Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-024-03845-1 OPEN

Self-control failures, as judged by themselves

Leonhard K. Lades [1]✉ & Liam Delaney[2]

The existence of self-control failures is often used to legitimize public policy interventions. The argument is that reducing self-control failures can make people better off, as judged by themselves. However, there is only scarce evidence on the frequency and welfare costs of self-control failures. This paper presents a survey method that allows us to measure self-control failures in everyday life and to identify their welfare costs in terms of associations with experienced subjective well-being. We present novel survey evidence using this method and discuss its implications for behavioural welfare economics and behavioural public policy.

[1] Behavioural Science Centre and Economics, University of Stirling, Stirling, UK. [2] Department of Psychological & Behavioural Science, London School of Economics & Political Science, London, UK. ✉email: l.k.lades@stir.ac.uk

Introduction

Self-control problems are among the most frequently discussed topics in moral philosophy, political economics, psychology, and behavioural public policy. Influential philosophical antecedents of modern Economics such as the works of Adam Smith and David Hume placed strong emphasis on the importance of self-control in understanding economic decision making and governance. More recent contributions stress the importance of changing preferences over time and dynamic inconsistencies (Ainslie 1975; Laibson 1997), establish the importance of visceral influences (Loewenstein 1996), argue that willpower is the greatest human strength (Baumeister and Tierney 2011),[1] show that self-control correlates with a number of beneficial life outcomes (Moffitt et al. 2011), and demonstrate that self-control conflicts are a common phenomenon in people's everyday lives (Hofmann et al. 2012).

Policy-makers and academics sometimes argue that the adequate response to the identification of self-control problems is to design policies that help individuals to overcome their self-control problems in order to encourage decisions that make individuals *better off, as judged by themselves* (Thaler and Sunstein 2008; Sunstein 2018; Ericson and Laibson 2019).[2] Such policies include the provision of social security benefits and illiquid retirement savings accounts, restrictions on payday borrowing, compulsory education laws, bans to smoke indoors, bans to purchase alcohol during the night, cooling down periods when shopping online, and taxes on "sin goods". Additionally, behavioural policies such as nudges that do not limit freedom of choice are suggested to reduce self-control failures (Thaler and Sunstein 2008). Duckworth et al. (2018) review the most important behavioural and psychological ways to reduce self-control failures.

However, not everybody agrees with the idea of policy intervention as a legitimate response to findings of frequent self-control failures. One argument against policy intervention focuses on the difficulty for policy-makers to identify what *better off, as judged by themselves* actually means (Rizzo and Whitman 2008; Sugden 2017; Sunstein 2019; Grubiak et al. 2024).[3] When short-term preferences conflict with long-term preferences, identifying individuals' true preferences is difficult as it is not clear whether satisfying the short-term preference or the long-term preference would make individuals better off (Read 2006). For example, is eating an unhealthy dessert a self-control failure or a legitimate manifestation of instantaneous preferences? Is the decision to renege on a resolution to go to the gym the result of present bias or the result of new information? Is under-saving an error of over-optimism or the result of the realization that we only live once? Is the long-term preference itself context-dependent and thus constructed and of limited normative weight (Grubiak et al. 2024)? Highlighting such difficulties, Sugden (2017) argues that there is very little evidence on the extent to which people actually make decisions against their own interests as they define them,[4] and Ericson and Laibson (2019) ask as an open question "*What welfare criterion should we use to evaluate inter-temporal trade-offs?*"

Even if many behaviours were the result of self-control failures, policy interventions might still not be warranted by the *better off, as judged by themselves* criterion. Additionally, the policy-maker would need to show in a cost-benefit analysis that self-control failures have welfare costs and make individuals worse off (e.g., Berg and Gigerenzer 2010). If self-control failures are not costly, little can be said that would legitimize policies that aim to reduce self-control failures to make individuals better off. Nevertheless, currently applied policies that aim to overcome self-control problems are almost never based on solid empirical evidence on the frequency and costs of self-control failures.

This paper approaches these issues associated with self-control problems from a pragmatic, empirical perspective. Most discussions in behavioural public policy about the legitimacy of policies to overcome self-control failures and about the welfare consequences of self-control failures are conceptual. However, we aim to contribute to these discussions with some data by making at least three contributions. First, the paper suggests using a method that allows us to empirically quantify and characterize self-acknowledged self-control failures defined as everyday behaviours that are enacted, although people attempted to resist enacting the behaviour. The method is an adaptation of the Day Reconstruction Method (Kahneman et al. 2004; Knabe et al. 2010; Doyle et al. 2017; Lades et al. 2022) used to measure desires and decision-making processes in people's everyday lives as suggested by Delaney and Lades (2017). We present survey evidence using this method to show that the method allows us to measure the frequency of self-control failures in daily life, to identify in which domains and situations self-control failures are particularly likely to occur, and to identify whether some people are more likely than others to have self-control failures.

Second, the paper illustrates that it is possible to obtain data that can contribute to cost-benefit analyses of the effects of self-control failures. Self-control problems are typically characterized by trade-offs between immediate pleasures and long-term costs. We focus on the short-term welfare effects of engaging or not engaging in tempting and potentially pleasurable activities assuming that experienced subjective well-being (SWB) is a valid short-term welfare measure in this context (Dolan and Kahneman 2008). We show that these short-term welfare effects can be measured using the Day Reconstruction Method. An alternative way to measure self-control failures and contemporary SWB is experience sampling where participants are contacted multiple times a day on their smartphones (Hofmann et al. 2014). However, the participant burden of day reconstruction studies is lower and large-scale studies that are often required to inform policy are more feasible with this method.

Finally, the paper discusses whether and how data on everyday self-control failures and their short-term welfare associations can and should be used to inform public policies. For example, data on everyday decision-making processes may be helpful to "purify" preferences, i.e., to isolate the true, normatively relevant preferences from the influence of weakness of will on choices. Moreover, a convincing argument (at least from a welfare perspective) for a policy intervention to overcome self-control failures could be made if the satisfaction of a short-term temptation did not provide short-term welfare. In this case, a policy intervention that reduces self-control failures might increase long-term welfare without reducing short-term welfare and would thus be acceptable from a welfare perspective and the *better off, as judged by themselves* criterion.

The remainder of the paper is structured as follows. Section 2 discusses different definitions of self-control failures and suggests that subjective well-being data are a valid source for measuring the short-term welfare effects of self-control failures. Section 3 presents the approach we use to measure self-acknowledged self-control failures in everyday life, as well as experienced subjective well-being at the same point in time. Section 4 presents data to illustrate that it is indeed possible to measure self-control failures and their associations with experienced subjective well-being in everyday life. Section 5 explores how such everyday data may be used in future work in behavioural public policy to inform questions related to the design and the welfare assessment of policies that aim to reduce self-control failures. The last section concludes.

What are self-control failures and are they costly?

Sequential self-control failures Self-control failures can be understood in at least two ways: sequentially and in parallel (Read 2008).[5] Most of the economic inter-temporal choice literature on present bias and instant gratification uses the sequential definition and understands self-control failures as dynamic inconsistencies where people's preferences change nearer to the decision time from preferring a larger later reward to preferring a smaller sooner reward (Ainslie 1975; Laibson 1997; Frederick et al. 2002; Ericson and Laibson 2019). In this literature, self-control is about "the present" versus "the future" and self-control failures are present when individuals act against previously made plans. In a typical case of a sequential self-control failure, an individual comes up with a plan at Time 1, acts at Time 2 in a way that is not in line with the plan, and regrets the action at Time 3.

There are at least two problems with the sequential definition of self-control failures, one descriptive, the other normative. First, a robust identification of self-control failures based on dynamic inconsistencies is difficult, because people's preferences can change for a multitude of reasons (Kurzban et al. 2013; Inzlicht et al. 2014; Rizzo 2016). For example, new information about the internal states of the individual and/or new information about the outside world might have become available at Time 2. For policy-makers, it might be impossible to distinguish between self-control failures and other reasons of dynamically inconsistent choices.

The second problem of the sequential definition of self-control failures is the difficulty to defend the claim that the Time 2 preferences are of less normative value than the preferences at Time 1 and Time 3. While many behavioural economic models at least implicitly assume that choices that are in line with long-term preferences make them better off, *as judged by themselves* (O'Donoghue and Rabin 1999; Thaler and Sunstein 2008; Binder and Lades 2015), people might also have a fully rational mantra at Time 2: "*Enjoy life now. This is not a rehearsal*" as discussed by Sunstein (2016, p. 48). It is possible that satisfying the Time 2 preferences maximizes well-being and that people are better off when satisfying short-term preferences even when they are conflicting (Berg and Gigerenzer 2010). Moreover, Sugden (2008) argues that there is no normative reason to favour long-term over short-term preferences. One way to overcome (or ignore) this normative problem is to rely on some external moral standard and external judgments about the ethicality or "rightness" of an action to inform policies (Read 2006; Sunstein 2019). However, this approach departs from the *as judged by themselves* criterion to legitimize policy interventions and thus attracts other criticism.

Parallel self-control .failures The second way to define self-control failures is in line with the psychological literature on desire-goal conflicts. In this literature, "*self-control is not about present versus future; it is about deliberation versus affect*" (Loewenstein 2018, p. 97) and both can be present at the same point in time, i.e. in parallel. In parallel models, self-control failures are viewed as the result of intra-individual conflicts between the short-term preference to satisfy immediate desires and the long-term preferences to stick to higher-order goals. Self-control problems arise when short-term preferences conflict with long-term preferences, for example when our desires to eat high-calorific foods, smoke cigarettes, or buy the newest consumption goods conflict with long-term goals related to healthy, wealthy, and happy living (Thaler and Shefrin 1981; Hoch and Loewenstein 1991; Loewenstein 1996; Gul and Pesendorfer 2001; Baumeister 2002; Milkman et al. 2008; Moffitt et al. 2011; Duckworth et al. 2016; Lades et al. 2017). Self-control failures arise when individuals attempt to resist enacting the immediate desire, but fail in this resistance attempt (Kotabe and Hofmann 2015; Lades and Hofmann 2019).

The advantage of the parallel definition of self-control failures is that we can identify self-control failures unambiguously as behaviours that are conducted despite resistance attempts. Observing that individuals attempt to resist enacting a short-term desire (i.e., observing that individuals use willpower or alternative reactive self-control strategies) suggests that in the moment of choice the individuals themselves consider it costly in the long run to enact the tempting desire. The observation of a behaviour that is enacted despite a resistance attempt is thus a clear signal that people would have preferred a different course of action. The behaviour resulting from the self-control failure is hence not the behaviour that makes people better off, as judged by themselves and we do not have to deviate from the *as judged by themselves* criterion. For these reasons, we adopt the parallel definition of self-control failures in this paper.

This definition of self-control failures, however, is relatively restrictive and does not capture all behaviours one might reasonably characterise as a self-control failure. This restrictiveness is due to the requirement that people actively resist a desire for it to be a self-control failure. But desires that conflict with long-term goals might be enacted automatically and very fast without opportunities to become aware of a problem and attempt to resist the desire. Only later might people realise that their behaviour was not in their long-term interests and that they should have resisted it. Defining self-control failures as desires that are enacted despite a resistance attempt excludes these behaviours. Nevertheless, the benefit of our definition is that we avoid the problems related to the sequential definition of self-control failures.

The costs and benefits of self-control .failures As mentioned in the introduction, identifying self-control failures is not sufficient to legitimize policy interventions. At least from a welfare perspective, one would additionally need to show that self-control failures are costly. When analysing the costs and benefits of self-control failures, it is typically assumed that the benefits arise in the short-term and that the costs arise in the long-term.[6] For example, satisfying the short-term desire to eat unhealthy food is pleasurable immediately and the detrimental health consequences are only delayed. Hence, a convincing argument (at least from a welfare perspective) for a policy intervention to overcome self-control failures could be made if the satisfaction of the short-term desire does not provide short-term benefits. For example, Gruber and Mullainathan (2006) suggest that the short-run and the long-run welfare of smokers who were on the margin of quitting improved as a result of the imposition of smoking taxes. There might not always be a trade-off between either increasing short-term welfare or increasing long-term welfare and some interventions might increase long-term welfare without short-term costs.

One way to measure the short-term effects of policy interventions is to measure the experienced subjective well-being (SWB) associated with the satisfaction of short-term desires. Since the short-term benefits of enacting temptations are often understood in terms of Benthamite pleasure and absence of pain,[7] this experiential short-term welfare definition seems warranted in this context (Dolan and Kahneman 2008). Previous research has analysed the links between self-control and simultaneous experiences. For example, Hofmann et al. (2014) use experience sampling technology to show that people are happier when they enact non-conflicting desires compared to situations in which they enact conflicting desires (i.e., in situations in which self-control failures are present).[8] We

follow this literature and present a study in which we measure self-control failures and their associations with SWB at the same point in time.

Method

We present a study in which we use a modified version of the Day Reconstruction Method (DRM) which was introduced by Kahneman et al. (2004). The DRM is a diary technique that helps study participants to recall what they did and how they felt "yesterday". It is typically used to measure the momentary happiness (or experienced utility) individuals felt for a full day "yesterday", i.e., one day before the data is collected. But the DRM can also be used to directly measure desires (or subjective preferences, or what people want) in everyday life, their determinants, and whether desires are successfully or unsuccessfully resisted (Delaney and Lades 2017). In this paper, we present data from a DRM study with University students that measured both self-control failures in everyday life and simultaneous subjective well-being ratings to illustrate the potential of the method for use in behavioural welfare economics and behavioural public policy.

.Procedure As it is common in DRM studies, participants first completed a private time-usage diary (on paper) in which they recalled and reported the previous day as a discrete set of episodes (see Supplementary Information in Delaney and Lades 2017). Following Kahneman et al. (2004), we asked participants to think of their day as a film divided into multiple episodes. Participants could divide their day into as many episodes as they saw fit, with a maximum of 20 episodes. We explicitly mentioned that participants could take their paper diary home after the completion of the questionnaire and that the diary was merely an aid to recalling what happened and how they felt yesterday.

After the participants had finished the diary, they went through each episode on the computer screen and answered questions about each one. In particular, they indicated where they were (e.g., at home or at the university), who they interacted with (e.g., friends, spouse/significant other, or nobody), and how they felt on a scale from 0 (not at all) to 6 (very much) regarding 11 states ("happy", "frustrated/annoyed", "enjoying myself", "angry/hostile", "stressed", "tired", "hungry", "thirsty", "drunk", "in control", and "focused"). We averaged their responses to "happy" and "enjoying myself" to calculate a positive affect score and "frustrated/annoyed", "angry/hostile", and "stressed" to create the negative affect score. Participants also provided information about their age, gender, marital status, country of origin, educational attainment, income, and life satisfaction.

Most importantly, we asked participants several questions about the desires (or short-term preferences) they experienced during each episode. We used the questions proposed by Hofmann et al. (2012) and asked "Did you feel any desires during this episode?" with a list of desire domains to select up to three from (Eat; Drink; Drink alcohol; Smoke; Sex; Use media; Spend money; Social contact; Leisure; Postpone; Work; Sport; Sleep; Other). We measured desire strength by asking "How strong was the desire you mentioned?" (on a scale from 1 to 7), conflict strength by asking "Sometimes, we prefer not to act upon our desires. On a scale from 1 to 7, how strongly was satisfying the desire at odds with one or more of your general life goals, aspirations, or plans that you have made before?" (on a scale from 1 to 7), resistance by asking "Did you attempt to resist satisfying the desire?" (with "Yes" and "No" as answer options), and desire enactment by asking "Did you satisfy this desire (even just a little bit)?" (again with "Yes" and "No" as answer options). We define self-control failures as situations in which participants attempted to resist enacting a desire but nonetheless enacted it (see Fig. 2). Since participants could select multiple desires in each episode, the data is structured on three levels with desires nested in episodes and episodes nested in individuals.

.Participants 248 individuals from the local University came to computer labs to participate in our study.[9] The participants were aged between 18 and 53 (M 23.29, SD 6.04), 67.47 per cent were female, 57.66 per cent had a college degree, 77.82 per cent were single, 13.31 per cent were married, 48.19 per cent were from the UK, and 57.26 per cent indicated that they had an annual income below £25.000.

Analysis .strategy We present the results in three subsections. In sub-section 4.1., we present descriptive results on the prevalence of self-control failures in the everyday lives of our participants. In sub-section 4.2., we show that it is possible to identify individual-specific, episode-specific, and desire-specific factors that predict whether a desire leads to a self-control failure. For this, we analyse the data on the desire level, i.e., the data set includes one row for each desire. We use logit regressions to predict the binary dependent variable "presence of a self-control failure" for each desire. Multi-level models with three levels are needed as desires are nested in episodes (i.e., there might be more than one desire for each episode), and episodes are nested in individuals (i.e., people recorded usually more than one episode). In sub-section 4.3., we test whether enacting short-term desires has short-term costs (or benefits) in terms of subjective well-being. We test whether the SWB that participants experience in episodes with self-control failures differs from the SWB felt in episodes without self-control failures. Since the information we have on SWB (i.e., positive and negative affect) are on the episode-level, we specify regression models differentiating between two levels, accounting for episodes being nested in individuals. More details are presented in the respective subsections below.

Results

How prevalent are self-control failures? The 248 participants reported 3082 episodes ($M = 1243$; SD = 3.37) and 3676 desires ($M = 14.80$; SD = 7.60) altogether. In 2127 episodes, at least one desire was present (see Fig. 1a). Altogether 2126 desires were enacted and at least one desire was enacted in 1414 episodes (see Fig. 1b). Of all desires, 531 (14.45 per cent) led to a self-control failure, which is in line with previous research (Hofmann et al. 2012) suggesting that desire is an important aspect of everyday life and that up to 14 per cent of all desires lead to self-control failures. At least one self-control failure was present in 420 episodes (see Fig. 1c). Figure 2 illustrates that we define self-control failures as desires that are enacted despite a resistance attempt. Similarly, self-control successes are defined as desires that are not enacted following a resistance attempt. However, to avoid self-control failures, people can use proactive or reactive self-control strategies (Kotabe and Hofmann 2015; Duckworth et al. 2016; Lades and Hofmann 2019). What we define as self-control successes as illustrated in Fig. 2 is limited to the successful use of *reactive* self-control strategies.[10]

Figure 3 summarizes the distribution of desires across the different desire domains. The number of all desires in each domain is represented by the height of the bars. Each bar is then partitioned into (i) desires that individuals did not attempt to resist, (ii) desires that individuals successfully resisted (self-control successes), and (iii) desires that individuals attempted to resist but nevertheless enacted (self-control failures). Self-control failures were most likely when participants had the desire to use social media and to postpone something, and participants were particularly successful in resisting the desire to sleep. The average

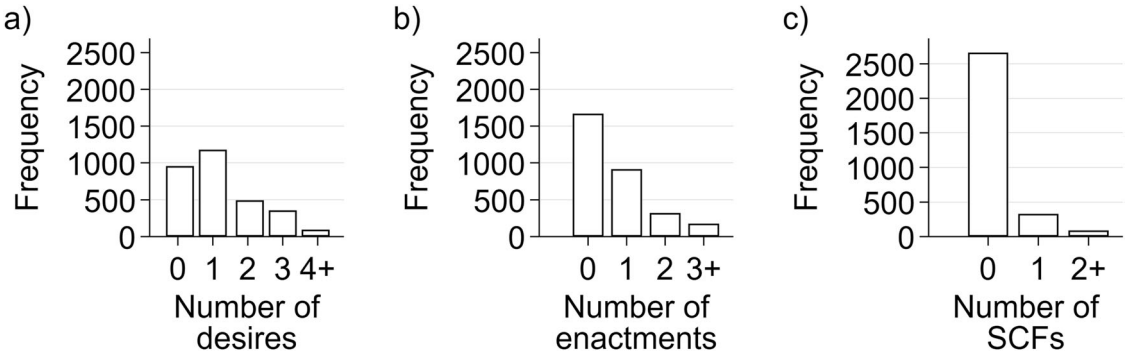

Fig. 1 Descriptive statistics about of self-reported desires, enactments, and self-control failures. a Number of desires reported by participants per episode. **b** Number of desires that participants enacted per episode. **c** Number of desires that participants enacted despite a resistance attempt (self-control failures) per episode.

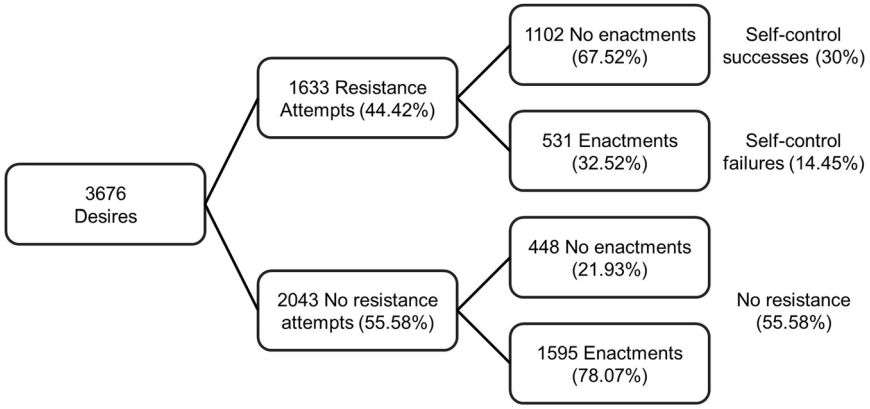

Fig. 2 Number of desires, resistance attempts, and desire enactments in the composition that illustrates the definitions of self-control failures and self-control successes.

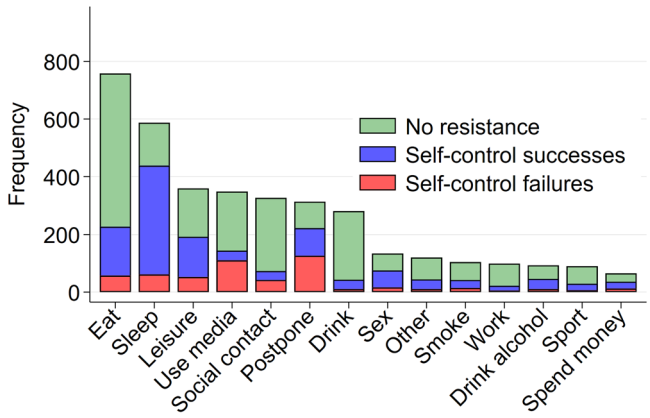

Fig. 3 Desires, resistance attempts, and self-control failures by desire domains.

desire strength was 5.04 (SD = 0.99) and the average conflict strength was 3.25 (SD = 2.06). Table 1 presents more details on the number of desires, resistance attempts, enactment, self-control failures, self-control successes, average desire strength, and average conflict strength by desire domain. For example, the strongest desires were in the domains "Other" and "Sport" and the strongest conflicts in the domains "Postponing" and "Spending Money". Figure 4 presents the diurnal patterns of the six most frequently mentioned desires with the same partitioning as in Fig. 3. For example, the desire to eat something is most prevalent around lunch which provides some confidence in the data.

What predicts self-control failures? We test whether the presence of a self-control failure can be explained by a set of individual-specific, episode-specific, and desire-specific variables using the following model

$$SCF_{ijk} = \beta_0 + \beta_1 \boldsymbol{Des}_{ijk} + \beta_2 \boldsymbol{Epi}_{ij} + \beta_3 \boldsymbol{Ind}_i + u_k + v_{jk} + \epsilon_{ijk}, \quad (1)$$

where SCF_{ijk} is the binary variable indicating the presence of a self-control failure of individual i, in episode j, of desire k The vector $\boldsymbol{Des}_{ijk}$ represents desire type, desire strength, and conflict strength. We included these desire-specific independent variables to identify whether self-control failures are more prevalent in some domains than in others and whether stronger and more conflicting desires lead to a higher probability of self-control failures as suggested in previous self-control research (Hofmann et al. 2012, Delaney and Lades, 2017). The vector $\boldsymbol{Epi}_{ij}$ represents the episode-specific predictors location, social interaction, and time of the day. These predictors are included to identify whether people are particularly prone to self-control failures in specific situations which might inform the design of behavioural public policy interventions. The episode-specific variables can also be interpreted as control variables as, for example, desire type, desire strength, and conflict strength might differ across locations, social situations, and time of the day. The vector $\boldsymbol{Ind}_i$ represents age, gender, marital status, education, country of origin, and income

Table 1 Information about desires by type.

	Desires	Resistance	Enactment	SCF[a]	SCS[b]	Mean desire strength	Mean conflict strength
Eat	758	227	469	57	170	5.04	2.66
Drink	280	43	228	10	33	5.06	2.17
Drink alcohol	93	46	52	10	36	4.61	3.58
Smoke	104	42	68	14	28	5.15	3.39
Sex	134	75	45	16	59	5.01	3.51
Use media	348	144	309	110	34	4.78	3.45
Spend money	65	36	35	12	24	4.89	4.17
Social contact	326	73	242	42	31	4.98	2.84
Leisure	359	192	183	52	140	4.97	3.73
Postpone	313	222	190	126	96	5.06	4.90
Work	99	22	57	5	17	5.05	2.32
Sport	90	29	43	6	23	5.23	2.89
Sleep	587	438	149	61	377	5.19	3.46
Other	120	44	56	10	34	5.40	3.18
Sum|Av[c]	3676	1633	2126	531	1102	5.03[c]	3.30[c]

[a]Self-control failures.
[b]Self-control successes.
[c]Mean desire strength and mean conflict strength.

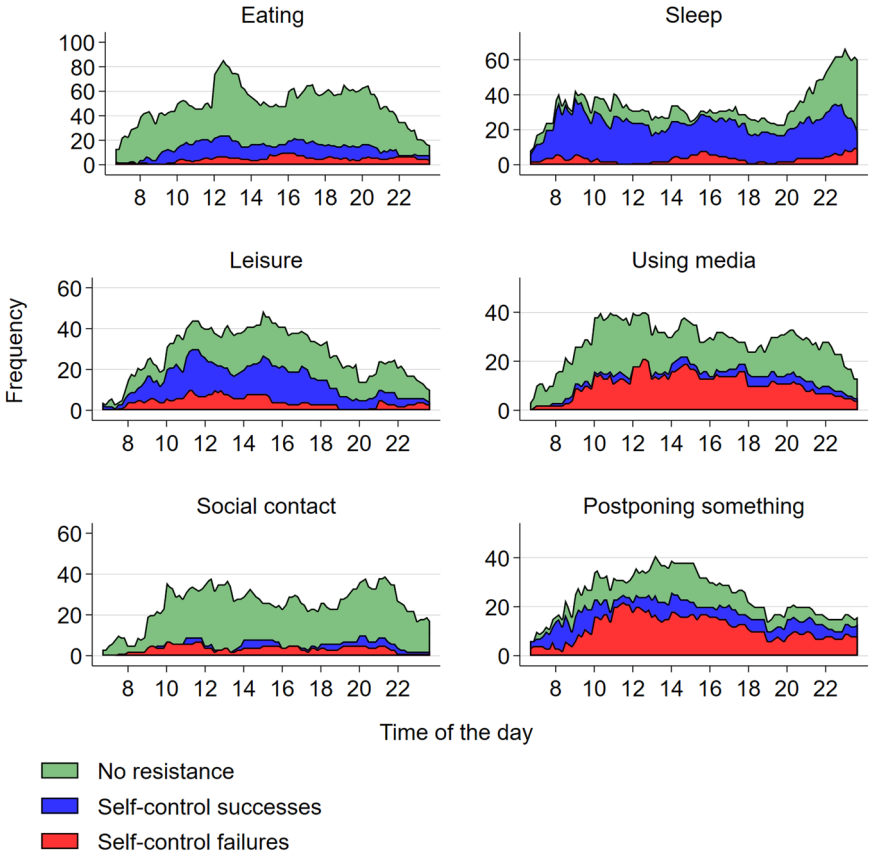

Fig. 4 Diurnal patterns of desires, resistance attempts, and self-control failures of the six most common desires over the day.

which are included as control variables, but whose inclusion also might lead to interesting insights. The error terms u_k, v_{jk}, and ϵ_{ijk} represent the structure of the data with error terms on the desire level, episode-level, and individual-level, respectively. In a second model, we run the same regression but focus only on those desires that participants had attempted to resist (illustrated by the upper half in Fig. 2). This allows us to identify the determinants of how successful people are once they attempt to resist which can be interpreted as the strength of willpower.

Table 2, model (1) shows the factors that explain whether a given desire led to a self-control failure. The model is based on the 3676 desires that were distributed across 2127 episodes and 248 participants. In terms of individual-specific determinants, it shows that age predicts a significantly lower probability that a desire leads to a self-control failure. It also shows that self-control failures are more likely when the desire is strong and when the conflict is strong, replicating earlier findings (Hofmann et al. 2012). We controlled for the desire

domains in the model (e.g., eating, sleeping, etc.), but do not present the 14 domain-specific coefficients in the table for presentation reasons. Instead, Fig. 5, panel A shows the predicted margins of a self-control failure by domain for the same model. As also suggested by Fig. 3, self-control failures are most likely in the domains "Postponing" and "Using media" and least likely in domains "work" and "non-alcoholic drinks".

Table 2, model (2) investigates the 1633 desires that participants attempted to resist (as seen in the upper half of Fig. 2). Such desires were present in 1149 episodes of 237 participants. This model presents factors that explain whether existing resistance attempts are successful (and not whether participants attempted to resist at all). Hence, the coefficients here can be interpreted as a form of reactive self-control where individuals aim to resist existing desires rather than aiming to reduce the occurrence of unwanted desires pre-actively (Kotabe and Hofmann 2015; Duckworth et al. 2016; Lades and Hofmann 2019). The model shows that strong desires are more difficult to resist, but conflict strength is not significantly predictive. Combined with the significant effect of conflict strength on self-control failures in model (1), the insignificance of conflict strength in model (2) suggests that stronger conflicts increase the probability that people attempt to resist the desire (i.e., increase the probability that people take the upper route in Fig. 2) but that the success of a resistance attempt is not dependent on the strength of the conflict. Moreover, people are more likely to fail in their resistance attempts when they are at home (compared to being at the university or at another place). In terms of individual-specific variables, females and participants from the UK are less likely to unsuccessfully attempt to resist an existing desire. Again, we do not present the domain-specific coefficients in the table but present the marginal effects of the domains in Fig. 5, panel B. As can also be seen in Fig. 3 when comparing self-control failures and self-control successes, once participants attempted to resist a desire, they were most likely to enact the desire nevertheless in the domains "Using media" "Postponing" and "Social contact" and least likely to enact desires in the domain "sleep".

Table 2 Multi-level logit regressions predicting the presence of a self-control failure (=enactment despite resistance attempt). Model (1) uses all desires and model (2) uses the desires that participants attempted to resist.

	(1)	(2)
Variables	**Self-control failure**	**Self-control failure**
Desire Strength	0.136**	0.532***
	(0.0665)	(0.103)
Conflict Strength	0.205***	−0.00128
	(0.0342)	(0.0485)
Location. Base = At home.		
At the University	0.0449	−0.743***
	(0.168)	(0.240)
Another place	−0.0989	−0.593**
	(0.177)	(0.252)
Social interaction	−0.103	0.174
	(0.149)	(0.212)
Hour of the day	0.00290	0.0113
	(0.0129)	(0.0184)
Age	−0.0341*	−0.0200
	(0.0195)	(0.0272)
Female	−0.155	−0.553**
	(0.198)	(0.275)
Marital status. Base = Single.		
Married	0.222	0.256
	(0.294)	(0.411)
Divorced	−0.212	0.0805
	(0.330)	(0.445)
College degree	0.252	0.351
	(0.201)	(0.276)
From the UK	−0.544***	−0.951***
	(0.185)	(0.261)
Annual income. Base = Below £10,000.		
£10,000-£24,999	0.196	0.297
	(0.247)	(0.334)
£25,000+	0.320	0.523
	(0.243)	(0.332)
Rather not say	−0.0659	0.0317
	(0.254)	(0.343)
Constant	−3.485***	−3.302***
	(0.625)	(0.881)
Desire type	Yes	Yes
Observations	3675	1633
Number of episodes	2127	1149
Number of individuals	248	237

Standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

The short-term benefits of self-control .failures As described in section 2, many researchers assume that self-control failures provide short-term benefits but are costly in the long-term. The DRM data does not provide information about the long-term benefits of satisfying short-term desires, but the presence of a resistance attempt suggests that individuals are themselves aware of long-term costs and would preferably not enact the short-term desire. We thus assume that individuals subjectively believe that enacting the short-term desire has costs and we ignore the question of whether the enactment has objectively long-term costs. The implicit assumption here is that resistance attempts,

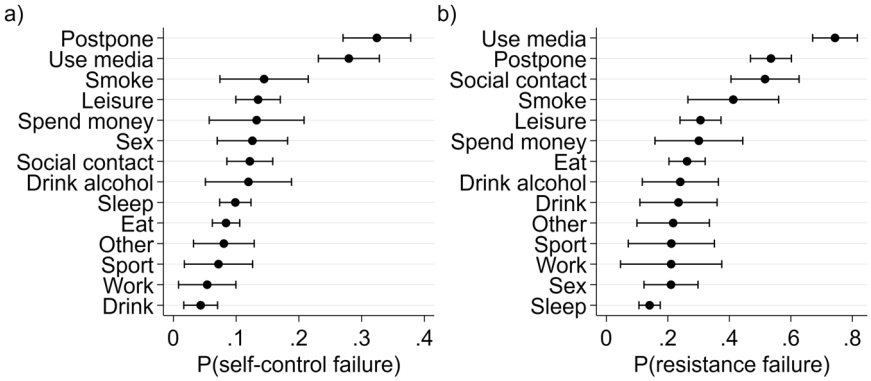

Fig. 5 Probability of self-control failures and unsuccessful resistance attempts by desire domain. a Probability that a desire becomes a self-control failure in the different domains. **b** Probability that a desire that is resisted is enacted and thus becomes a self-control failure in the different domains.

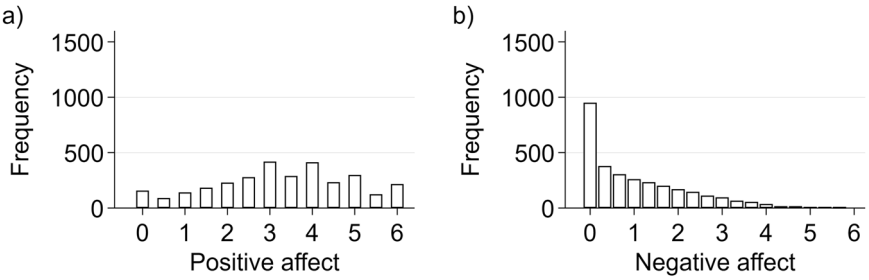

Fig. 6 Distributions of affect across all 3082 episodes. a Distribution of positive affect. **b** Distribution of negative effect.

i.e., the use of self-control, are driven by the desire to stick to long-term goals to avoid long-term costs.

To explore whether enacting short-term desires has short-term costs (or benefits) in terms of subjective well-being, we compare the positive and negative affect ratings in episodes with self-control failures with the positive and negative affect ratings in episodes without self-control failures. Figure 6 shows the distributions of this sub-section's dependent variables, i.e., positive affect (Panel A) and negative affect (Panel B) as measured in all 3082 episodes.

We test whether the SWB that participants experience in episodes with self-control failures differs from the SWB felt in episodes without self-control failures. Since the information we have on SWB (i.e., positive and negative affect) are on the episode-level, we specify regression models differentiating between two levels, accounting for episodes being nested in individuals. In linear multi-level models, we first regress the positive and the negative affect individual i experiences in episode j on the number of self-control failures in that episode, as summarized by

$$\boldsymbol{A}_{ij} = \beta_0 + \beta_1 NumSCF_{ij} + \beta_2 \boldsymbol{Epi}_{ij} + \beta_3 \boldsymbol{Ind}_{\boldsymbol{i}} + v_j + \epsilon_{ij}, \qquad (2)$$

where $\boldsymbol{A}_{ij}$ is a vector indicating the positive and the negative affect of individual i in episode j and $NumSCF_{ij}$ is the number of self-control failures in the same episode. The vectors $\boldsymbol{Epi}_{ij}$ and $\boldsymbol{Ind}_{\boldsymbol{i}}$ are control variables similar to the definitions for Eq. (1) with the exception that we also include a measure for life satisfaction as an individual-specific control variable in Eq. (2). Another difference between Eqs. (1) and (2) is that the latter does not include control variables for the desire domains, because this would lead to multicollinearity with the variable that indicates the number of desires.

The results of this regression model, presented in column 1 of Table 3, show that the positive affect in episodes with one ($b = -0.273$; $p < 0.001$) or two or more ($b = -0.4333$; $p = 0.004$) self-control failures is significantly lower than in episodes without self-control failures. The predicted positive affect ratings in episodes with no self-control failure, one self-control failure, and two or more self-control failures are 4.32 (SE = 0.06), 4.04(SE = 0.09), and 3.88 (SE = 0.16), respectively.

This approach, however, underestimates the strength of the negative association between self-control failures and positive affect. Previous research shows that people feel better in episodes in which they enact desires compared to episodes in which they do not enact desires (e.g., Hofmann et al. 2013) and whenever there is a self-control failure a desire is by definition enacted. Hence, in another regression model, we control for the number of desire enactments in episode j ($NumEnac_{ij}$) to neutralize the association between enactment and affect. We also control for the number of desires people had in episode j ($NumDes_{ij}$) to avoid measuring the association between having more or less desires and SWB. Column 2 in Table 3 shows that the associations between the number of self-control failures and positive affect indeed become stronger. These coefficients compare the positive affect of episodes with and without self-control failures while holding constant the number of desires and the number of enacted desires in the episode. The predicted positive affect ratings in episodes with no self-control failure, one self-control failure, and two or more self-control failures based on model 2 are now 4.35 (SE = 0.06), 3.90 (SE = 0.1), and 3.54 (SE = 0.17), respectively.

Columns 3 and 4 in Table 3 show that the patters are reversed for negative effect, albeit somewhat weaker. For example, based on model 4 the predicted negative affect ratings for no self-control failure, one

Table 3 Multi-level linear models predicting affect by presence of self-control failures.

Variables	(1) Positive affect	(2) Positive affect	(3) Negative affect	(4) Negative affect
Number of self-control failures. Base = Zero.				
One	−0.273***	−0.443***	0.149**	0.186***
	(0.0803)	(0.0858)	(0.0591)	(0.0634)
Two or more	−0.433***	−0.807***	0.256**	0.395***
	(0.151)	(0.163)	(0.111)	(0.121)
Number of enacted desires. Base = Zero.				
One		0.532***		−0.258***
		(0.0702)		(0.0519)
Two		0.850***		−0.600***
		(0.111)		(0.0819)
Three or more		1.371***		−0.848***
		(0.155)		(0.115)
Number of desires. Base = Zero.				
One		−0.565***		0.331***
		(0.0690)		(0.0510)
Two		−0.590***		0.501***
		(0.0957)		(0.0708)
Three		−1.023***		0.855***
		(0.117)		(0.0866)
Four or more		−1.186***		0.971***
		(0.179)		(0.133)
Hour of the day	0.0410***	0.0404***	−0.0263***	−0.0232***
	(0.00479)	(0.00479)	(0.00353)	(0.00354)
Location. Base = At home.				
At University	−0.353***	−0.283***	0.145***	0.107**
	(0.0674)	(0.0664)	(0.0497)	(0.0492)
At another place	0.0532	0.0796	−0.0176	−0.0211
	(0.0629)	(0.0619)	(0.0464)	(0.0458)
Social interaction	0.440***	0.404***	0.0346	0.0517
	(0.0552)	(0.0543)	(0.0406)	(0.0401)
Life satisfaction	0.0592***	0.0560***	−0.0250***	−0.0231***
	(0.00833)	(0.00829)	(0.00695)	(0.00678)
Constant	2.496***	2.823***	3.049***	2.756***
	(0.365)	(0.364)	(0.301)	(0.295)
Observations	3082	3082	3082	3082
Number of groups	248	248	248	248

All models are controlled for age, gender, marital status, college-education, country of origin, and income. Standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

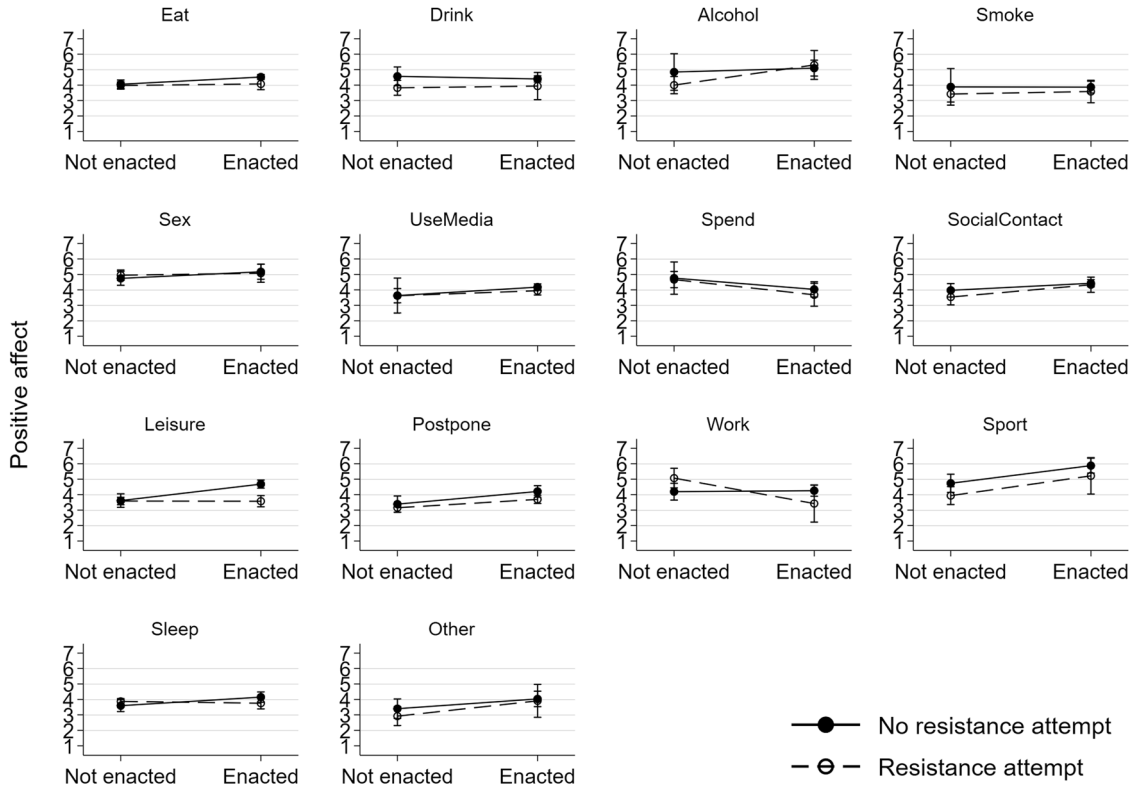

Fig. 7 Positive affect in episodes in which the respected desire was present and enacted (vs not enacted) and resisted (vs not resisted). The open circles on the right-hand-side represent the positive affect in episodes with self-control failures.

self-control failure, and two or more self-control failures are 2.07 (SE = 0.05), 2.25 (SE = 0.07), and 2.565 (SE = 0.13), respectively.

Models (2) and (4) also show that the number of enacted desires is positively associated with positive affect and negatively associated with negative affect (holding constant the number of desires present in the episode). This pattern replicates findings from Hofmann et al. (2013),[11] and suggests that preference satisfaction can indeed be a source of well-being. However, comparing the coefficients for desire enactment with the coefficients for self-control failures shows that they are rather similar in strength with reversed signs. This suggests that the enactment of a desire increases positive affect, but only in the absence of a self-control failure.

A more detailed analysis of the associations between desire enactment and experienced SWB (with and without resistance attempts) is precluded by the data structure of the above analysis, because the dependent affect variable is measured on the episode-level and the desire-related independent variables are measured on the higher desire level. Multiple desires, resistance attempts, and self-control failures can be present in the same episode. However, we can analyse the data for each desire domain independently, which allows us to present one-to-one associations between self-control failures in a specific domain and contemporaneous experienced SWB. This is possible because information about the same desire domain cannot appear more than once in each episode. To do so, we estimate models such as

$$\begin{aligned} \boldsymbol{A}_{ij} = {} & \beta_0 + \beta_1 EatEnact_{ij} + \beta_2 EatResist_{ij} + \beta_3 EatEnact * EatResist \\ & + \beta_4 \boldsymbol{Epi}_{ij} + \beta_5 \boldsymbol{Ind}_{\boldsymbol{i}} + \nu_j + \epsilon_{ij} \end{aligned} \quad (3)$$

for the subset of episodes in which a desire to eat was present. In Eq. (3), $EatEnact_{ij}$ indicates whether individual i enacted a desire to eat in episode j, $EatResist_{ij}$ is a dummy indicating whether the individual resisted enacting a desire to eat, and the other variables are as above. The interaction between $EatEnact_{ij}$ and $EatResist_{ij}$ represents a self-control failure. We run these regressions 14 times for each desire type separately and predict the affect for all four combinations of resistance attempts (yes or no) and enactments (yes or no). The resulting four datapoints for each desire domain are presented in Fig. 7 in 14 panels. These four datapoints correspond to the for boxes on the right in Fig. 2. Self-control failures are represented by the open circles on the right-hand side of each panel. While this approach allows us to present the one-to-one associations between self-control failures and experienced SWB, the number of observations is reduced a lot, and results need to be interpreted particularly carefully.

The analysis illustrated in Fig. 7 suggests that there might be some situations in which the satisfaction of a short-term desire that is a self-control failure does not provide short-term benefits. For example, in episodes in which people had a desire for leisure, positive affect is higher when they enacted that desire without a resistance attempt ($b = 1.13$, $p < 0.001$). However, when comparing episodes in which people had a desire for leisure and attempted to resist it, positive affect is not significantly different across episodes with and without enactments of the desire for leisure ($b = -0.022$; $p = 0.918$). The difference between these two associations represented by the slopes of the two lines in the leisure panel in Fig. 7 is statistically significant ($b = -1.077$; $p = 0.001$). The difference between the slopes of the two lines is also statistically significant in the domain sleep ($b = -0.669$; $p = 0.035$) and work ($b = -1.708$; $p = 0.027$).

We can also compare the positive affect ratings across episodes in which a not-resisted desire is enacted and episodes in which a resisted desire is enacted. This analysis suggests that over most desire domains, positive affect ratings were lower when desires were enacted despite a resistance attempt (i.e., self-control failures) than when desires were enacted without a resistance attempt. This difference is statistically significant in the domains

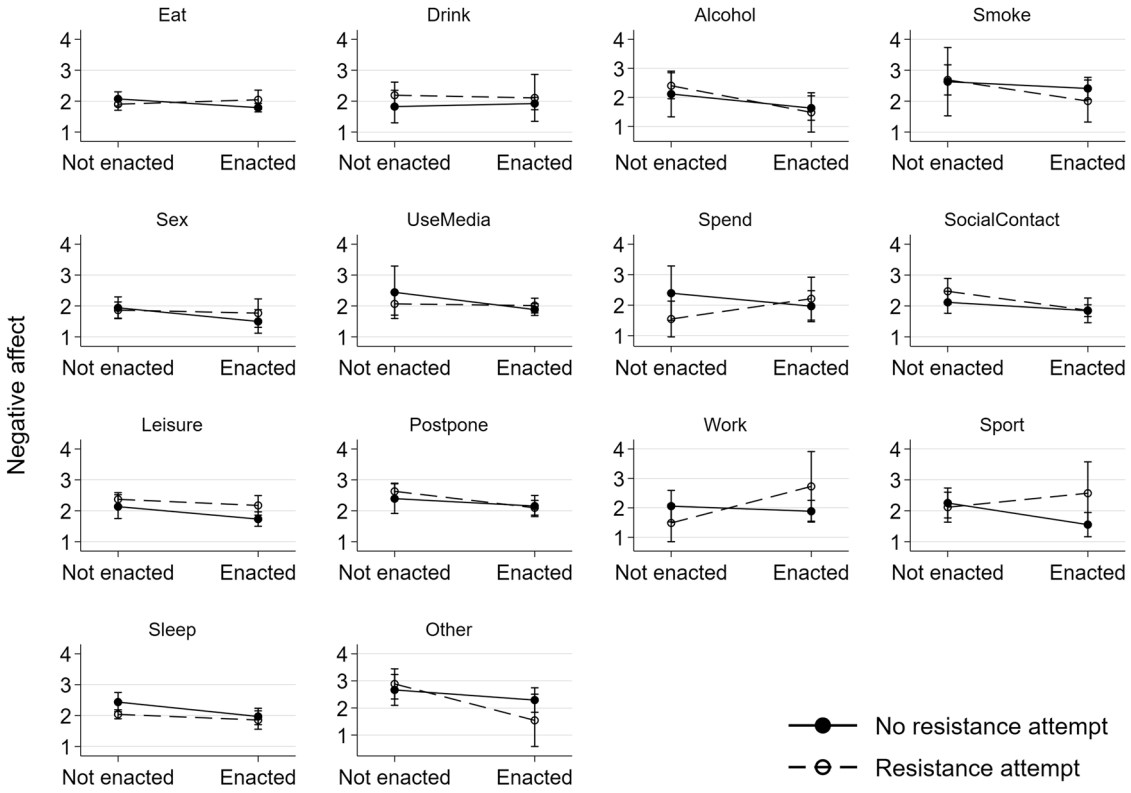

Fig. 8 Negative affect in episodes in which the respected desire was present and enacted (vs not enacted) and resisted (vs not resisted). The open circles on the right-hand-side represent the positive affect in episodes with self-control failures.

eating ($b = -0.45$; $p = 0.021$), leisure ($b = -1.11$; $p < 0.001$), and postponing ($b = -0.51$; $p = 0.019$). For example, given that people enacted a desire for leisure, their contemporaneous positive affect was 3.58 (SE = 0.18) if they had resisted the desire and 4.68 (SE = 0.13) if they had not resisted it. Figure 8 presents the same analysis for negative affect and similar exploratory observations are possible

Discussion

The previous section showed that it is possible to measure self-acknowledged self-control failures and that it is also possible to measure experiential correlates of self-control failures which can be interpreted as short-term welfare consequences. This section discusses whether data like this can and should be used to inform (behavioural) policy-making.

Notes of .caution Let us offer a few warnings. The approach to measure everyday self-acknowledged self-control failures presented here has some limitations and we view the study as an empirical demonstration of the possibility to measure self-control failures and contemporaneous affective states. We would not yet be willing to derive strong and general welfare implications from this particular study for a number of reasons. First, the data we present in section 4 is based on a sample of students and the patterns we find might be specific to that population. Second, policies tend to be domain-specific, and the study presented in this paper deals with self-control failures more generally across multiple domains. If policy-makers are interested in whether or not self-control failures occur in one particular domain of life (e.g., unhealthy eating), it would be essential to conduct domain-specific studies asking the target population whether they had specific food-related self-control failures. Such domain-specific studies are also beneficial from a statistical point of view as it will be possible to calculate one-to-one associations between self-control failures and subjective well-being as presented in Fig. 7 with larger sample sizes. Our analysis of the associations between affect ratings and self-control failures as presented in Table 3 is limited because the dependent variables (positive and negative affect) are measured on the episode-level and desire-related independent variables (including self-control failures) are measured on the desire level. Analysing the data at the desire level would yield biased estimates of the standard errors and inflated Type I error rates (Foster-Johnson and Kromrey 2018). Hence, we used less informative summary variables such as "number of desires" that can be interpreted at the episode-level. However, these summary variables do not provide a one-to-one association between desire and affect measures and using them can lead to inaccurate standard error estimations and biased regression parameters (Foster-Johnson and Kromrey 2018). Another limitation is that the day reconstruction method is a relatively new measurement tool that is potentially subject to survey measurement problems, including recall bias, social desirability bias, and certain individual-specific response styles. Further methodological work is required to improve the measures, test their reliability and validity, and to better understand how responses are affected by question wording, respondent understanding, and other features of the survey design. Additionally, the results we presented are correlational and people might have selected into their environments to some extent based on unobserved or unobservable characteristics. Moreover, it is possible that affect is a determinant of self-control failures (rather than the other way around) as for example suggested by Tice et al. (2007). If this was the case, the association between self-control failures and affect would be biased upwards and we would overestimate the welfare costs of self-control failures. The method would need to be

integrated into field studies in order to gather causal evidence in future work before informing policy. Our study is exploratory as we did not pre-register the study and did not correct for multiple hypothesis testing. Finally, we focus on the *better off, as judged by themselves* criterion to evaluate policies and this is an approach that highlights welfare. There are other normative criteria to evaluate policies, such as autonomy, freedom, or negative externalities that we ignore here. Overall, we view this study as a contribution to the discussion within behavioural public policy about how and whether data about self-control failures in everyday life can and should inform policy. We do not claim that this discussion has yet concluded but hope to make substantial and interesting contributions.

The use of experiential well-being data for policy .evaluation We measured the short-term welfare effects of self-control failures in terms of subjective well-being (SWB). In the last decade, SWB measures have received increasing attention in economics and psychology (Frey and Stutzer 2002; Boyce et al. 2010; Knabe et al. 2010; Layard 2011; Laffan et al. 2024), and it is becoming more and more common to evaluate policies based on their effects on SWB (Frey and Stutzer 2002; Kahneman and Sugden 2005; Dolan and White 2007; Dolan and Kahneman 2008; Layard 2011; Odermatt and Stutzer 2017).[12] SWB can be defined and measured in at least two ways: as a cognitive construct in terms of life satisfaction and as an experience in terms of momentary happiness. We use the latter definition as we focus on momentary correlates of behaviours. Some researchers have argued that this momentary definition is less prone to cognitive biases as it is relatively easy for people to answer questions about their feelings (Kahneman and Sugden 2005). However, it is clear that there are potentially disadvantages to using momentary measures over general measures also, in particular the difficulty in using momentary measures as indices of overall life evaluation. As we argue below, momentary measures may be particularly suitable for evaluating welfare changes in life domains with high degree of regularized behaviour.

Implications for behavioural welfare .economics Notwithstanding the notes of caution presented above and the limits of subjective well-being measures, we argue that data from everyday life may contribute in several ways in the future to behavioural welfare economics and public policy-making more broadly. In fact, data from everyday life might help with a big challenge that behavioural economic findings about bounded rationality pose for the neoclassical economic welfare analysis. Welfare economics is based on the revealed preference approach which assumes that desire fulfilment is the origin of welfare and that individuals rationally make decisions that maximize their welfare (Hausman 2012). However, if people do not always act rationally, as shown by behavioural economists, choices do not always reveal the preferences whose satisfaction maximizes welfare.[13] Accordingly, welfare economists who acknowledge behavioural economic findings of bounded rationality need to develop alternative welfare criteria to measure whether a policy makes people better off, as judged by themselves (for a recent summary, see Sugden 2018). Data from everyday life might inform such alternative behavioural welfare economic approaches.

Distinguishing between short-term and long-term preferences Data from everyday life as presented in this paper allows us to go beyond observed choices to identify what increases people's welfare. We can additionally measure the decision-making process that led to these choices. Information about how choices came about can be helpful in various ways. For example, a policy-maker who observes that people eat unhealthily does not know whether this behaviour is in line with people's goals or whether it results from a self-control failure. Data as presented here allow to make that distinction. This is relevant from the perspective of several authors who distinguish between short-run and long-run preferences and put more normative weight on the long-run preference as these are more persistent and dynamically consistent (O'Donoghue and Rabin 1999; Ericson and Laibson 2019). In line with this argument, the fulfilment of non-conflicting and not resisted desires could be given more normative weight than the satisfaction of desires that conflict with higher-order goals and that people attempted to resist. That in our data in some desire domains the satisfaction of unresisted desires seems to be associated with higher experienced SWB than the satisfaction of resisted desires provides some support for the higher normative weight of unresisted desire satisfaction.

Purified and correctly informed .preferences Information about decision-making mechanisms might also be helpful to "purify preferences", i.e., to isolate the true, normatively relevant preferences from the influence of decision-making biases and weakness of will on choice. Data from everyday life may thus provide an additional way to clean preferences, and this approach could be added to the list of strategies to reveal true preferences as discussed by Beshears et al. (2008). For example, we measured the extent to which short-term desires are conflicting with higher-order goals and a person's long-run stated objectives. One could use this measure of conflict strength to reduce the normative weight put on the satisfaction of conflicting desires. However, even with information about the decision-making processes that underlie choices, preference purification will be far from straightforward. Moreover, a number of researchers argue that the preference purification approach is based on a mistaken view of human behaviour as it presupposes a rational core with true preferences that might not exist (Infante et al. 2016; Sugden 2018).

An alternative to preference purification is to identify the subset of "correctly informed choices" for which good evidence exists that these choices are not mistaken, and to use only these correctly informed choices as a welfare benchmark (Bernheim 2016). It is possible to use the DRM to identify the subset of choices that are in line with both short-term and long-term preferences (i.e., those desires that are not resisted and enacted, Box 4 in Fig. 2) and these could be interpreted as correctly informed choices. Future work should measure whether the enactment of these unresisted desires causally leads to stronger improvements in subjective well-being than the enactment of desires that individual attempt to resist do.

Paternalism without trade-offs? Data on everyday self-control failures and their associations with SWB may also be used to identify behaviours that neither increase short-term nor long-term welfare. Table 3 suggests that the associations between desire satisfaction and experienced SWB at the same time are attenuated when the desire satisfaction is the result of a self-control failure (i.e., when people had attempted to resist enacting the desire). This pattern is also visible in some desire domains as shown in Figs. 7 and 8. If behaviours exist that do not increase short-term welfare (measured as experienced SWB) and at the same time are detrimental to long-term welfare (in the sense that they are not in line with long-term goals), these behaviours should be prime targets for policy intervention. Data from everyday life could help to identify these instances and policy-attention could be focused towards the domains where interventions increase long-term goals without reducing short-term welfare. Such policies would be warranted based on the *better off, as judged by themselves* criterion. This argument is in line with Gruber and Mullainathan

(2006) who show that marginal smokers benefit from the imposition of smoking taxes both in the short-term (their short-run well-being is improved) and the long-term (they are able to quit smoking). The approach presented here could potentially allow for the identification of other examples like this.

Desire fulfilment theories and the absence of .desire Table 3 also shows that desire enactment is positively associated with positive affect and negatively associated with negative affect. This is also evident in some desire domains in Figs. 7 and 8. This strengthens the argument that desire fulfilment is a meaningful ingredient of individual welfare and indicates that policies that help people to satisfy their desires might increase welfare already in the short-run. At the same time, Table 3 shows that having more desires is associated with lower positive affect and higher negative affect. This suggests that desire fulfilment might not be the only way to higher SWB, and that an alternative way to increase welfare might be to reduce the amount of desires in everyday life. More generally, the extent to which people can shape their internal motivations and selection of external environments to improve meaning and capacity for enjoyment is a key theme in a number of psychological literatures and the implications for economics remain underexplored despite the pioneering contribution of Scitovsky (1976) on this question.

Conclusion

This paper showed that it is possible to quantify the prevalence of self-acknowledged self-control failures in daily life and to measure the subjective short-term welfare effects of self-control failures in terms of associations with experienced well-being. The paper also argues that this type of data can be used to inform behavioural welfare economics and public policy more generally. In terms of future work, it will be necessary to conduct methodological research to test the reliability and validity of the data provided by tools such as the day reconstruction method. It will also be necessary to conduct domain-specific studies that identify everyday self-control failures in specific domains in those populations that might be targeted by public policies. Finally, everyday measures should be incorporated in experimental designs in order to measure the causal effects of experiments and policy-changes on self-control failures and associated subjective well-being in everyday life.

Data availability

The data and Stata analysis code are available on the website of the Open Science Framework: https://osf.io/x3de7/.

Received: 13 May 2024; Accepted: 23 September 2024;
Published online: 08 October 2024

Notes

1 The claim that self-control is a limited resource that can become depleted when used which can lead to self-control failures is very popular but does not seem to be true (Hagger et al. 2016; Vohs et al. 2021).
2 In their review of the behavioural economic literature on inter-temporal choice, Ericson and Laibson (2019) suggest that it is an empirical regularity that "Many public policies seem paternalistic and aimed at addressing present-focused behaviour."
3 There are a number of additional arguments in favour of and against (soft) paternalism as a response to behavioural economic insights. These are the subject of intensive debate (Hausman and Welch 2010; Conly 2012; Grüne-Yanoff 2012; Rebonato 2012; Sunstein 2016; Schmidt and Engelen 2020; Lades and Delaney, 2022), which we do not review here.
4 Sugden (2017) summarizes his main argument as follows: "*My claim is that self-acknowledged self-control problems are a lot less common than many behavioural economists or policy-makers seem to think. Even if behavioural economists or policymakers feel confident that people's lifestyle choices are based on some kind of error, they should not jump to the conclusion that the error is a self-acknowledged failure of self-control or that ..,it is what those people themselves would call an error*" (p. 22). In a response to Sugden, Sunstein (2018) presents some preliminary data based on a single survey question that suggests that 70 per cent of the 200 study participants believe that they have some kind of self-control problems.
5 Duckworth et al. (2018) present an alternative classification and differentiate between three classes of self-control models: (i) models of multiple sequential selves with dynamically inconsistent preferences (Ainslie 1975), (ii) models with multiple coexisting selves (Thaler and Shefrin 1981; e.g., Fudenberg and Levine 2006), and (iii) multiple-attribute models (e.g., Berkman et al. 2017).
6 This is true for "leisure goods" (that have immediate benefits and future costs) and for "investment goods" (that have immediate costs and future benefits) the pattern is reversed (see DellaVigna 2018). However, one could re-interpret investment goods as "not spending effort" in the short-term.
7 Jeremy Bentham starts his Principles of Morals and Legislation with the following sentences: "Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do." (Bentham 1789)
8 Also when interpreting SWB cognitively, more self-control is linearly correlated with more life satisfaction (Wiese et al. 2018).
9 The data of 145 of these participants is used in Delaney and Lades (2017) to show the associations between present bias (as measured in incentivized delay discounting tasks) and everyday self-control failures (as measured in the way described above).
10 People might also proactively organise their lives to avoid problematic desires in the first place. Since the series of questions in our survey starts with existing desires, we do not have data for these proactive self-control successes.
11 Hofmann et al. (2014) find that desire satisfaction predicts higher positive affect but only in the absence of resistance attempts. Once individuals attempt to resist enacting the desire, its satisfaction does not predict higher positive affect. They argue that self-referential feelings of guilt and shame can explain this.
12 Using subjective measures for utility is by no means a new idea in economics. For example, Edgeworth described a "hedonimeter" that would allow economists to directly measure utility as discussed in Colander (2007). However, attempts to incorporate direct subjective measures of well-being and introspective data on decision processes fell out of favour in economics in the late 19th century and the tradition of "sensory utilitarianism" (Witt 2005) gave way to formal utilitarian accounts based on axioms of rational choice and economics became a science of choice (Kahneman 1999; Read 2007). It seems to us that measurement developments such as those presented in the present paper suggest that economists should not disregard the possibility to measure utility directly.
13 Even if individuals do not have all relevant information at their disposal, choices do not necessarily reveal true preferences (Hausman 2000).

References

Ainslie G (1975) Specious reward: a behavioral theory of impulsiveness and impulse control. Psychol Bull 82(4):463–496. https://doi.org/10.1037/h0076860
Baumeister RF (2002) Yielding to temptation: self-control failure, impulsive purchasing, and consumer behavior. J Consum Res 28(04):670–676. https://doi.org/10.1086/338209
Baumeister RF, Tierney J (2011) Willpower: rediscovering the greatest human strength. New York, NY, The Penguin Press
Bentham, J. (1789). From An Introduction to the Principles of Morals and Legislation. Printed in the Year 1780, and Now First Published. In Literature and Philosophy in Nineteenth Century British Culture, Routledge, pp 261–268
Berg N, Gigerenzer G (2010) As-if behavioral economics: neoclassical economics in disguise? Hist Econ Ideas 18(1):133–165
Berkman ET, Hutcherson CA, Livingston JL et al. (2017) Self-control as value-based choice. Curr Dir Psychol Sci 26(5):422–428. https://doi.org/10.1177/0963721417704394
Bernheim BD (2016) The good, the bad, and the ugly: a unified approach to behavioral welfare economics. J Benefit-Cost Anal 7(1):12–68
Beshears J, Choi JJ, Laibson D et al. (2008) How are preferences revealed? J Public Econ 92(8):1787–1794
Binder M, Lades LK (2015) Autonomy-enhancing paternalism. Kyklos 68(1):3–27
Boyce CJ, Brown GDA, Moore SC (2010) Money and happiness: rank of income, not income, affects life satisfaction. Psychol Sci 21(4):471–475. https://doi.org/10.1177/0956797610362671
Colander D (2007) Retrospectives: edgeworth's hedonimeter and the quest to measure utility. J Econ Perspect 21(2):215–225. https://doi.org/10.1257/jep.21.2.215
Conly S (2012) Against autonomy: justifying coercive paternalism. Cambridge University Press, Cambridge

Delaney L, Lades LK (2017) Present bias and everyday self-control failures: a day reconstruction study. J Behav Decis Mak 30(5):1157–1167. https://doi.org/10.1002/bdm.2031
DellaVigna S (2018) Structural behavioral economics. In: Bernheim BD, DellaVigna S, Laibson D (eds.) Handbook of behavioral economics: applications and foundations 1. Handbook of behavioral economics - foundations and applications 1, vol 1. North-Holland pp. 613–723. https://doi.org/10.1016/bs.hesbe.2018.07.005
Dolan P, Kahneman D (2008) Interpretations of utility and their implications for the valuation of health. Econ J 118(525):215–234. https://doi.org/10.1111/j.1468-0297.2007.02110.x
Dolan P, White MP (2007) How can measures of subjective well-being be used to inform public policy? Perspect Psychol Sci 2(1):71–85. https://doi.org/10.1111/j.1745-6916.2007.00030.x
Doyle O, Delaney L, O'Farrelly C et al. (2017) Can early intervention improve maternal well-being? Evidence from a randomized controlled trial. PLoS ONE 12(1):1–25. https://doi.org/10.1371/journal.pone.0169829
Duckworth AL, Gendler TS, Gross JJ (2016) Situational strategies for self-control. Perspect Psychol Sci 11(1):35–55. https://doi.org/10.1177/1745691615623247
Duckworth AL, Milkman KL, Laibson D (2018) Beyond willpower: strategies for reducing failures of self-control. Psychol Sci Public Interest 19(3):102–129. https://doi.org/10.1177/1529100618821893
Ericson KM, Laibson D (2019) Intertemporal choice. In: Bernheim BD, DellaVigna S, Laibson D (eds.) Handbook of behavioral economics: applications and foundations 1. vol 2. North-Holland pp. 1–67. https://doi.org/10.1016/bs.hesbe.2018.12.001
Foster-Johnson L, Kromrey JD (2018) Predicting group-level outcome variables: an empirical comparison of analysis strategies. Behav Res 50(6):2461–2479. https://doi.org/10.3758/s13428-018-1025-8
Frederick S, Loewenstein G, O'Donoghue T (2002) Time discounting and time preference: a critical review. J Econ Lit 40(2):351–401. https://doi.org/10.1257/002205102320161311
Frey BS, Stutzer A (2002) What can economists learn from happiness research? J Econ Lit 40(2):402–435. https://doi.org/10.1257/002205102320161320
Fudenberg D, Levine DK (2006) A dual-self model of impulse control. Am Econ Rev 96(5):1449–1476. https://doi.org/10.1257/aer.96.5.1449
Gruber J, Mullainathan S (2006) Do cigarette taxes make smokers happier? In: Happiness and public policy. Palgrave Macmillan, UK pp. 109–146
Grubiak KP, Isoni A, Sugden R et al. (2024) Taking the new year's resolution test seriously: eliciting individuals' judgements about self-control and spontaneity. Behav Public Policy 8(1):1–23. https://doi.org/10.1017/bpp.2021.41
Grüne-Yanoff T (2012) Old wine in new casks: libertarian paternalism still violates liberal principles. Soc Choice Welf 38(4):635–645
Gul F, Pesendorfer W (2001) Temptation and self-control. Econometrica 69(6):1403–1435. https://doi.org/10.1111/1468-0262.00252
Hagger MS, Chatzisarantis NLD, Alberts H et al. (2016) A multilab preregistered replication of the ego-depletion effect. Perspect Psychol Sci 11(4):546–573. https://doi.org/10.1177/1745691616652873
Hausman DM (2000) Revealed preference, belief, and game theory. Econ Philos 16(1):99–115. https://doi.org/10.1017/S0266267100000158
Hausman DM (2012) Preference, value, choice, and welfare. Cambridge University Press, Cambridge, UK
Hausman DM, Welch B (2010) Debate: to nudge or not to nudge. J Political Philos 18(1):123–136
Hoch SJ, Loewenstein G (1991) Time-inconsistent preferences and consumer self-control. J Consum Res 17(4):492–507. https://doi.org/10.1086/208573
Hofmann W, Baumeister RF, Förster G et al. (2012) Everyday temptations: an experience sampling study of desire, conflict, and self-control. J Pers Soc Psychol 102(6):1318–1335. https://doi.org/10.1037/a0026545
Hofmann W, Kotabe H, Luhmann M (2013) The spoiled pleasure of giving in to temptation. Motiv Emot 37(4):733–742. https://doi.org/10.1007/s11031-013-9355-4
Hofmann W, Luhmann M, Fisher RR et al. (2014) Yes, but are they happy? effects of trait self-control on affective well-being and life satisfaction. J Personal 82(4):265–277
Infante G, Lecouteux G, Sugden R (2016) Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics. J Econ Methodol 23(1):1–25
Inzlicht M, Schmeichel BJ, Macrae CN (2014) Why self-control seems (but may not be) limited. Trends Cogn Sci 18:127–133. https://doi.org/10.1016/j.tics.2013.12.009
Kahneman, D. (1999). Objective happiness. In: Kahneman, D., Diener, E., & Schwarz, N. (eds) Wellbeing: The foundations of hedonic psychology, New York, Russell Sage Foundation, pp 3–25
Kahneman D, Krueger AB, Schkade DA et al. (2004) A survey method for characterizing daily life experience: the day reconstruction method. Science 306(5702):1776–1780. https://doi.org/10.1126/science.1103572
Kahneman D, Sugden R (2005) Experienced utility as a standard of policy evaluation. Environ Resour Econ 32(1):161–181. https://doi.org/10.1007/s10640-005-6032-4
Knabe A, Rätzel S, Schöb R et al. (2010) Dissatisfied with life but having a good day: time-use and well-being of the unemployed. Econ J 120(547):867–889. https://doi.org/10.1111/j.1468-0297.2009.02347.x
Kotabe HP, Hofmann W (2015) On integrating the components of self-control. Perspect Psychol Sci 10(5):618–638. https://doi.org/10.1177/1745691615593382
Kurzban R, Duckworth A, Kable JW et al. (2013) An opportunity cost model of subjective effort and task performance. Behav Brain Sci 36(6):661–679. https://doi.org/10.1017/S0140525X12003196
Lades LK, Delaney L (2022) Nudge FORGOOD. Behav Public Policy 6(1):75–94
Lades LK, Egan M, Delaney L et al. (2017) Childhood self-control and adult pension participation. Econom Lett 161:102–104. https://doi.org/10.1016/j.econlet.2017.09.025
Lades LK, Hofmann W (2019) Temptation, self-control, and inter-temporal choice. J Bioecon 21(1):47–70. https://doi.org/10.1007/s10818-019-09284-2
Lades LK, Martin L, Delaney L (2022) Informing behavioural policies with data from everyday life. Behav Public Policy 6(2):172–190. https://doi.org/10.1017/bpp.2018.37
Laffan K, Sunstein C, Dolan P (2024) Facing it: assessing the immediate emotional impacts of calorie labelling using automatic facial coding. Behav Public Policy 8(3):572–589. https://doi.org/10.1017/bpp.2021.32
Laibson D (1997) Golden eggs and hyperbolic discounting. Q J Econ 112(2):443–477. https://doi.org/10.1162/003355397555253
Layard R (2011) Happiness: lessons from a new science. Penguin, UK
Loewenstein G (1996) Out of control: visceral influences on behavior. Organ Behav Hum Decis Process 65(3):272–292. https://doi.org/10.1006/obhd.1996.0028
Loewenstein G (2018) Self-control and its discontents: a commentary on Duckworth, milkman, and Laibson. Psychol Sci Public Interest 19(3):95–101. https://doi.org/10.1177/1529100619828401
Milkman KL, Rogers T, Bazerman MH (2008) Harnessing our inner angels and demons: what we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making. Perspect Psychol Sci 3(4):324–338. https://doi.org/10.1111/j.1745-6924.2008.00083.x
Moffitt TE, Arseneault L, Belsky D et al. (2011) A gradient of childhood self-control predicts health, wealth, and public safety. Proc Natl Acad Sci USA 108(7):2693–2698. https://doi.org/10.1073/pnas.1010076108
Odermatt R, Stutzer A (2017) Subjective well-being and public policy. In: Diener E, Oishi S, Tay L (eds.) Handbook of well-being. DEF Publishers, Salt Lake City
O'Donoghue T, Rabin M (1999) Doing it now or later. Am Econ Rev 89(1):103–124. https://doi.org/10.1257/aer.89.1.103
Read D (2006) Which side are you on? the ethics of self-command. J Econ Psychol 27(5):681–693. https://doi.org/10.1016/j.joep.2006.04.005
Read D (2007) Experienced utility: utility theory from Jeremy Bentham to Daniel Kahneman. Think Reason 13(1):45–61
Read D (2008) Personal utilitarianism: multiple selves and their search for the good life
Rebonato R (2012) Taking liberties: a critical examination of libertarian paternalism. New York, NY, Palgrave Macmillan
Rizzo MJ (2016) Behavioral economics and deficient willpower: searching for akrasia. Georget J Law Public Policy 14:789
Rizzo MJ, Whitman DG (2008) The knowledge problem of new paternalism. SSRN Scholarly Paper ID 1310732. Social Science Research Network, Rochester, NY
Schmidt AT, Engelen B (2020) The ethics of nudging: an overview. Philos Compass 15(4):e12658. https://doi.org/10.1111/phc3.12658
Scitovsky T (1976) The joyless economy: the psychology of human satisfaction. Oxford University Press, USA
Sugden R (2008) Why incoherent preferences do not justify paternalism. Const Political Econ 19(3):226–248. https://doi.org/10.1007/s10602-008-9043-7
Sugden R (2017) Do people really want to be nudged towards healthy lifestyles? Int Rev Econ 64(2):113–123. https://doi.org/10.1007/s12232-016-0264-1
Sugden R (2018) The community of advantage: a behavioural economist's defence of the market. Oxford, Oxford University Press
Sunstein CR (2016) The ethics of influence: government in the age of behavioral science. Cambridge University Press, New York, NY
Sunstein CR (2018) "Better off, as judged by themselves": a comment on evaluating nudges. Int Rev Econ 65(1):1–8. https://doi.org/10.1007/s12232-017-0280-9
Sunstein CR (2019) On freedom. Princeton, NJ, Princeton University Press
Thaler RH, Shefrin HM (1981) An economic theory of self-control. J Polit Econ 89(2):392–406. https://doi.org/10.1086/260971
Thaler RH, Sunstein CR (2008) Nudge: Improving decisions about health, wealth, and happiness. Newhaven, Yale University Press
Tice DM, Baumeister RF, Shmueli D et al. (2007) Restoring the self: positive affect helps improve self-regulation following ego depletion. J Exp Soc Psychol 43(3):379–384. https://doi.org/10.1016/j.jesp.2006.05.007
Vohs KD, Schmeichel BJ, Lohmann S et al. (2021) A multisite preregistered paradigmatic test of the ego-depletion effect. Psychol Sci 32(10):1566–1581. https://doi.org/10.1177/0956797621989733

Wiese CW, Tay L, Duckworth AL et al. (2018) Too much of a good thing? exploring the inverted-u relationship between self-control and happiness. J Pers 86(3):380–396. https://doi.org/10.1111/jopy.12322

Witt U (2005) From sensory to positivist utilitarianism and back: the rehabilitation of naturalistic conjectures in the theory of demand. Pap. Econ. Evol. 2005(07):1–31

Acknowledgements

We would like to thank participants in various seminars and presentations in Belfast, Cologne, Dublin, Göttingen, Kassel, London, Newcastle, Nice, Rotterdam, Stirling, Wageningen, and Warwick and in particular Daniel Read and two anonymous reviewers for valuable comments and the Scottish Institute for Research in Economics (SIRE) for financial support. All remaining errors are ours.

Author contributions

All authors contributed equally to this work.

Competing interests

The authors declare no competing interests.

Ethical approval

This study was performed in line with the principles of the Declaration of Helsinki. Ethical approval was granted on March, 27th, 2014 by the Ethics Committee of the Stirling Management School (SMS) at the University of Stirling (no ethical approval number was provided).

Informed consent

Before starting the study, participants received detailed written information about the research digitally on the first page of the survey. This information included the voluntary nature of the study, participants' right to withdraw from the study at any time, the anonymity and confidentiality that was guaranteed, and the use of the data for scientific research. Participants provided their informed consent in digital form by clicking an "I agree" button at the bottom of the informed consent page before beginning the study.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-024-03845-1.

Correspondence and requests for materials should be addressed to Leonhard K. Lades.