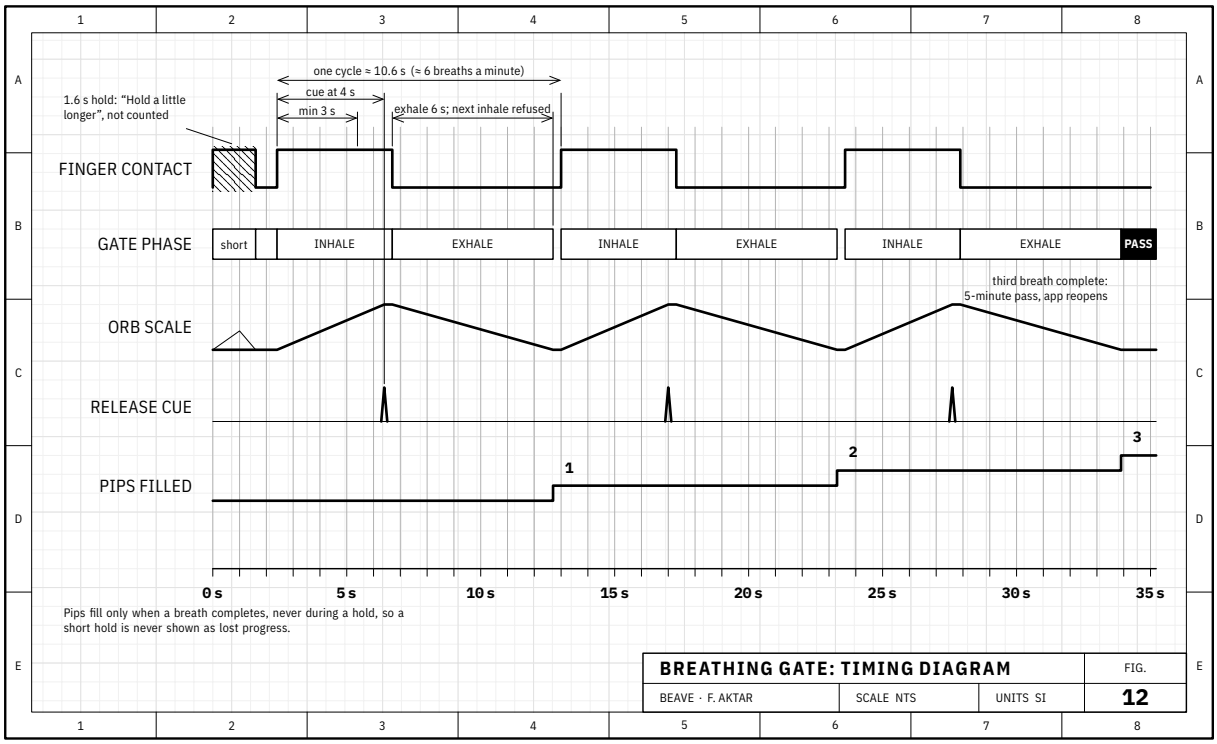


Figure 12. Timing of one successful pass through the breathing gate. Rows, from the top: finger contact, gate phase, orb size, release cue and progress pips. A hold shorter than the 3 s minimum (shaded) is rejected, the release cue comes at 4 s, and a new breath cannot start before the 6 s exhale has ended. A pass costs at least 27 s; one cycle lasts about 10 s, or six breaths a minute.

Three strands of evidence went into the gate. The first concerns friction at the moment an app is launched. In a six-week field experiment with 280 participants, *one sec*, which inserts a short delay and an option to cancel whenever a target app is opened, led people to abandon 36% of their attempts to open those apps and to make 37% fewer attempts overall; by the sixth week actual openings were down by 57%, and the option to cancel turned out to be the most effective component (Grüning et al., 2023).[3] Effortful lockout tasks before an app opens (Kim et al., 2019) and vibrations that start once a self-set limit has been passed (Okeke et al., 2018) have also reduced use.

The second is that a pause gives the reflective system a chance to step in (M4). College smokers who were given brief instruction in watching a craving until it passed, so-called 'urge surfing', smoked fewer cigarettes over the following week than those who were not (Bowen & Marlatt, 2009).

The third is that slow breathing with a long exhale has effects of its own. It is associated with more parasympathetic activity and lower arousal (Zaccaro et al., 2018), a rate of around six breaths a minute matches the resonance frequency used in heart-rate-variability biofeedback (Lehrer & Gevirtz, 2014), and in a randomised comparison, exhale-focused breathing improved mood and lowered breathing rate more than mindfulness meditation did (Balban et al., 2023). Four seconds in and six out comes to exactly six breaths a minute. Balban and colleagues were testing five minutes of daily practice over a month, however, which is a long way from thirty seconds spent in front of a blocked app, so I take their results as a justification for the shape of the pause and not as evidence that it changes behaviour.

[3]The developer of *one sec* is one of the study's authors, a position not unlike that of the present author (see Declarations).

The gate is absent, on purpose, in three situations. Inside a focus session the only exit is the session's own pause button, so that the gate cannot serve as an uncounted way round the free tier's two pauses or water down a verified session; when Beave Police is on, the gate is withdrawn from rule blocks because the user has asked for stricter enforcement; and during offline hours the way through is simply to go offline.

Anti-doomscroll: a stopping point in endless feeds. Blocking decides whether an app opens but does nothing about how long a session lasts once it has. *Anti-doomscroll* (Gold, in beta) measures scrolling in the apps a user chooses and interrupts long sittings (Figure 13). It works from Android accessibility events, which report that a scroll took place and how far it went, and it is set up so that the content of the screen is not read; scroll data stay on the phone and are kept out of analytics entirely.

A *sitting* adds up active scrolling time, with each gap between scrolls counting for at most 60 seconds, so that reading a long post or watching a video counts while putting the phone down does not. Leaving the app or locking the screen pauses the sitting, coming back within ten minutes of the last scroll resumes it, and ten minutes without scrolling starts a new one. These two values decide, to a large extent, whether the feature ever fires at all. If a sitting were reset on every exit, a one-second visit to the home screen would wipe it, and with a short gap cap (30 seconds, for instance) someone who watched each short video to the end would hardly ever reach the limit, leaving only the fastest scrollers to be interrupted.

When a sitting reaches the user's limit (15, 30, 45 or 60 minutes, 30 by default), the next scroll brings up a break panel that stays for ten minutes. Its only button leaves the app, since a button that dismissed the panel and left the user on the feed would make the break a formality. Scrolling in a watched app during those ten minutes brings the same panel straight back, but its clock is not restarted by each return, which would turn the break into a penalty that grows without the user being able to see why; and since ten minutes without scrolling ends a sitting, the next one starts from zero once the break is over.

This is the one block outside a focus session that has no priced way through. Earlier versions offered a five-minute snooze through the breathing gate and kept the fixed break as an opt-in hard mode; the break is now the only mode. It stays within the spirit of P1 in three respects: the feature is opt-in and the user chooses its apps and its limit, the break ends by itself after ten minutes, and it closes a feed and not the phone, since every other app remains usable.

The design follows Rixen et al. (2023) in assuming that people mostly stop scrolling for reasons that lie outside the app, and the break supplies the interruption that the feed was designed to leave out. A reminder that cannot be interacted with is easily ignored, and a panel with a costless 'keep scrolling' button turns the break into one more thing to dismiss, which is why the panel has a single button and that button leaves the feed.

Unlock check: a stated purpose at each unlock. The interventions described so far act once an app has been chosen. *Unlock check* (Gold, opt-in) acts a step earlier, at the unlock itself, which is where a checking habit begins (M1; Oulasvirta et al., 2012). After each unlock a panel over the home screen asks what the user is there for and offers the apps they used most over the past week as one-tap answers, together with 'Something else'; choosing an app opens it. The apps the user is trying to cut down on, meaning any app named in one of their rules, are left out of the choices on purpose. A useful app is therefore one tap away and a distracting one can only be reached through 'Something else', and it is this asymmetry, not a delay, that supplies the friction.

If the user then moves into an app they are cutting down on, a single nudge names what

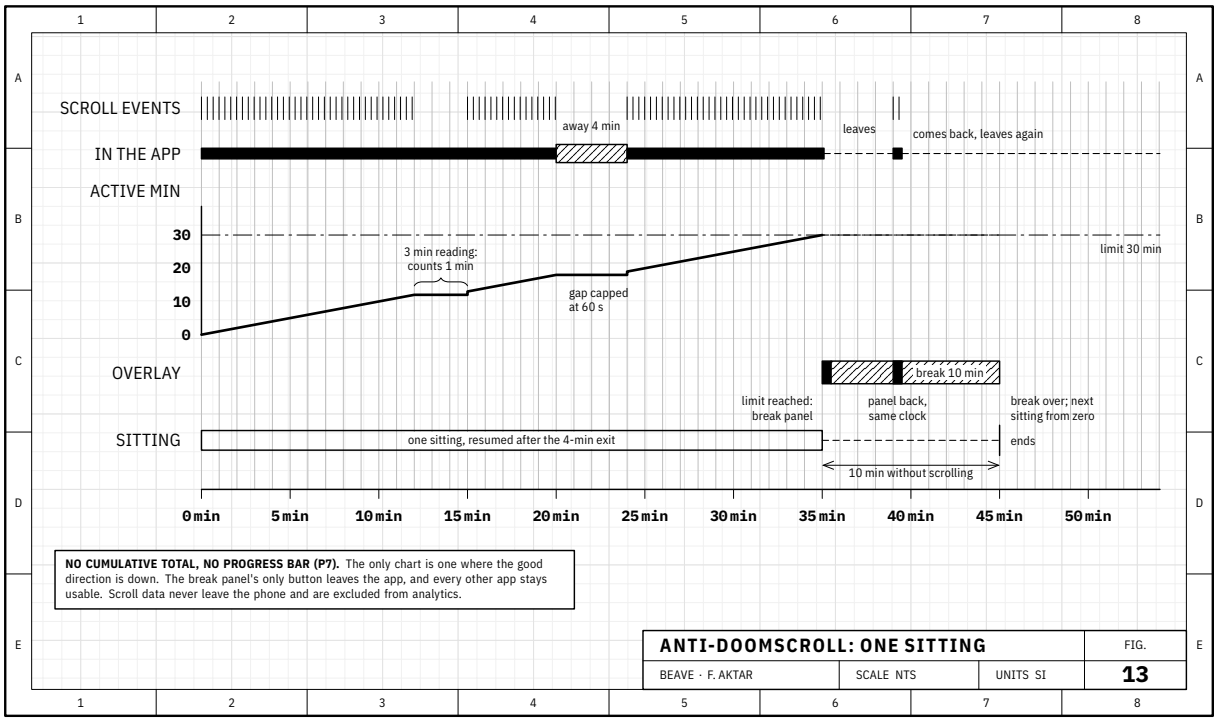


Figure 13. One scrolling sitting under a 30-minute limit. Scroll events (ticks) add up active time, each gap counting for at most 60 s, and an absence from the app of under ten minutes continues the same sitting. At the limit, the next scroll brings up a ten-minute break panel whose only button leaves the app. Coming back during the break brings the panel back without restarting its clock, and since ten minutes without scrolling ends a sitting, the break also closes it. There is no running total and no progress bar (P7).

they came for ('You came for WhatsApp') and offers a way back or 'Keep going'. Keeping going accepts that app for the rest of the unlock, so the same app is never questioned twice, while the stated purpose remains in force for any further drift. An optional check-in, after 5, 10 or 15 minutes (10 by default), reminds the user how long ago they stated their purpose, which covers the case the nudge cannot, namely too long spent in the app they came for (M6).

The answer is a goal stated at the moment of the cue, and the nudge holds the user to it, borrowing from implementation intentions the idea of tying a response to a specific situation (Gollwitzer, 1999). In terms of the cycle described by Tran et al. (2019), the question meets the trigger and the check-in breaks into the stretch of time that would otherwise go unnoticed. Users rate habitual checking as the least meaningful time they spend on their phones (Lukoff et al., 2018), and the design lets use with a stated purpose through in a single tap. Because the panel appears dozens of times a day it is kept cheap: there is no pause before the choices, and the question rotates through five wordings so that it is not answered by rote. It is not shown during a call, a focus session or Full Lockout, while offline hours are in force, when the unlock came from tapping a notification (the notification is then the reason), or within three minutes of the screen going off if a reason has already been given. Under Police, switching Unlock check off, narrowing its hours or lengthening the check-in counts as loosening. The apps a user names stay on the phone, and analytics record only counts, never app names. The design itself has not been tested, and I offer it as an application of the evidence on implementation intentions and not as a finding in its own right.

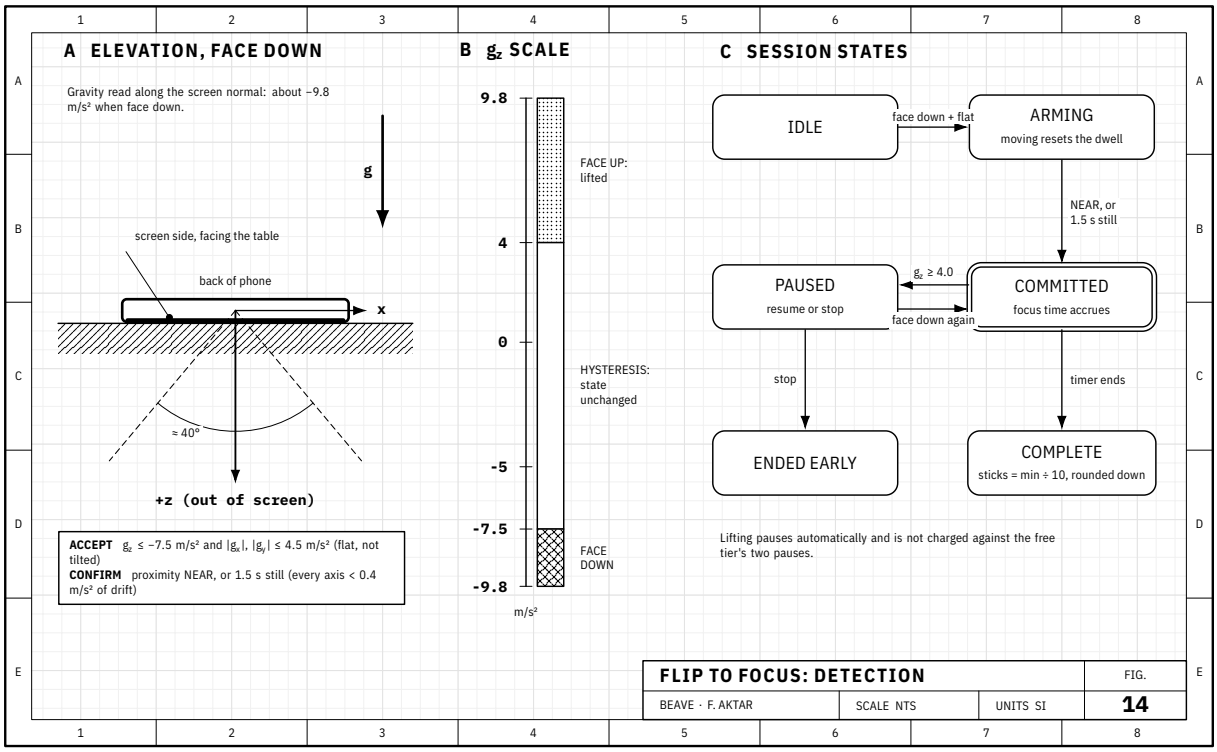


Figure 14. Flip detection. Left: the device axes and the gravity vector in the face-down position, with the accepted angular range. Centre: thresholds along the z axis, including the dead band between face-down and face-up. Right: session states; lifting the phone pauses the session and laying it down resumes it. Focus time accrues only in the COMMITTED state.

Lighter prompts. Three lighter interventions keep usage in view without blocking anything. A *time-used pill* appears in a corner of the screen for four seconds when a watched app is opened, and again every 30 minutes, showing the time spent in that app today, and touches pass straight through it. It is aimed at time distortion (M6) as it happens and relies on the reactivity of self-monitoring, the long-standing observation that keeping track of a behaviour tends to change it (Korotitsch & Nelson-Gray, 1999). A *limit notification* arrives at 80% of the daily limit, early enough to do something about it, and a *recap* at 21:00 sums up the day; these two fire at most once a day each.

5.6 Layer 5: Reinforce

Flip to focus: sessions verified by the device. A focus session can be started with a button or by laying the phone face down (Figure 14). Orientation comes from the accelerometer: gravity has to register through the back of the phone at 7.5 m/s^2 or more, with the horizontal components below 4.5 m/s^2, which separates a phone lying flat from one that is merely tilted. If there is a proximity sensor it confirms the placement straight away. If not, the phone has to stay face down and still for 1.5 s, moving by less than 0.4 m/s^2 on every axis, a threshold chosen to lie between the jitter of a phone on a table and the tremor of one held in the hand. Lifting the phone (gravity above 4.0 m/s^2 through the screen) pauses the session and putting it back down resumes it,[4] and these automatic pauses do not count against the free tier's allowance.

Flipping does two things. It takes the phone out of sight and out of easy reach, which is about the strongest form of situation modification an app can bring about (M7; Duckworth

[4]The gap between the two thresholds acts as hysteresis. With a single threshold, small rotations of a phone lying close to it would toggle the session on and off.

et al., 2016); the evidence for a cognitive cost of the phone's mere presence is mixed (Ward et al., 2017; Ruiz Pardo & Minda, 2022), but the cost of incoming notifications is better established (Stothart et al., 2015). The second, and more important, is that session length is measured by the sensors without any input from the user, and it is this that lets a focus session carry some weight with other people (P6).

Sticks, dams and rivers. Each completed session earns one stick per ten minutes of focus, and the sticks go into a dam on a real river in one of ten regions drawn from map data. The rate is linear so that splitting sessions up brings no advantage; under a schedule that paid short sessions disproportionately, twelve five-minute sessions might outscore one unbroken hour, which would reward exactly the fragmented attention the app is supposed to reduce. Filling every river in every region takes 1,705 sticks, or roughly 284 hours of verified focus. Session presets run from 10 minutes to 3 hours, the longest worth 18 sticks.

The dam and the offline-hours total (Section 5.4) are the only totals in Beave that only ever go up, and both count time away from the phone, in line with P7. Effort is known to increase as a visible goal comes closer (Kivetz et al., 2006), and a review of 24 empirical studies of gamification found mostly positive effects that depended heavily on the context and on the users involved (Hamari et al., 2014). Forest's virtual tree, mentioned in Section 3, is perhaps the best-known use of the idea in this area, and Lyngs et al. (2019) point out that it asks the user to refrain from doing something, where a virtual pet would normally ask them to do something in order to feed it.

Pomodoro runs. A focus session can also be split into a Pomodoro run of two to six rounds, either 25 minutes with 5-minute breaks or 50 minutes with 10-minute breaks. Each round is an ordinary session, verified, enforced and credited like any other, so flip detection, Full Lockout and the session's block list all apply and every completed round counts as a session. Sticks are credited on the run's total focus time and not round by round, which keeps the rate linear: four 25-minute rounds earn the same 10 sticks as 100 minutes in one sitting, where crediting each round separately would earn eight. Breaks lift every block. In a run started by flipping, the sensor stays armed through the break, so laying the phone face down again starts the next round, and giving up part of the way through keeps the rounds already finished.

5.7 Social features

Self-reported effort carries little weight with other people because it costs nothing to claim (anyone can say they studied for three hours). Activity trackers changed this for exercise, since a run recorded by a watch does not depend on trusting the runner, and flip verification gives focus time a similar standing: the phone measured the session, so an hour of focus becomes a record that others can rely on. Beave's social features are built on that property.

There is good evidence that other people affect how much effort we put in. Using five years of data on about 1.1 million runners, with variation in the weather where their friends lived serving as a natural experiment, Aral & Nicolaides (2017) found running to be socially contagious; when a runner's friends ran more on a given day, so did the runner, and less active runners influenced more active ones but not the other way round. In the STEP UP trial, 602 adults were randomised to a control arm or to one of three gamified arms built on support, collaboration or competition (Patel et al., 2019). All three raised daily steps above control during the intervention, by between 637 and 920 steps a day, but the competition arm, whose members received a weekly leaderboard ranking them against two other participants, gained the most and was the only arm still ahead of control during follow-up, by 569 steps a day. Leaderboards seem to work largely as a form of

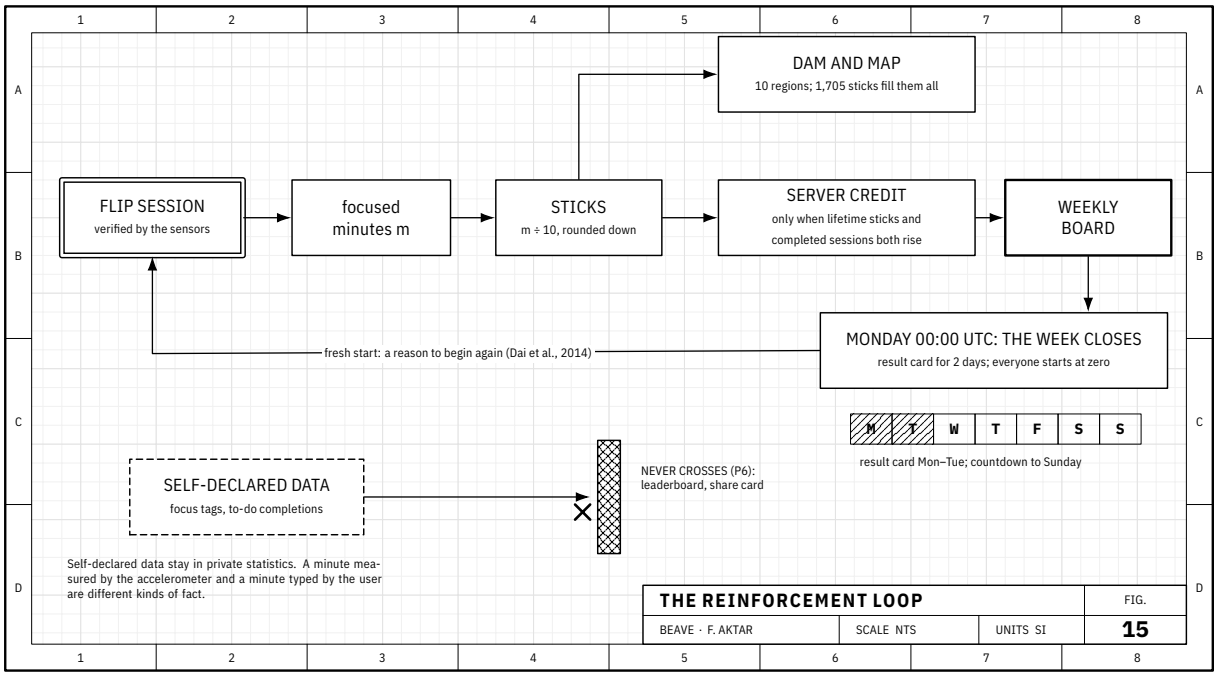


Figure 15. The reinforcement loop. A verified session earns sticks at a fixed rate. The sticks fill rivers on the map and, through a server-verified credit, count towards the weekly leaderboard. The week ends at 00:00 UTC on Monday, when all weekly counts start again. Self-declared data such as tags and tasks are kept away from every surface other people can see.

goal setting (Landers et al., 2017), progress shared with others has larger effects than progress kept private (Harkin et al., 2016), and comparison is most informative when it is with people similar to oneself (Festinger, 1954).

The same literature also shows how leaderboards go wrong, and this shaped Beave's version more than anything else (Figure 16). Competitive motivation falls as the number of competitors rises (Garcia & Tor, 2009), and it is concentrated near meaningful standards such as the top few ranks, so that the difference between 202nd and 203rd place matters very little to anyone (Garcia et al., 2013). A cumulative global ranking comes close to the least motivating arrangement there is, since a newcomer enters thousands of places down, behind people with months of sticks, with no realistic chance of climbing.

The leaderboard in Beave therefore opens on *This week* on every visit, with the all-time ranking one tap away. Weeks end at 00:00 UTC on Monday, one shared week for everybody, and temporal landmarks of this kind are known to prompt people to make a fresh start (Dai et al., 2014). The interface talks about the week finishing and never about a 'reset'. Its heading counts down to the end of the week, and for two days afterwards a card shows the user's final rank and the sticks they added to their dam. Users outside the top ten see the two entries just above them and the one just below, so they are compared with people they could realistically catch. Several things were left out deliberately, including any count of the sticks needed to overtake the next person, a follow button and a friend list, and only verified effort ever reaches the board.

A finished session can also produce a *share card*, a square image taken just after the session showing the user's handle, how long they focused, the sticks they earned, the region those went to and the date. Badges mark milestones of verified effort: 10, 100 and 1,000 sticks; 10, 100 and 1,000 flips; 100 hours of focus; five Pomodoro runs with every round finished; and 100 hours spent offline inside offline hours. They mark progress and are not a second currency, resting on

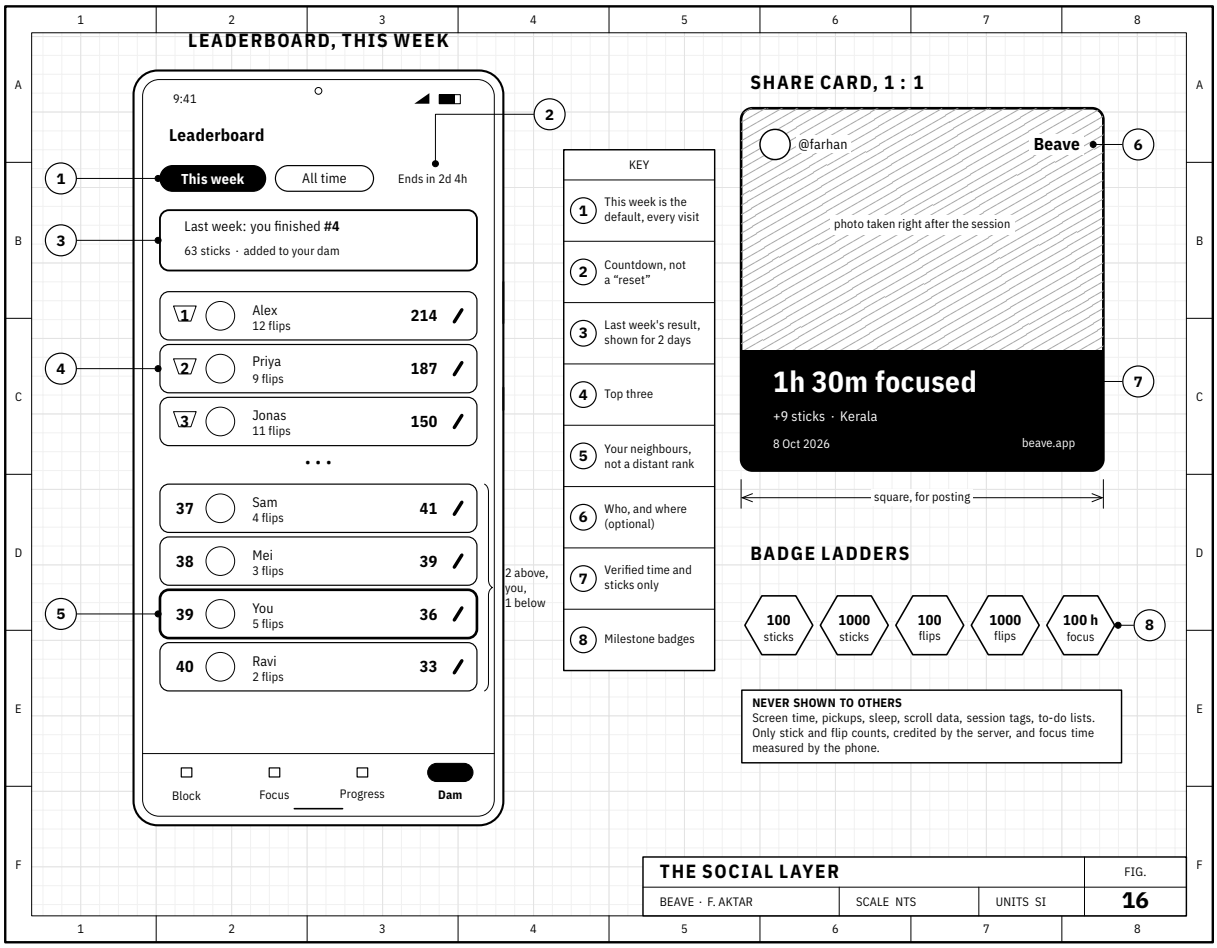


Figure 16. Social features. Left: the weekly leaderboard on the Dam tab. It opens on This week (1), counts down to the end of the week (2), shows the previous week's result for two days (3), and below the top three (4) shows the user's nearest neighbours in place of a distant rank (5). Right: a square share card carrying the user's handle (6) and verified figures only (7), and the badge ladders (8). Usage, sleep and scroll data appear on none of these surfaces.

the same idea as the leaderboard, that a visible record of verified effort is itself worth working for.

A place on the board is only worth having if the record behind it can be trusted, and the board is protected accordingly. Weekly stick and flip counts are written only by a function on the server, and only when the user's lifetime stick total and number of completed sessions go up together, and the phone cannot write to those fields at all. Self-declared information such as session tags ('study', 'work') or to-do items never appears on the board or on a share card, even though including it would probably increase engagement, because a minute confirmed by the accelerometer and a typed claim to have been studying are different kinds of evidence (P6). No usage, sleep or scroll data appear anywhere social.

5.8 Privacy and ethics

Digital wellbeing tools are persuasive technologies too, and raise the same ethical questions as the products they are meant to counter (Peters et al., 2018; Monge Roffarello & De Russis, 2019). Beave keeps to five constraints. (i) Detailed usage, sleep and scroll data, and the apps named at Unlock check, are stored only on the phone; for the same reason the wellness benchmark is built from published averages instead of other users' data, and the scroll-measuring feature is excluded

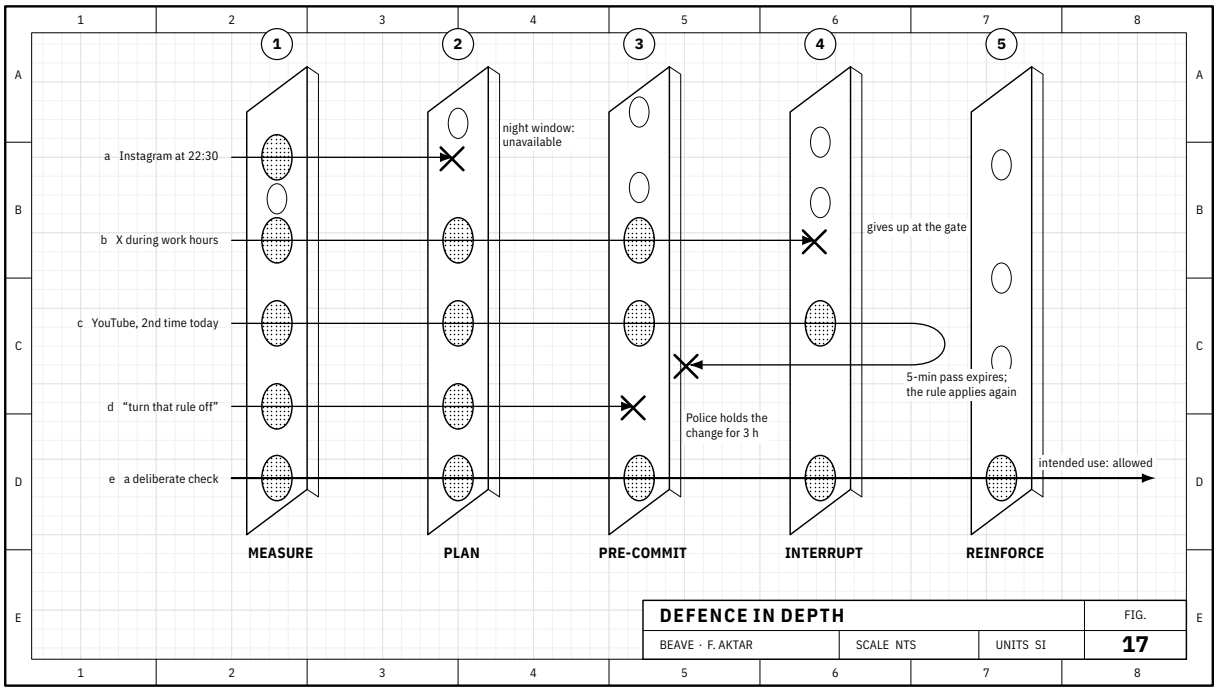


Figure 17. Defence in depth. Each layer is a barrier with gaps in it. Five impulses are traced from left to right; most are stopped or turned aside early, one gets through the breathing gate and is caught again when its five-minute pass expires, and one passes every layer, which is the intended outcome when the user really does want the app. The aim is to regulate the flow, and zero is not the target.

from analytics altogether. (ii) No way out of a block requires payment (P4). (iii) Features that can hold a block (Police, Full Lockout and Anti-doomscroll's break) are opt-in and explained before they are switched on. (iv) If a paid feature lapses, whatever it was enforcing stops, so that no off switch is ever left behind a paywall. (v) The projections shown during onboarding are calculated from the user's measured week and contain no invented percentages.

6 Discussion

6.1 The case for layering

Safety engineering offers a useful way of thinking about this. In Reason's (2000) 'Swiss cheese' model of accidents, every defensive layer has holes, and harm only gets through when the holes in successive layers happen to line up. Much the same holds for self-regulation (Figure 17). A plan with nothing to enforce it is dropped at the first hot moment, enforcement with no plan behind it seems arbitrary and invites a fight, and an exit that costs something but has no commitment behind it is little more than a speed bump; and all of them rely on the measurement underneath being right. An impulse that gets past one layer, for instance by completing the breathing gate, runs into the next: the pass expires after five minutes, the daily limit keeps counting, Police holds any attempt to loosen the limit, and the evening recap reports what happened.

6.2 The claim in the title

No single feature, app or habit resolves problematic smartphone use, and the evidence does not support anyone who says it does. What the literature does provide is a fairly clear account of the mechanisms that keep compulsive use going and, for most of them, independent experimental evidence that a particular counter-measure helps. The approach I propose lies in the combination of accurate measurement, advance planning in if-then form, commitment that resists being

undone on impulse, a modest cost between urge and app, and a visible place for the time won back to go. My argument is that each of these is necessary and none is sufficient on its own. Beave is one way of building the combination, and there are certainly others.

6.3 Limitations

No outcome data. Beave has not been evaluated in a controlled trial. Every claim of efficacy in this paper refers to a mechanism demonstrated elsewhere, with different tools and different populations, and the interaction between layers, on which the central argument of the framework depends, has not been tested.

Contested evidence. Some of the findings cited here are disputed or have failed to replicate, among them the cognitive cost of a phone's mere presence (Ruiz Pardo & Minda, 2022) and the average effect of nudges once publication bias is accounted for (Maier et al., 2022). I have tried to rely on such findings only where other evidence points the same way, and to say so where it does not.

Platform constraints. Enforcement depends on capabilities Android provides (usage statistics, overlay windows and persistent background services) that iOS does not offer in an equivalent form. It also depends on phone manufacturers not killing background services, which some of them do aggressively, hence the periodic heartbeat on which the service restarts itself. Scroll measurement depends on what each app reports and therefore varies from app to app and may change when they are updated.

Circumvention. Anything installed on a user's own phone can be uninstalled. Beave's response is to reduce the motivation to remove it (P1, P3) and not to make removal hard, and it does not use device-administrator privileges. Whether that trade-off holds up is an empirical question.

Self-selection. People who install an app of this kind are already motivated to change, which limits how far any results generalise and flatters uncontrolled before-and-after comparisons, one more reason for the randomised design below.

7 Proposed evaluation

The evaluation is set out in enough detail to be pre-registered. It is designed to estimate both the effect of the framework as a whole and the separate contribution of its in-the-moment components (Figure 18), borrowing from Allcott et al. (2022) the use of device-logged outcomes and an app that records use in every arm.

Design. There are two parts. *Part A* is an eight-week randomised comparison in which new users are allocated either to the full framework or to a monitoring-only condition that keeps measurement and the Progress tab but enforces no rules and delivers no interruptions. A four-week follow-up then tests whether any change persists, as it did after deactivation in Allcott et al. (2020) and as one would expect if new habits had formed (Lally et al., 2010). *Part B* is a micro-randomised trial nested in the full-framework arm (Klasnja et al., 2015): at each block, the breathing gate is randomly either shown or replaced by a plain 'Open anyway' button, and each time 80% of the limit is crossed, the notification is randomly sent or withheld. This estimates the proximal effect of each in-the-moment component (Nahum-Shani et al., 2018) with far fewer participants than a full factorial trial would need.[5]

[5]I leave the sample size open until a pilot provides variance estimates for logged screen time, which differs widely between users.

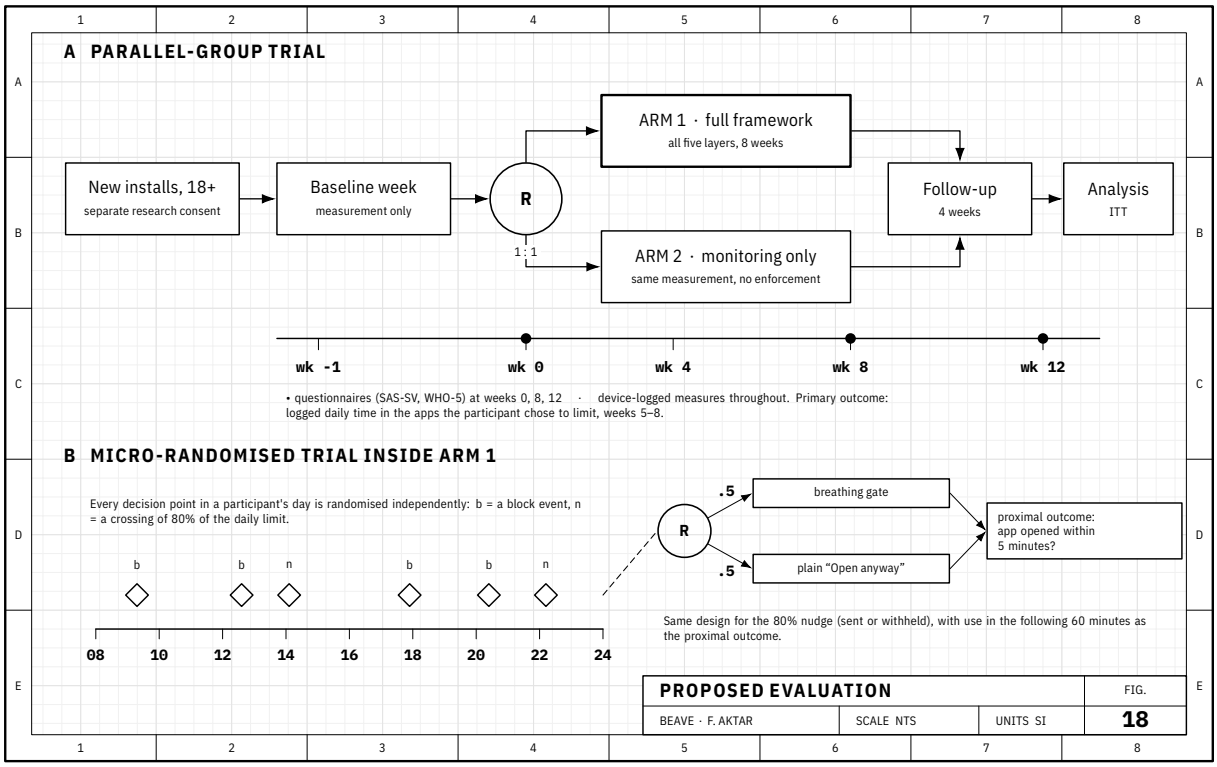


Figure 18. The proposed evaluation. Part A (top): eight weeks comparing the full framework with monitoring only, followed by four weeks of follow-up. Part B (bottom): within the full-framework arm, a micro-randomised trial of two in-the-moment components at each point where they could apply. All measures are logged by the device except two validated questionnaires.

Outcomes. The primary outcome is device-logged daily screen time in the apps each participant chose to limit, averaged over weeks 5 to 8. Secondary outcomes are total screen time, pickups, inferred sleep and late-night pickups, all logged by the device to avoid the self-report bias discussed under M6, together with the short form of the Smartphone Addiction Scale (Kwon et al., 2013) and the WHO-5 Well-Being Index (Topp et al., 2015) at baseline, week 8 and week 12. For Part B, the proximal outcomes are whether the blocked app is opened within five minutes and how much the phone is used in the hour after a notification.

Analysis and ethics. The primary analysis is by intention to treat, with mixed-effects models, and Part B uses weighted and centred least squares, the standard estimator for micro-randomised trials. Taking part would require explicit consent separate from Beave's ordinary privacy terms, because the study has to upload measurements the app would otherwise keep on the phone. The protocol would be pre-registered and the results published whichever way they come out.

8 Conclusion

People who struggle with their phones are rarely short of information about how much they use them; what they lack is structure at the moments that matter. A good deal is now known about what sustains compulsive use, and most of the mechanisms involved have a known counter, whether accurate measurement for time distortion, if-then plans for cue-driven habits, commitment devices for present bias, a priced pause for impulses, a stopping point in endless feeds or a verified reward to compete with the feed itself. Used one at a time, none of these is likely to achieve much. Combined in layers, held to principles that protect autonomy and privacy, and

enforced by software that does not depend on willpower at the moment willpower is weakest, they amount to a credible way of regulating phone use without banning it. Beave is my attempt to build that combination; whether it does what it is meant to do is an empirical question, and Section 7 describes how I think it should be answered.

Declarations

Competing interests. The author develops Beave and receives income from its paid tier. Readers should weigh the paper with this in mind; it is also the reason no claims of efficacy are made here and an independent, pre-registered evaluation is proposed in their place.

Medical disclaimer. Beave is a digital wellbeing tool. It is not a medical device and is not intended to diagnose or treat any condition. Anyone experiencing significant distress related to phone use should consider seeking professional help.

Data. No user data were analysed for this paper. Parameter values are taken from Beave's source code.

Version. This paper describes Beave 3.8.5 for Android as of October 2026. Details may change in later versions.

References

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82*(4), 463–496. https://doi.org/10.1037/h0076860

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, *110*(3), 629–676. https://doi.org/10.1257/aer.20190658

Allcott, H., Gentzkow, M., & Song, L. (2022). Digital addiction. *American Economic Review*, *112*(7), 2424–2463. https://doi.org/10.1257/aer.20210867

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond self-report: Tools to compare estimated and real-world smartphone use. *PLOS ONE*, *10*(10), e0139004. https://doi.org/10.1371/journal.pone.0139004

Aral, S., & Nicolaides, C. (2017). Exercise contagion in a global social network. *Nature Communications*, *8*, 14753. https://doi.org/10.1038/ncomms14753

Balban, M. Y., Neri, E., Kogon, M. M., Weed, L., Nouriani, B., Jo, B., Holl, G., Zeitzer, J. M., Spiegel, D., & Huberman, A. D. (2023). Brief structured respiration practices enhance mood and reduce physiological arousal. *Cell Reports Medicine*, *4*(1), 100895. https://doi.org/10.1016/j.xcrm.2022.100895

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Billieux, J., Maurage, P., Lopez-Fernandez, O., Kuss, D. J., & Griffiths, M. D. (2015). Can disordered mobile phone use be considered a behavioral addiction? An update on current evidence and a comprehensive model for future research. *Current Addiction Reports*, *2*(2), 156–162. https://doi.org/10.1007/s40429-015-0054-y

Bowen, S., & Marlatt, A. (2009). Surfing the urge: Brief mindfulness-based intervention for college student smokers. *Psychology of Addictive Behaviors*, *23*(4), 666–671. https://doi.org/10.1037/a0017127

Brand, M., Young, K. S., Laier, C., Wölfling, K., & Potenza, M. N. (2016). Integrating psychological and neurobiological considerations regarding the development and maintenance of specific Internet-use disorders: An Interaction of Person-Affect-Cognition-Execution (I-PACE) model. *Neuroscience & Biobehavioral Reviews*, *71*, 252–266. https://doi.org/10.1016/j.neubiorev.2016.08.033

Brand, M., Wegmann, E., Stark, R., Müller, A., Wölfling, K., Robbins, T. W., & Potenza, M. N. (2019).

The Interaction of Person-Affect-Cognition-Execution (I-PACE) model for addictive behaviors: Update, generalization to addictive behaviors beyond internet-use disorders, and specification of the process character of addictive behaviors. *Neuroscience & Biobehavioral Reviews*, *104*, 1–10. https://doi.org/10.1016/j.neubiorev.2019.06.032

Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics*, *2*, 671–698. https://doi.org/10.1146/annurev.economics.102308.124324

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review*, *151*(3), 1211–1254. https://doi.org/10.2307/3312889

Chang, A.-M., Aeschbach, D., Duffy, J. F., & Czeisler, C. A. (2015). Evening use of light-emitting eReaders negatively affects sleep, circadian timing, and next-morning alertness. *Proceedings of the National Academy of Sciences*, *112*(4), 1232–1237. https://doi.org/10.1073/pnas.1418490112

Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science*, *60*(10), 2563–2582. https://doi.org/10.1287/mnsc.2014.1901

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2016). Situational strategies for self-control. *Perspectives on Psychological Science*, *11*(1), 35–55. https://doi.org/10.1177/1745691615623247

Elhai, J. D., Dvorak, R. D., Levine, J. C., & Hall, B. J. (2017). Problematic smartphone use: A conceptual overview and systematic review of relations with anxiety and depression psychopathology. *Journal of Affective Disorders*, *207*, 251–259. https://doi.org/10.1016/j.jad.2016.08.030

Elhai, J. D., Vasquez, J. K., Lustgarten, S. D., Levine, J. C., & Hall, B. J. (2018). Proneness to boredom mediates relationships between problematic smartphone use with depression and anxiety severity. *Social Science Computer Review*, *36*(6), 707–720. https://doi.org/10.1177/0894439317741087

Elster, J. (1979). *Ulysses and the Sirens: Studies in rationality and irrationality*. Cambridge University Press.

Exelmans, L., & Van den Bulck, J. (2016). Bedtime mobile phone use and sleep in adults. *Social Science & Medicine*, *148*, 93–101. https://doi.org/10.1016/j.socscimed.2015.11.037

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Appleton-Century-Crofts.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Garcia, S. M., & Tor, A. (2009). The N-effect: More competitors, less competition. *Psychological Science*, *20*(7), 871–877. https://doi.org/10.1111/j.1467-9280.2009.02385.x

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on Psychological Science*, *8*(6), 634–650. https://doi.org/10.1177/1745691613504114

Giné, X., Karlan, D., & Zinman, J. (2010). Put your money where your butt is: A commitment contract for smoking cessation. *American Economic Journal: Applied Economics*, *2*(4), 213–235. https://doi.org/10.1257/app.2.4.213

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*(7), 493–503. https://doi.org/10.1037/0003-066X.54.7.493

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, *38*, 69–119. https://doi.org/10.1016/S0065-2601(06)38002-1

Grüning, D. J., Riedel, F., & Lorenz-Spreen, P. (2023). Directing smartphone use through the self-nudge app one sec. *Proceedings of the National Academy of Sciences*, *120*(8), e2213114120. https://doi.org/10.1073/pnas.2213114120

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical

studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE. https://doi.org/10.1109/HICSS.2014.377

Harkin, B., Webb, T. L., Chang, B. P. I., Prestwich, A., Conner, M., Kellar, I., Benn, Y., & Sheeran, P. (2016). Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychological Bulletin*, *142*(2), 198–229. https://doi.org/10.1037/bul0000025

Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., Hazen, N., Herman, J., Katz, E. S., et al. (2015). National Sleep Foundation's sleep time duration recommendations: Methodology and results summary. *Sleep Health*, *1*(1), 40–43. https://doi.org/10.1016/j.sleh.2014.12.010

Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology*, *102*(6), 1318–1335. https://doi.org/10.1037/a0026545

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science*, *4*(2), 162–176. https://doi.org/10.1111/j.1745-6924.2009.01116.x

Holte, A. J., & Ferraro, F. R. (2023). True colors: Grayscale setting reduces screen time in college students. *The Social Science Journal*, *60*(2), 274–290. https://doi.org/10.1080/03623319.2020.1737461

Hunt, M. G., Marx, R., Lipson, C., & Young, J. (2018). No more FOMO: Limiting social media decreases loneliness and depression. *Journal of Social and Clinical Psychology*, *37*(10), 751–768. https://doi.org/10.1521/jscp.2018.37.10.751

Kardefelt-Winther, D. (2014). A conceptual and methodological critique of internet addiction research: Towards a model of compensatory internet use. *Computers in Human Behavior*, *31*, 351–354. https://doi.org/10.1016/j.chb.2013.10.059

Kim, J., Park, J., Lee, H., Ko, M., & Lee, U. (2019). LocknType: Lockout task intervention for discouraging smartphone app use. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). ACM. https://doi.org/10.1145/3290605.3300927

Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research*, *43*(1), 39–58. https://doi.org/10.1509/jmkr.43.1.39

Klasnja, P., Hekler, E. B., Shiffman, S., Boruvka, A., Almirall, D., Tewari, A., & Murphy, S. A. (2015). Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychology*, *34*(Suppl.), 1220–1228. https://doi.org/10.1037/hea0000305

Ko, M., Yang, S., Lee, J., Heizmann, C., Jeong, J., Lee, U., Shin, D., Yatani, K., Song, J., & Chung, K.-M. (2015). NUGU: A group-based intervention app for improving self-regulation of limiting smartphone use. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 1235–1245). ACM. https://doi.org/10.1145/2675133.2675244

Korotitsch, W. J., & Nelson-Gray, R. O. (1999). An overview of self-monitoring research in assessment and treatment. *Psychological Assessment*, *11*(4), 415–425. https://doi.org/10.1037/1040-3590.11.4.415

Kushlev, K., Proulx, J., & Dunn, E. W. (2016). "Silence your phones": Smartphone notifications increase inattention and hyperactivity symptoms. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1011–1020). ACM. https://doi.org/10.1145/2858036.2858359

Kwon, M., Kim, D.-J., Cho, H., & Yang, S. (2013). The Smartphone Addiction Scale: Development and validation of a short version for adolescents. *PLOS ONE*, *8*(12), e83558. https://doi.org/10.1371/journal.pone.0083558

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, *112*(2), 443–478. https://doi.org/10.1162/003355397555253

Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, *40*(6), 998–1009. https://doi.org/10.1002/ejsp.674

Landers, R. N., Bauer, K. N., & Callan, R. C. (2017). Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behavior*, *71*, 508–515. https://doi.org/10.1016/j.chb.2015.08.008

Lehrer, P. M., & Gevirtz, R. (2014). Heart rate variability biofeedback: How and why does it work? *Frontiers in Psychology*, *5*, 756. https://doi.org/10.3389/fpsyg.2014.00756

Levenson, J. C., Shensa, A., Sidani, J. E., Colditz, J. B., & Primack, B. A. (2016). The association between social media use and sleep disturbance among young adults. *Preventive Medicine*, *85*, 36–41. https://doi.org/10.1016/j.ypmed.2016.01.001

Lin, Y.-H., Lin, Y.-C., Lee, Y.-H., Lin, P.-H., Lin, S.-H., Chang, L.-R., Tseng, H.-W., Yen, L.-Y., Yang, C. C. H., & Kuo, T. B. J. (2015). Time distortion associated with smartphone addiction: Identifying smartphone addiction via a mobile application (App). *Journal of Psychiatric Research*, *65*, 139–145. https://doi.org/10.1016/j.jpsychires.2015.04.003

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, *65*(3), 272–292. https://doi.org/10.1006/obhd.1996.0028

Loid, K., Täht, K., & Rozgonjuk, D. (2020). Do pop-up notifications regarding smartphone use decrease screen time, phone checking behavior, and self-reported problematic smartphone use? Evidence from a two-month experimental study. *Computers in Human Behavior*, *102*, 22–30. https://doi.org/10.1016/j.chb.2019.08.007

Lukoff, K., Yu, C., Kientz, J., & Hiniker, A. (2018). What makes smartphone use meaningful or meaningless? *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(1), 1–26. https://doi.org/10.1145/3191754

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. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–17). ACM. https://doi.org/10.1145/3411764.3445467

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. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). ACM. https://doi.org/10.1145/3290605.3300361

Lyngs, U., Lukoff, K., Slovak, P., Seymour, W., Webb, H., Jirotka, 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. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). ACM. https://doi.org/10.1145/3313831.3376672

Maier, M., Bartoš, F., Stanley, T. D., Shanks, D. R., Harris, A. J. L., & Wagenmakers, E.-J. (2022). No evidence for nudging after adjusting for publication bias. *Proceedings of the National Academy of Sciences*, *119*(31), e2200300119. https://doi.org/10.1073/pnas.2200300119

Mark, G., Iqbal, S., & Czerwinski, M. (2017). How blocking distractions affects workplace focus and productivity. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers* (pp. 928–934). ACM. https://doi.org/10.1145/3123024.3124558

Meng, S.-Q., Cheng, J.-L., Li, Y.-Y., Yang, X.-Q., Zheng, J.-W., Chang, X.-W., Shi, Y., Chen, Y., Lu, L., et al. (2022). Global prevalence of digital addiction in general population: A systematic review and meta-analysis. *Clinical Psychology Review*, *92*, 102128. https://doi.org/10.1016/j.cpr.2022.102128

Mertens, S., Herberz, M., Hahnel, U. J. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences*, *119*(1), e2107346118. https://doi.org/10.1073/pnas.2107346118

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, *46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Monge Roffarello, A., & De Russis, L. (2019). The race towards digital wellbeing: Issues and opportunities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM. https://doi.org/10.1145/3290605.3300616

Monge Roffarello, A., & De Russis, L. (2023). Achieving digital wellbeing through digital self-control tools: A systematic review and meta-analysis. *ACM Transactions on Computer-Human Interaction*, *30*(4), 1–66. https://doi.org/10.1145/3571810

Montag, C., Lachmann, B., Herrlich, M., & Zweig, K. (2019). Addictive features of social media/messenger platforms and freemium games against the background of psychological and economic theories. *International Journal of Environmental Research and Public Health*, *16*(14), 2612. https://doi.org/10.3390/ijerph16142612

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89*(1), 103–124. https://doi.org/10.1257/aer.89.1.103

Okeke, F., Sobolev, M., Dell, N., & Estrin, D. (2018). Good vibrations: Can a digital nudge reduce digital overload? In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 1–12). ACM. https://doi.org/10.1145/3229434.3229463

Olson, J. A., Sandra, D. A., Colucci, É. S., Al Bikaii, A., Chmoulevitch, D., Nahas, J., Raz, A., & Veissière, S. P. L. (2022). Smartphone addiction is increasing across the world: A meta-analysis of 24 countries. *Computers in Human Behavior*, *129*, 107138. https://doi.org/10.1016/j.chb.2021.107138

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, *3*(2), 173–182. https://doi.org/10.1038/s41562-018-0506-1

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, *16*(1), 105–114. https://doi.org/10.1007/s00779-011-0412-2

Panova, T., & Carbonell, X. (2018). Is smartphone addiction really an addiction? *Journal of Behavioral Addictions*, *7*(2), 252–259. https://doi.org/10.1556/2006.7.2018.49

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, *5*(11), 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Patel, M. S., Small, D. S., Harrison, J. D., Fortunato, M. P., Oon, A. L., Rareshide, C. A. L., et al. (2019). Effectiveness of behaviorally designed gamification interventions with social incentives for increasing physical activity among overweight and obese adults across the United States: The STEP UP randomized clinical trial. *JAMA Internal Medicine*, *179*(12), 1624–1632. https://doi.org/10.1001/jamainternmed.2019.3505

Peters, D., Calvo, R. A., & Ryan, R. M. (2018). Designing for motivation, engagement and wellbeing in digital experience. *Frontiers in Psychology*, *9*, 797. https://doi.org/10.3389/fpsyg.2018.00797

Przybylski, A. K., & Weinstein, N. (2013). Can you connect with me now? How the presence of mobile communication technology influences face-to-face conversation quality. *Journal of Social and*

Personal Relationships, *30*(3), 237–246. https://doi.org/10.1177/0265407512453827

Radtke, T., Apel, T., Schenkel, K., Keller, J., & von Lindern, E. (2022). Digital detox: An effective solution in the smartphone era? A systematic literature review. *Mobile Media & Communication*, *10*(2), 190–215. https://doi.org/10.1177/20501579211028647

Reason, J. (2000). Human error: Models and management. *BMJ*, *320*(7237), 768–770. https://doi.org/10.1136/bmj.320.7237.768

Rixen, J. O., Meinhardt, L.-M., Glöckler, M., Ziegenbein, M.-L., Schlothauer, A., Colley, M., Rukzio, E., & Gugenheimer, J. (2023). The loop and reasons to break it: Investigating infinite scrolling behaviour in social media applications and reasons to stop. *Proceedings of the ACM on Human-Computer Interaction*, *7*(MHCI), 1–22. https://doi.org/10.1145/3604275

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research Reviews*, *18*(3), 247–291. https://doi.org/10.1016/0165-0173(93)90013-P

Ruiz Pardo, A. C., & Minda, J. P. (2022). Reexamining the "brain drain" effect: A replication of Ward et al. (2017). *Acta Psychologica*, *230*, 103717. https://doi.org/10.1016/j.actpsy.2022.103717

Schelling, T. C. (1984). Self-command in practice, in policy, and in a theory of rational choice. *American Economic Review*, *74*(2), 1–11.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Sharma, B., Lee, S. S., & Johnson, B. K. (2022). The dark at the end of the tunnel: Doomscrolling on social media newsfeeds. *Technology, Mind, and Behavior*, *3*(1), 144–156. https://doi.org/10.1037/tmb0000059

Sheeran, P., & Webb, T. L. (2016). The intention-behavior gap. *Social and Personality Psychology Compass*, *10*(9), 503–518. https://doi.org/10.1111/spc3.12265

Sohn, S. Y., Rees, P., Wildridge, B., Kalk, N. J., & Carter, B. (2019). Prevalence of problematic smartphone usage and associated mental health outcomes amongst children and young people: A systematic review, meta-analysis and GRADE of the evidence. *BMC Psychiatry*, *19*, 356. https://doi.org/10.1186/s12888-019-2350-x

Steindl, C., Jonas, E., Sittenthaler, S., Traut-Mattausch, E., & Greenberg, J. (2015). Understanding psychological reactance: New developments and findings. *Zeitschrift für Psychologie*, *223*(4), 205–214. https://doi.org/10.1027/2151-2604/a000222

Stothart, C., Mitchum, A., & Yehnert, C. (2015). The attentional cost of receiving a cell phone notification. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4), 893–897. https://doi.org/10.1037/xhp0000100

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1

Strathern, M. (1997). 'Improving ratings': Audit in the British University system. *European Review*, *5*(3), 305–321.

Thornton, B., Faires, A., Robbins, M., & Rollins, E. (2014). The mere presence of a cell phone may be distracting: Implications for attention and task performance. *Social Psychology*, *45*(6), 479–488. https://doi.org/10.1027/1864-9335/a000216

Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 Well-Being Index: A systematic review of the literature. *Psychotherapy and Psychosomatics*, *84*(3), 167–176. https://doi.org/10.1159/000376585

Tran, J. A., Yang, K. S., Davis, K., & Hiniker, A. (2019). Modeling the engagement-disengagement cycle of compulsive phone use. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM. https://doi.org/10.1145/3290605.3300542

Twenge, J. M., & Campbell, W. K. (2018). Associations between screen time and lower psychological well-being among children and adolescents: Evidence from a population-based study. *Preventive*

Medicine Reports, *12*, 271–283. https://doi.org/10.1016/j.pmedr.2018.10.003

Verplanken, B., & Wood, W. (2006). Interventions to break and create consumer habits. *Journal of Public Policy & Marketing*, *25*(1), 90–103. https://doi.org/10.1509/jppm.25.1.90

Ward, A. F., Duke, K., Gneezy, A., & Bos, M. W. (2017). Brain drain: The mere presence of one's own smartphone reduces available cognitive capacity. *Journal of the Association for Consumer Research*, *2*(2), 140–154. https://doi.org/10.1086/691462

Watson, N. F., Badr, M. S., Belenky, G., Bliwise, D. L., Buxton, O. M., Buysse, D., Dinges, D. F., Gangwisch, J., Grandner, M. A., et al. (2015). Recommended amount of sleep for a healthy adult: A joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. *Sleep*, *38*(6), 843–844. https://doi.org/10.5665/sleep.4716

Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological Review*, *114*(4), 843–863. https://doi.org/10.1037/0033-295X.114.4.843

Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology*, *67*, 289–314. https://doi.org/10.1146/annurev-psych-122414-033417

Wood, W., Tam, L., & Witt, M. G. (2005). Changing circumstances, disrupting habits. *Journal of Personality and Social Psychology*, *88*(6), 918–933. https://doi.org/10.1037/0022-3514.88.6.918

World Health Organization. (2019). *International classification of diseases for mortality and morbidity statistics* (11th rev.). World Health Organization.

Zaccaro, A., Piarulli, A., Laurino, M., Garbella, E., Menicucci, D., Neri, B., & Gemignani, A. (2018). How breath-control can change your life: A systematic review on psycho-physiological correlates of slow breathing. *Frontiers in Human Neuroscience*, *12*, 353. https://doi.org/10.3389/fnhum.2018.00353

Zimmermann, L. (2021). "Your screen-time app is keeping track": Consumers are happy to monitor but unlikely to reduce smartphone usage. *Journal of the Association for Consumer Research*, *6*(3), 377–382. https://doi.org/10.1086/714365

A Parameter reference

Table 5. Parameters referred to in Section 5, with the rationale for each value.

Parameter	Value	Rationale
Measurement		
Enforcement tick	500 ms (2 Hz)	Fast enough that a blocked app is covered almost as it opens.
Usage cache	30 s	Usage changes a minute at a time; walking the event log has a cost.
Baseline window	7 complete days	Excludes today's partial total; the OS log holds about a week.
Sleep gap minimum	3 h unbroken	Shorter gaps are naps or idle periods.
Gap-breaking activity	>60 s in foreground	Reading a notification should not count as waking.
Sleep window	circular mean ±4 h	Learns night-shift patterns; fixed fallback below 5 days or circular SD > 3 h.
Screen-time exclusions	Beave, alarm apps, dialler	A screen kept on by an alarm or a call is not use; tied to the app, not to a time window.
Wellness score scale	0–100, whole points	A week that improves moves the number visibly.
Wellness bands	from 25, 40, 55, 70, 85	Six fixed bands, Poor to Excellent; the typical user (65) falls in Average.

Parameter	Value	Rationale
Plan		
Goal dial range	15 min to 75% of baseline	At least a 25% cut; a goal can never be an increase.
Goal dial opening value	50% of baseline	Ambitious but adjustable.
Morning and night windows	≥1 h each, default 2 h	Long enough to cover the most cue-laden periods.
First-day grace	usage counted from plan creation	Avoids blocking for usage that preceded the plan; written once.
Free rule allowance	1 (plus the daily limit)	Caps creation, never possession.
Pre-commit		
Police cooldowns	1 h, 3 h, 24 h, 3 days	Long enough for a hot state to pass; 1 h is the picker default.
Lockout redirect interval	700 ms	Re-asserted on every tick, so one missed transition is not an escape.
Offline-hours default	22:00–07:00	Covers the pre-sleep period and the night.
Offline-hours kept apps	up to 10	For apps needed online at night; adding one is loosening under Police.
Interrupt		
Breathing cycles	3	Enough to require deliberate effort; short enough not to provoke reactance.
Inhale cue, minimum hold	4 s, 3 s	Prevents the gate being tapped through.
Exhale	6 s	Exhale-weighted; with the inhale, six breaths per minute.
Pass duration	5 min per app	Enough for a deliberate task; expires automatically, and dies with the service.
Limit notification	80% of daily limit	Arrives while the user can still act.
Recap	21:00	Reviews the day before the night window.
Time-used pill	4 s on open, every 30 min	Brief enough not to become wallpaper.
Scroll limit rungs	15, 30, 45, 60 min (default 30)	15 min is roughly 30 short videos.
Scroll gap cap	60 s per gap	Reading counts; absence does not.
Sitting resume window	10 min since last scroll	A brief trip out of the app does not reset the clock.
Scroll break	10 min, fixed	Ends the sitting; returning does not restart the clock.
Unlock check questions	5, in rotation	A question that never changes is answered by rote.
Unlock check check-in	off, 5, 10, 15 min (default 10)	Catches too long spent in the app the user came for.
Unlock check grace	3 min after screen-off	A screen that timed out mid-task does not ask again.
Reinforce		
Face-down threshold	$g_z \leq -7.5$ m/s², horizontal ≤ 4.5 m/s²	Flat and face down, not merely tilted.
Face-up threshold	$g_z \geq 4.0$ m/s²	Hysteresis prevents spurious pauses.
Still-dwell fallback	1.5 s, motion < 0.4 m/s²	Between table noise and hand tremor; for phones without a usable proximity sensor.
Stick rate	1 per 10 focused minutes	Linear, so fragmentation earns nothing extra.
Session presets	10, 20, 30, 60, 180 min	Short entry points; the rate rewards longer sessions proportionally.
Pomodoro runs	25/5 or 50/10 min, 2–6 rounds	Sticks credited on the run's total, so the rate stays linear.
Free pauses	2 per session	Lifting the phone does not use one.
Week boundary	Monday 00:00 UTC	One shared week for a global board.