Year: 2020

Know your weaknesses: Sophisticated impulsiveness motivates voluntary self-restrictions

Soutschek, Alexander ; Tobler, Philippe N

Abstract: Restricting one's access to temptations (precommitment) facilitates the achievement of long-term goals. The sophisticated impulsiveness model of precommitment posits that impulsive agents who are aware that they are impulsive should show the strongest preference for precommitment. Empirically however, two central predictions of this theoretical notion remained untested: whether impulsiveness causally drives the demand for precommitment and whether the willingness to precommit depends on metacognitive awareness of one's impulsiveness. Here, we tested these predictions in three independent experiments. Participants performed a delay discounting task in which they could precommit to larger-later rewards. The results of Experiment 1 provide causal evidence that reducing impulse control capacities increases precommitment demand. Moreover, Experiments 2 and 3 support the hypothesis that metacognitive awareness of one's impulsiveness moderates the relationship between impulsiveness and precommitment. Together, our data put the sophisticated impulsiveness model of precommitment on strong empirical foundations.

DOI: https://doi.org/10.1037/xlm0000833

Originally published at:
Soutschek, Alexander; Tobler, Philippe N (2020). Know your weaknesses: Sophisticated impulsiveness motivates voluntary self-restrictions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 46(9):1611-1623.
DOI: https://doi.org/10.1037/xlm0000833

Running head: Sophisticated impulsiveness and precommitment

Know your weaknesses: sophisticated impulsiveness motivates voluntary self-restrictions

Alexander Soutschek[1, 2] and Philippe N. Tobler[2, 3]

[1] Department of Psychology, Ludwig Maximilan University Munich, Munich, Germany

[2] Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland

[3] Neuroscience Center Zurich, University of Zurich, Swiss Federal Institute of Technology Zurich, Zurich, Switzerland

Keywords: self-control, impulse control, impulsiveness, metacognition, reward

Word count: 4995 words

Author contributions:

AS and PNT developed study concept and design; AS performed research and statistical analyses; AS and PNT wrote manuscript

Correspondence should be addressed to:
Alexander Soutschek
Ludwig Maximilian University Munich
Department of Psychology
Leopoldstr. 13
80802 Munich, Germany
E-mail: alexander.soutschek@psy.lmu.de

Abstract

Restricting one's access to temptations (precommitment) facilitates the achievement of long-term goals. The sophisticated impulsiveness model of precommitment posits that impulsive agents who are aware that they are impulsive should show the strongest preference for precommitment. Empirically however, two central predictions of this theoretical notion remained untested: whether impulsiveness causally drives the demand for precommitment and whether the willingness to precommit depends on metacognitive awareness of one's impulsiveness. Here, we tested these predictions in three independent experiments. Participants performed a delay discounting task in which they could precommit to larger-later rewards. The results of Experiment 1 provide causal evidence that reducing impulse control capacities increases precommitment demand. Moreover, Experiments 2 and 3 support the hypothesis that metacognitive awareness of one's impulsiveness moderates the relationship between impulsiveness and precommitment. Together, our data put the sophisticated impulsiveness model of precommitment on strong empirical foundations.

Introduction

Impulse control enables biological agents, including humans, to resist short-term temptations for the sake of long-term goods. However, the ability to control impulses is limited (Muraven & Baumeister, 2000). Humans and animals may therefore have developed strategies that protect against impulse control failures. One such strategy is to prevent access to desired short-term temptations by making a binding choice for long-term rewards (“precommitment”) (O'Donoghue & Rabin, 1999; Strotz, 1956; Thaler & Shefrin, 1981). Due to the potential of precommitment for protecting impulsive individuals against control failures, precommitment devices have been designed for both economic and clinical purposes to counteract impulsiveness. For example, rates of saving for retirement can be increased by inaccessible investment funds (David Laibson, Repetto, & Tobacman, 1998), or pathological gamblers can benefit from slot machines requiring to pre-set a maximum loss limit (Ladouceur, Blaszczynski, & Lalande, 2012). Thus, precommitment promises to be an effective means to promote impulse control, particularly in impulsive agents (for a review, see Bryan, Karlan, & Nelson, 2010).

Theoretical models posit that the demand for precommitment essentially results from impulsiveness (Kurth-Nelson & Redish, 2010, 2012; O'Donoghue & Rabin, 1999; Rachlin & Green, 1972; Thaler & Shefrin, 1981). In intertemporal decision-making, impulsive agents are present-biased and overvalue immediate over delayed rewards (hyperbolic discounting) (Frederick, Loewenstein, & O'Donoghue, 2002; D. Laibson, 1997; Strotz, 1956). With present bias, agents can show time inconsistent preferences. In other words, impulsive agents may change their preference from originally higher-valued later rewards to originally lower-valued sooner rewards when the time of the delivery of the sooner reward is getting closer. To avoid such preference reversals, impulsive agents should show a demand for committing to the long-term reward option. More elaborate accounts reserve this demand only for sophisticated agents who are aware of their own impulsiveness, but not for naïve agents (D. Laibson, 2015;

O'Donoghue & Rabin, 1999). In the following, we will refer to such an account of precommitment as “sophisticated impulsiveness” model.

To the extent that it is available, empirical evidence for these theoretical assumptions is mixed. While some studies support the claim that impulsive individuals benefit most from precommitment (Ashraf, Karlan, & Yin, 2006; Ladouceur et al., 2012), others found no evidence that impulsive populations such as individuals with drug addiction or low-income groups particularly benefit from precommitment devices (Fuller & Roth, 1979; Venti & Wise, 1990). Crucially, previous empirical studies only assessed precommitment demand in low-impulsiveness versus high-impulsiveness groups (Bryan et al., 2010), but such correlative analyses do not allow drawing conclusions about the *causal* relationship between impulsiveness and precommitment. In other words, it is unclear whether impulsiveness causally drives the willingness to restrict access to the source of temptation, as claimed by the sophisticated impulsiveness framework. Moreover, also the sophistication assumption, i.e. that precommitment demand depends on the match between perceived and actual impulse control, has not been empirically tested. O'Donoghue and Rabin (1999) consider the fact that people use precommitment devices as evidence for sophistication, but it has never been shown that sophistication leads to precommitment.

Due to the mixed evidence for central assumptions of the sophisticated impulsiveness account, alternative models of precommitment have been developed. At variance with the idea that impulsiveness drives precommitment demand, an alternative account posits that precommitment should not be chosen by impulsive but by patient agents because the decision to precommit itself requires impulse control (Noor, 2007). Due to the lack of conclusive empirical evidence, however, it is not possible to decide between these conflicting accounts on the motives of precommitment decisions.

Here, we empirically tested two central assumptions of the sophisticated impulsiveness account of precommitment that have been challenged by alternative accounts. First, we tested the causal role of impulse control for motivating precommitment decisions by assessing the impact of experimentally enhanced impulsiveness on precommitment demand (Experiment 1). Second, we examined the hypothesized moderating role of sophistication. Thus, we tested whether impulsive agents who are aware of their impulse control deficits show a stronger demand for precommitment than patient or naïve agents (Experiments 2 and 3). We find evidence in support of both assumptions. Our findings suggest that the sophisticated impulsiveness account provides a plausible explanation for the motives underlying voluntary self-restrictions.

EXPERIMENT 1

In Experiment 1, we administered an intertemporal decision task that allowed measuring both impulsiveness and the willingness to precommit within the same experimental task (Crockett et al., 2013; Soutschek et al., 2017). To establish a *causal* link between impulse control and precommitment demand, we experimentally manipulated impulse control capacities. Prior to the precommitment task one group of participants had to perform a 3-back task. This task is cognitively taxing because it requires continuous maintenance and updating of three items in working memory (Braver et al., 1997). As control task, another group of participants performed a 0-back task that did not impose any working memory demands. High working memory demands were previously found to impair impulse control processes in intertemporal choice relative to a less demanding control condition (Blain, Hollard, & Pessiglione, 2016). We note that we do not make any strong theoretical assumptions about the nature of such a depletion-like effect and that the theoretical claims formulated by the ego depletion framework (Baumeister & Vohs, 2016; Hagger, Wood, Stiff, & Chatzisarantis, 2010) were controversially debated recently (Carter & McCullough, 2014; Hagger et al., 2016; Inzlicht, Schmeichel, & Macrae, 2014). Instead, we only assume that having dealt with high demands on executive control can impair the exertion of control processes in a subsequent task

(Blain et al., 2016; Dang, Bjorklund, & Backstrom, 2017). Impairing executive control (and thus impulse control) allowed us to distinguish between theoretical accounts on the causal relationship between impulsiveness and precommitment. According to the sophisticated impulsiveness account, impulsiveness enhances the demand for precommitment, such that reducing impulse control capacities should increase preference for precommitment. In direct opposition, if precommitment itself requires impulse control (Noor, 2007), experimentally reduced impulse control should result in fewer precommitment decisions.

Materials and Methods

Participants

Seventy-nine heterosexual males (mean age = 22.4 year, age range 18-33 years) volunteered in the study after having given informed consent. Participants were reimbursed with 50 Swiss francs for their participation. The study was approved by the local ethics committee. Power analyses suggested that 37 participants per group (thus a total of 74 participants) were required to detect a statistical effect (alpha = 5%) with a probability of 80%, given the effect size (Cohen's d = 0.48) from a previous study manipulating intertemporal choices with working memory demands (Blain et al., 2016).

Stimuli and task design

Precommitment task. To measure the individual willingness to precommit, we applied a decision task we had already used in previous studies (Crockett et al., 2013; Soutschek et al., 2017). Participants made choices between immediately available smaller sooner (SS) and larger later (LL) rewards that were delivered after temporal delays of 4-10 s. Due to the relatively short delays, we used erotic stimuli as primary reinforcers consumable at the time of delivery instead of monetary rewards (Prevost, Pessiglione, Metereau, Clery-Melin, & Dreher, 2010). To approximately match the subjective value of SS and LL rewards across participants, all participants rated the attractiveness of 300 pictures of women in lingerie on a Likert-scale of 0-10 seven days before the main experimental session. We used these ratings to construct individualized stimulus sets by first removing all pictures rated as unattractive (ratings of 0 or 1) and then splitting the remaining pictures at each individual's median. We defined pictures above and below the median as LL and SS rewards, respectively. We never presented the same picture more than once in the main experimental session to avoid saturation effects.

To measure individual differences in impulsiveness and willingness to precommit, the task comprised two conditions: precommitment and willpower (Figure 1). In the *precommitment* condition, participants first decided in a pre-delay phase (4 s) to either have a free choice between the SS and the LL option during a subsequent delay phase, or to precommit to the LL option. If participants precommitted to the LL reward, the SS option was not available for selection during the delay phase, after which the image for the LL reward was presented. In contrast, if participants preferred having free choice, they could choose the SS option during the indicated temporal delay. If they chose this option, the SS reward was presented immediately. If they did not, the LL reward was presented automatically after the delay phase. Participants indicated their choice by pressing the left control key (for the option shown on the left side of the screen) or the right control key (for the option shown on the right side of the screen) on a standard keyboard. The assignment of options to screen sides was counterbalanced across participants. Depending on participants' choices, either the SS (immediately after selection) or the LL reward (after the delay) was presented for 2.5 s, followed by a variable inter-trial interval (ITI) in which a fixation cross was presented on the centre of the screen. If participants chose the LL option, the length of the ITI was 1.5 s, whereas in case of SS choices the remaining duration of the terminated delay phase was added to these 1.5 s. This ensured that participants could not finish the task faster by choosing the SS reward, and we informed participants about this procedure before the start of the experiment.

The *willpower* condition provided a measure of impulsiveness independent of the possibility to precommit. Willpower trials started directly with a delay phase (omitting the pre-delay phase), where the SS reward was available for immediate selection and the LL reward was presented automatically after the indicated temporal delay. This delay phase terminated immediately if participants chose the SS option (in this case, the remaining part of the terminated delay phase was again added to the ITI). Thus, the willpower condition assessed participants' ability to suppress the temptation to choose the immediately available SS reward while waiting for the LL reward.

N-back task. Participants performed a letter version of the n-back task in either a 3-back (depletion group; N = 43) or a 0-back (control group; N = 36) condition (Braver et al., 1997). The sample size incidentally happened to be larger in the 3-back than the 0-back condition because the majority of no-shows (who took part only in the first session for the picture rating task but did not show up for the main experimental session) had been assigned to the 0-back condition. In both conditions, participants were presented a stream of letters (b, c, d, g, p, t, w). To increase task difficulty, we used letters that were phonologically similar in German (Soutschek & Schubert, 2016; Soutschek, Strobach, & Schubert, 2013). In the 3-back condition, participants were instructed to press the space bar on a keyboard if the currently presented letter was identical to the letter presented three trials before, such that participants permanently had to maintain and update three letters in working memory. In the 0-back condition, in contrast, participants were instructed to respond to a letter defined at the start of a block. Thus, the 0-back task required no updating and less maintenance (only one instead of three letters) than the 3-back task. Each trial of the n-back task started with the presentation of a letter for 2000 ms, followed by a fixation cross (1500 ms). If the n-back letter was a target stimulus, the response had to be executed before the start of the next trial.

Procedure

Experiment 1 followed a pre-test/post-test design that allowed assessing how the working memory manipulation (3-back vs. 0-back) changed behaviour in the precommitment task (Figure 2A). In the pre-test session, participants first performed six blocks (each 6 trials, with balanced numbers of 4, 7, and 10 s delay trials) of the precommitment task (three blocks of the precommitment and of the willpower condition each). To increase the distinctiveness of the conditions, the frames of the choice options were shown in different colours (the assignment of conditions to colours was counterbalanced across participants). Moreover, before the start of a block, a cue presented for 2 s alerted participants to the upcoming condition (e.g., "blue task" for the precommitment condition). After the pre-test, participants performed three blocks (each 62 trials) of either the 3-back (working memory group) or 0-back (control group) task in order to manipulate their impulse control capacities. After the n-back task, participants were presented again with six blocks of the precommitment task (post-test).

Statistical analysis

In the precommitment task, we analysed differences in the percentage of precommitment decisions for the precommitment condition and the percentage of LL reward deliveries for the willpower condition. As recommended for pre-test/post-test designs (Dugard & Todman, 1995), we assessed behaviour in the post-test session with an analysis of covariance (ANCOVA) where we entered the individual pre-test data as covariate to control for pre-existing baseline differences. In the n-back task, we analysed performance rates, defined as the difference between hit rate (percentage of correctly detected targets) and false alarms (percentage of erroneous responses to non-targets). All analyses were run with IBM SPSS 22. The alpha threshold was set to 5%.

Results

We first evaluated the effectiveness of the working memory manipulation in the n-back task. Participants showed a significantly lower performance rate in the 3-back (67.0% correct responses) compared with the 0-back (99.9%) condition, $t(77) = 9.50$, $p < 0.001$, Cohen's d = 2.22, indicating that the 3-back task was mentally more demanding than the 0-back task (Figure 2B). In addition, we replicated previous findings (Crockett et al., 2013; Soutschek et al., 2017) that in the pre-test participants obtain more LL rewards in the precommitment (69%) than in the willpower (61%) condition, $t(78) = 5.21$, $p < 0.001$, Cohen's d = 0.61, consistent with the view that the possibility to precommit improves self-control.

Next, we tested how performing the 3-back vs. the 0-back task changed decisions in the precommitment and willpower conditions. For this purpose, we computed ANCOVAs on the post-test data (percentage of precommitment decisions in the precommitment condition and percentage of LL reward deliveries in the willpower condition) with the factors Group (0-back vs. 3-back), Condition (Precommitment vs. Willpower), and Delay (Table 1 and Table 2). In addition, we added mean choices from the pre-test as covariate to control for individual baseline differences. This analysis yielded a significant Group × Condition interaction, $F(1, 75) = 6.44$, $p = 0.01$, $\text{eta}^2_{\text{partial}} = 0.079$, suggesting that performing the 3-back relative to the 0-back condition differentially affected decision-making in the precommitment and willpower conditions. By extension, this finding cannot be explained by the assumption that the 3-back task demands generally changed the value of SS or LL rewards, because in this case we should have observed the same result pattern in the precommitment and the willpower condition.

To test whether working memory demands differentially affected behaviour in the willpower and the precommitment conditions, we computed separate ANCOVAs for these conditions, which included all factors as the omnibus ANCOVA reported above except the factor Condition (precommitment vs. willpower). In the willpower condition, a significant Group × Delay interaction, $F(2, 152) = 6.82$, $p = 0.001$, $\text{eta}^2_{\text{partial}} = 0.082$, indicated that performing the 3-back task reduced the capacity to wait for LL rewards as function of delay length. Post-hoc tests revealed that the 3-back, relative to the 0-back, task resulted in fewer LL outcomes only for a delay of 4 s, $t(76) = 2.64$, $p = 0.01$, not of 7 or 10 s, both $t < 1$, both $p > 0.59$ (Figure 2C and Table 3). The impact of previous 3-back (relative to 0-back) performance might be more pronounced for shorter than for longer delays due to floor effects, i.e. cognitive demands could not (statistically significantly) further reduce LL outcomes at longer delays due to the already relatively low number of LL outcomes at these delays. In fact, in the baseline pre-test session 20% of all participants never waited for the LL reward at delays of 10 s, compared to only 1% and 5% at delays of 4 s and 7 s, respectively. Similarly, in the precommitment task 4%, 8%, and 19% of all participants never waited for the LL outcome at the delays of 4, 7, and 10 s, respectively. When removing these 26 participants from the analysis, the main effect of Group showed a trend-level effect, $F(1, 50) = 3.49$, $p = 0.068$, $eta^2_{partial} = 0.065$, while the Group × Delay was clearly above the statistical threshold, $F(2, 100) < 1$, $p = 0.54$, $eta^2_{partial} = 0.012$. When controlling for floor effects, there was thus no longer evidence that group effects on willpower are stronger on shorter than on longer delays.

In addition, when conducting the same analysis (including all participants) on the decision time until participants made a choice for the SS reward, we observed a main effect of Group, $F(1, 76) = 4.88$, $p = 0.03$, $eta^2_{partial} = 0.060$, but no Group × Delay interaction, $F < 1$, $p = 0.48$, $eta^2_{partial} = 0.010$. Specifically, in line with the assumption of impaired impulse control, participants more quickly gave in to choosing the SS reward after 3-back (mean = 2.46 s) compared with after 0-back (mean = 3.23 s) performance. Taken together, the findings of less LL outcomes in the 3-back group is consistent with previous results that performing a highly demanding working memory task impairs the subsequent exertion of impulse control in intertemporal choice (Blain et al., 2016).

In the precommitment condition, increased working memory load increased demand for precommitment with long delays. Specifically, a significant Group × Delay interaction, $F(2,$

152) = 5.22, *p* = 0.006, $eta^2_{partial}$ = 0.064, was qualified by participants making more precommitment decisions after performing the 3-back compared with the 0-back task for delays of 10 s, *t*(76) = 2.22, *p* = 0.03, but not for delays of 4 or 7 s, both $t < 1$, both $p > 0.37$ (Figure 2D and Table 4). We note, though, that the Group × Condition × Delay interaction was not significant in the omnibus ANCOVA, *F*(2, 150) < 1, *p* = 0.81, $eta^2_{partial}$ = 0.003, such that our results do not allow concluding that working memory demands affect behaviour in the precommitment versus willpower condition at different delays. Nevertheless, the opposite directions of the effects in these conditions suggest that performing a demanding working memory task reduces the ability to exert impulse control and at the same time increases the propensity to precommit in order to avoid impulse control failures.

Lastly, we tested whether post-test minus pre-test changes in precommitment (at delays of 10 s, i.e. where we had observed significant effects) co-vary with post-test minus pre-test changes in LL outcomes in the willpower condition (at delays of 4 s). However, this correlation was not significant, r = -0.10, p = 0.40, suggesting that changes in willpower alone are not sufficient to explain changes in precommitment behaviour.

Discussion

When their ability to resist temptations was impaired after high mental effort, participants showed a stronger preference for precommitment to the LL reward option. They might thus have used the possibility to precommit to compensate their enhanced impulsiveness. This assumption is in line with the impulsiveness hypothesis according to which individuals should show a stronger willingness to restrict their access to the tempting option with increasing impulse control deficits. The finding that precommitment increased rather than decreased after mental effort exertion is at variance with the alternative claim that precommitment requires impulse control (Noor, 2007).

Our findings converges with a previous report showing that those who are more impulsive are more willing to precommit (Wertenbroch, 1998). However, this study allowed only correlative inferences on the relationship between impulsiveness and precommitment, whereas our experimental manipulations provide causal evidence that increased impulsiveness motivates precommitment decisions.

Our finding of increased impulsivity after working memory demands replicates previous reports (Blain et al., 2016). However, one might argue that the observed result pattern can alternatively be explained by a stronger desire to relinquish control after the 3-back compared with the 0-back condition. According to this view, participants make more SS reward choices in the willpower condition after 3-back (relative to 0-back) performance due to the increased costs of resisting the SS reward (Burger, 1989; Saunders, Milyavskaya, & Inzlicht, 2015). This account assumes that the behaviours in both the precommitment and the willpower condition are the consequence of lower perceived control, whereas we interpreted the lower number of LL outcomes in the willpower condition as reflecting failures in objective impulse control. Crucially, however, both approaches agree that precommitment is motivated by the lower perceived likelihood of receiving LL outcomes after high mental demands, such that they both support the claims of the sophisticated impulsiveness account. We note though that the non-significant result of the correlation analysis suggests variation in precommitment and willpower to have at least partially different sources. This appears to favour the assumption that the fewer LL outcomes in the willpower condition reflect objective impulse control failures, whereas only precommitment choices are driven by perceived impulsiveness (which depends on both objective impulsiveness and an individual’s metacognitive skills). Taken together, we thus believe that the most parsimonious explanation for our results is that strong working memory demands impaired the ability to exert self-control in intertemporal choice.

EXPERIMENT 2

In Experiment 2, we tested whether the impact of impulsiveness on precommitment (as suggested by Experiment 1) is moderated by agents' metacognitive awareness of their impulsiveness. In addition to the precommitment task employed in Experiment 1, we therefore assessed participants' impulse control beliefs as measure of sophistication. Thereby, we tested the prediction of the sophisticated impulsiveness account that sophisticated impulsive agents show the strongest demand for restricting their choice options to avoid impulse control failures.

Economic models make the minimal assumption that precommitment should be more strongly preferred by impulsive agents who are sophisticated (i.e., foresee future self-control failures) rather than naïve (i.e., do not anticipate future self-control problems) (D. Laibson, 2015; O'Donoghue & Rabin, 1999). However, these accounts do not distinguish between different aspects of metacognition such as metacognitive bias (the degree to which agents over- or underestimate their impulsiveness) and metacognitive accuracy (the accurateness of an agent's subjective self-control beliefs). To inform these economic models, and to ground them on more elaborated concepts of metacognition, we therefore tested the roles of metacognitive bias and accuracy in moderating the relationship between objective impulsiveness and precommitment.

Materials and Methods

Participants

Eighty-two healthy heterosexual males (mean age = 22.6 years, age range 18-26 years) volunteered in the study after having given informed consent. Participants were reimbursed with 50 Swiss francs for their participation. *A-priori* power analyses suggested that a sample of 80 participants is required to achieve a power of 80% (alpha = 5%), assuming a moderate correlation (r=0.3) between impulsiveness and willingness to precommit.

Stimuli and task design

Precommitment task. Participants performed the precommitment and willpower conditions of the precommitment task in the same way as in Experiment 1. In addition, we employed two further control conditions (Figure 1): the *choice* condition served to control for the possibility that an individual's decision *not* to precommit in the pre-delay phase of the precommitment condition might result from two different motivations, either an aversion against removing choice options or a preference for the SS over the LL reward. To distinguish between these two possibilities, the *choice* condition offered both the SS and the LL reward for selection in the pre-delay phase. Participants received the chosen reward after the indicated delay without having the possibility to reverse their choice. Choice behaviour in this condition thus assessed individual preferences for (binding choices of) LL over SS rewards in the pre-delay phase. Note that it was not necessary to control for this potential confound in Experiment 1 because we examined differences in the willingness to precommit between post-test and pre-test sessions.

In addition, we also controlled for participants' preference for signalling (to oneself and/or the experimenter) patience by making non-binding choices of LL rewards (*opt-out* condition). Similar to precommitment and choice trials, opt-out trials started with a pre-delay phase in which participants chose either to see the SS reward immediately after the pre-delay phase, or to have the free choice between the LL and the SS reward during the delay phase. In contrast to the precommitment and choice condition, it was therefore not possible to make binding choices of LL rewards in this condition. We assessed whether the possibility of making binding LL reward choices in the precommitment condition increased the probability of LL reward deliveries relative to the opt-out condition where no binding choices were possible. Thus, the opt-out condition allowed us to determine the efficiency of precommitment as self-control strategy.

Impulsiveness and confidence ratings. To test whether sophisticated impulsive individuals show the strongest precommitment demand, participants performed a rating task in which we asked them how likely they believed to be able to wait for the LL reward in the willpower condition. Participants rated their likelihood of waiting for the LL reward instead of choosing the SS reward separately for the three delays (4s, 7s, and 10s) on a scale that ranged from 0 (very unlikely) to 10 (very likely). These ratings thus provided a measure of an individual's impulsiveness beliefs that could be compared with the actual impulsiveness as observed in the willpower condition. After each impulsiveness rating, participants also rated their confidence regarding their impulsiveness beliefs on a scale from 0 (not confident at all) to 10 (very confident).

Procedure

Participants obtained detailed task instructions and performed 6 trials of each condition of the precommitment task (precommitment, choice, opt-out, willpower) as practice. The practice blocks were followed by the impulsiveness rating task. Then, participants performed a total of 4 blocks per task condition in randomized order, with each block containing 6 trials (with balanced numbers of 4, 7, and 10 s delays).

Data analysis

Comparisons between conditions of the precommitment task (choices in the pre-delay phases or outcomes after the delay phases of each trial) were conducted with paired-samples *t*-tests. In addition, we used general linear models (GLMs) to test the relationships between willingness to precommit, observed impulsive control (willpower condition), and self-reported impulsive control (impulsiveness ratings). To disentangle willingness to precommit from preference for LL over SS rewards (measured by the choice condition), we regressed precommitment decisions in the precommitment conditions on decisions in the choice condition. The resulting (unstandardized) residuals of this regression capture the unconfounded individual willingness to precommit. We then aimed to distinguish between two potential driving forces of the willingness to precommit: metacognitive bias and metacognitive accuracy. As measure of metacognitive bias, we subtracted observed impulsive control (%LL outcomes willpower condition) from the self-reported impulsive control (ratings). This score is zero if participants correctly estimated their impulsiveness (self-reported = observed impulse control), positive if they are too optimistic regarding their impulse control (self-reported > observed impulse control), and negative if they are too pessimistic (observed > self-reported impulse control). As measure of metacognitive accuracy, we computed the absolute value of metacognitive bias: the smaller the absolute difference between self-reported and observed control, the more accurately can participants estimate their true impulsiveness. To disentangle the impacts of metacognitive bias and accuracy, we computed two separate GLMs regressing precommitment residuals on fixed effect predictors for Observed impulse control, Confidence (to control for participants' confidence in their ratings), Metacognitive bias or accuracy, and all interaction terms. Parameters were estimated using restricted maximum likelihood estimations as implemented in SPSS.

Results

As for Experiment 1, we first validated the effectiveness of our experimental manipulations in the precommitment task. An ANOVA suggested that the number of LL outcomes significantly differed between the four experimental conditions, $F(3, 243) = 20.88$, $p < 0.001$, $\text{eta}^2_{\text{partial}} = 0.205$. We replicated the finding that the number of LL reward outcomes at the end of the delay phases was significantly higher in the precommitment condition (71%) than in the willpower condition (60%), $t(81) = 7.15$, $p < 0.001$. The precommitment condition also increased LL reward outcomes compared to the choice (64%) and opt-out (62%) conditions, both $t > 4.57$, $p < 0.001$, suggesting that the possibility to restrict one's access to temptations increased the likelihood of obtaining LL reward relative to conditions where precommitment was not possible. We moreover observed a higher number of LL outcomes in the choice compared with the willpower condition, $t(81) = 2.95$, $p = 0.004$, whereas the comparisons of the opt-out with the choice and willpower conditions yielded no significant results, both $t < 1.65$, $p > 0.10$ (Figure 3A and Table 5).

Next we investigated the prediction that sophisticated impulsive individuals (i.e., individuals who believe to be impulsive and actually are) have the strongest preference for precommitment. This hypothesis can be tested directly by assessing how observed impulse control and metacognitive accuracy (i.e., the difference between self-reported and observed impulse control) interact in motivating precommitment decisions. For that purpose, we analysed precommitment choices (after controlling for preference for LL outcomes in the choice condition; Figure 3B-D) with a GLM including predictors for Observed impulse control, Metacognitive accuracy, Confidence, and the interaction terms. While this analysis yielded no main effect of Observed impulse control, $t < 1$, $p = 0.66$, we found a significant Observed impulse control × Metacognitive accuracy interaction, beta = 0.08, $t(74) = 2.18$, $p = 0.03$. To resolve this interaction, we split our sample into four equal groups based on their observed impulse control scores. In the group with lowest impulse control, individuals were more likely to precommit if their metacognitive accuracy was close to zero, beta = 0.004, $t(20) = 2.15$, $p = 0.04$, i.e. if they were aware of their impulsiveness. By contrast, the less impulsive groups showed no significant relationship between metacognition and willingness to precommit, all $t < 1.12$, all $p > 0.27$ (Figure 3E/F). The finding that increased metacognitive awareness of one's impulsivity moderates the relationship between impulsiveness and precommitment is consistent with the assumptions of the sophisticated impulsiveness account.

In addition, we also tested for the possibility that precommitment decisions are driven by the degree to which participants over- or underestimate their impulsiveness (metacognitive bias) rather than by metacognitive accuracy. We therefore re-computed the above described

GLM on willingness to precommit, replacing the predictor Metacognitive accuracy with Metacognitive bias. This GLM showed no significant effects, all $t < 1.52$, all $p > 0.13$. Thus, our data suggest that metacognitive accuracy, rather than metacognitive bias, contributes to precommitment decisions. However, we would like to note that our measures for metacognitive bias and accuracy were strongly correlated, $r = 0.88$, $p < 0.001$, because participants generally tended to overestimate their impulse control abilities (independence would require that in the mean participants neither over- nor underestimate their impulse control). We thus refrain from drawing inferences regarding a stronger contribution of metacognitive accuracy relative to bias in motivating precommitment. In any case, our results support the hypothesized modulatory role of metacognition for precommitment.

Finally, we assessed the specificity of our results for binding precommitment choices by regressing non-binding LL reward choices in the opt-out condition on predictors for Observed impulse control, Metacognitive accuracy, Confidence, and the interaction terms (again controlling for LL reward preferences). This GLM yielded no significant effects, all $t < 1.72$, all $p > 0.09$, providing evidence for the specificity of our results for binding precommitment decisions.

Discussion

The data from Experiment 2 provide evidence for a moderating effect of metacognition on willingness to precommit, with impulsive (as compared to patient) individuals showing a stronger precommitment demand if they are aware of their impulsiveness. Thus, our findings provide empirical support for central assumptions of the sophisticated impulsiveness account of precommitment (D. Laibson, 2015; O'Donoghue & Rabin, 1999; Rachlin & Green, 1972; Thaler & Shefrin, 1981).

We note that both Experiments 1 and 2 used intertemporal choice tasks with erotic reinforcers. One might thus question whether our conclusions can be generalized to other types or domains of self-control, e.g. involving money. Experiment 3 therefore assessed whether the contribution of metacognitive awareness of one's impulsivity to precommitment is robust to choice situations involving monetary decisions.

EXPERIMENT 3

In Experiments 3, we employed an intertemporal choice task with decisions between monetary rewards at different temporal delays. To generalize our findings to different measures of metacognition, we assessed awareness of one's time preference based on confidence ratings after each intertemporal choice. The degree to which confident choices are less noisy than choices with low subjective confidence represents an established measure of the metacognitive awareness of one's economic preferences (De Martino, Fleming, Garrett, & Dolan, 2013). Experiment 3 therefore allowed assessing the robustness of the results of Experiment 2 to alternative measures of self-control and metacognitive accuracy.

Materials and Methods

Participants

30 healthy volunteers (mean age = 23.1 years, age range 20-28 years, 20 female) participated in the study after having given informed consent. Participants were reimbursed with 20 Swiss francs for their participation plus a bonus depending on their choices (see below).

Stimuli and task design

Confidence task. Participants performed a monetary intertemporal choice task where they made decisions between an immediately available SS reward (0-10 Swiss francs today, in steps of 1 Swiss franc) and a LL reward that was fixed to 10 Swiss francs and was delivered after a delay of 1-180 days (using the following delays: 1, 10, 20, 40, 80, 120, and 180 days). The SS and LL reward options were randomly presented on the left or right screen side.

Participants had to indicate their choice by pressing the left (for the left option) or right arrow key (for the right option) on a keyboard within 4 s. After each choice, participants had to indicate their confidence that they made the correct choice on a rating scale from 0 (low confidence) to 20 (high confidence) within a response interval of 3 s (Figure 4A).

Precommitment task. As for the confidence task, participants made choices between monetary rewards delivered in the future. One option consisted of a fixed monetary reward of 10 Swiss francs that was delivered after delays of 29-208 days (precommitment option; e.g., “10 Swiss francs in 38 days”; Figure 4B). When choosing this option, participants received the 10 Swiss francs after the given delay without having the possibility to reverse their choice. The other option (free-choice option) entailed an SS reward of 0-10 Swiss francs delivered after 28 days and a LL reward that was identical to the precommitment option (e.g., “5 Swiss francs in 28 days” or “10 Swiss francs in 38 days”). If participants decided for this free-choice option, they were re-contacted by the experimenter via email after 28 days and were asked to make a choice between the SS and LL rewards, with the delays adjusted for the bygone 20 days (in the current example, “5 Swiss francs today” or “10 Swiss francs in 10 days”). If they decided for the precommitment option, they received an information about the chosen option via email after 28 days without having the option to reverse their choice. The free-choice option thus allowed participants to reverse their current preference after 28 days, whereas with the precommitment option they made a binding choice for the LL reward.

Procedure

Participants performed the confidence task and the precommitment task in counterbalanced order. The confidence and precommitment tasks included a total of 168 and 108 trials, respectively. At the end of the experiment, one trial of the two tasks was randomly selected and the corresponding amount was paid after the given delay.

Data analysis

Data were analysed with mixed GLMs (MGLMs) as implemented in SPSS. The confidence task allowed us to measure an individual participant's metacognitive awareness of her time preferences. For that purpose, we first estimated each individual's preferences (as revealed by her choices) by fitting a standard hyperbolic discount function to the choices in the confidence task (equation 1):

(Equation 1) SV_{LL} = LL reward magnitude / (1 + k × delay)

Where SV_{LL} is the discounted subjective value of the LL reward and k is a participant-specific constant indicating the steepness of the hyperbolic discount function ("discount factor"). To translate subjective value into choices, we fitted the following softmax function to each participant's choices:

(Equation 2) P(choice of LL option) = $1 / (1 + e^{-\beta_{temp} \times (SV_{LL} - SV_{SS})})$

This function captures the likelihood of choosing the LL reward option as a function of the difference between the subjective value of the LL reward option (SV_{LL}) and the SS reward option (SV_{SS}). The free inverse temperature parameter β_{temp} captures how strongly participants relied on this value difference and temperature can be interpreted as noise. Parameters were estimated using maximum likelihood methods as implemented in Matlab (function fmincon). To measure metacognitive awareness of one's time preferences, we followed the approach described by De Martino et al. (2013). We first computed the difference between the value of the SS reward and the subjective value of the LL reward ($SV_{LL} - SV_{SS}$) based on the individual discount factors. We then regressed each individual participant's choices of LL versus SS rewards on z-standardized predictors for the subjective value difference ($SV_{LL} - SV_{SS}$), the confidence ratings, and the interaction term. The interaction term indicates the degree to which participants are aware of decision noise in the choice process and thus reflects a measures of metacognitive accuracy: the stronger the interaction between decision noise and confidence ratings, the better participants know their time preferences (metacognitive accuracy) (De

Martino et al., 2013). We note that, contrary to Experiment 2, this procedure allowed measuring only metacognitive accuracy but does not provide an established measure of metacognitive bias.

Based on this, we examined whether awareness of one's impulsivity predicted the willingness to precommit. As in Experiment 2, to disentangle the impact of precommitment demand from preference of LL outcomes on precommitment decisions, we first conducted a logistic regression that regressed choices in the precommitment task (0 = free-choice option, 1 = precommitment option) on the value difference between LL and SS options on each trial (using MGLMs as implemented in SPSS). The resulting residuals thus indicate the willingness to precommit after correcting for the preference for LL over SS rewards. These residuals were then submitted to a MGLM, which regressed them on predictors for Metacognitive accuracy (obtained from the confidence task), Commitment demand, and the interaction term. Commitment demand was computed by subtracting the value difference between SS and LL reward from the current perspective from the value difference between these options from the perspective in 28 days (when participants had to make a definite choice between the options). The higher this difference, the higher the likelihood of preference reversals (i.e., that a participant prefers the LL reward today and the SS reward in 28 days) and thus the potential benefit from precommitting. Based on the findings in Experiment 2, we expected to observe that higher metacognitive awareness of precommitment demand was associated with stronger preference for precommitment.

Results

An MGLM on the willingness to precommit revealed a negative effect of commitment demand, beta = -0.04, $t(8904) = 2.56$, $p = 0.01$, suggesting that participants with poor metacognition (Metacognitive accuracy = 0) were in fact less willing to precommit with increasing risk of preference reversals. This points to the subjective costliness of restricting one's action space and that participants who were not sensitive to the potential benefits of precommitment (due to their poor metacognition) were willing to precommit to the LL option predominantly when having a strong preference for the LL option. Importantly, we also observed a significant Commitment demand × Metacognitive accuracy interaction, beta = 0.04, $t(9943) = 2.78$, $p = 0.005$, indicating that increasing metacognitive skills strengthened the impact of commitment demand on choices to precommit. For illustration purpose, we conducted follow-up analyses by splitting the variable Metacognitive accuracy into quartiles. While commitment demand did not significantly predict precommitment in individuals with low-to-moderate metacognitive skills (lowest three quartiles), all beta > -0.04, all $t < 1.67$, all $p > 0.17$, participants with high metacognitive skills (highest quartile) increasingly preferred precommitment with higher commitment demand, beta = 0.06, $t(17) = 2.83$, $p = 0.01$ (Figure 4C). Again, the finding that metacognition increases the sensitivity to the expected benefit from precommitting is consistent with the predictions of the sophisticated impulsiveness account.

Discussion

The results of Experiment 3 suggest that metacognitive awareness of one's risk of preference reversals motivates precommitment choices. This replicates and extends the findings of Experiment 2, providing evidence for the robustness and generalizability of our conclusions to different types of self-control tasks, different types of reinforcers, and measures of metacognition.

GENERAL DISCUSSION

Restricting one's access to temptations is an effective strategy to protect against impulse control failure (Ariely & Wertenbroch, 2002; Bryan et al., 2010; Ladouceur et al., 2012; Schwartz et al., 2014). One account proposed a link between precommitment and the meta-cognitive awareness of impulsive agents of their impulsiveness (Kurth-Nelson & Redish, 2010, 2012; O'Donoghue & Rabin, 1999; Rachlin & Green, 1972; Thaler & Shefrin, 1981). However, empirical evidence for central assumptions of this framework was so far missing. Putting the theory on firmer ground, Experiment 1 corroborates the assumption that impulsiveness drives the demand for precommitment: reducing cognitive control resources causally increased both impulsiveness in intertemporal choices and the willingness to precommit. In addition, Experiments 2 and 3 corroborate the claimed moderating impact of sophistication: metacognitive accuracy strengthens the preference for precommitment with increasing objective impulsiveness (Experiment 2) or risk of preference reversals (Experiment 3). To the best of our knowledge, this is the first empirical demonstration that sophistication matters for precommitment, suggesting that sophisticated individuals prefer to restrict their access to temptations in order to protect themselves against impulse control failures. These findings are also consistent with neural data showing that the frontopolar cortex, a brain region related to metacognition, plays a crucial role in precommitment decisions (Crockett et al., 2013; Soutschek et al., 2017). Together, our results provide evidence for a crucial role of impulsiveness and sophistication for motivating precommitment choices.

The finding that metacognitive awareness of impulsiveness drives precommitment informs and may refine theoretical models of precommitment. It provides empirical support for the sophisticated impulsiveness account of precommitment (Kurth-Nelson & Redish, 2012; D. Laibson, 2015; O'Donoghue & Rabin, 1999; Thaler & Shefrin, 1981) but speaks against an alternative account assuming that precommitment decisions themselves require impulse control (Noor, 2007). We note that a further model claims that precommitment is motivated not by the avoidance of control lapses but by the elimination of the subjective costs of exerting control (Gul & Pesendorfer, 2001), a view that is supported by recent evidence (Toussaert, 2018). In analogy to the sophisticated impulsiveness model, metacognitive awareness of subjective self-control costs might moderate the willingness to precommit in the Gul-Pesendorfer account. Because it is plausible to assume that high impulsiveness is associated with high perceived self-control costs (Jimura, Chushak, Westbrook, & Braver, 2018), the sophisticated impulsiveness and the Gul-Pesendorfer account might be considered as alternative formulations for the role of metacognitive beliefs for precommitment, rather than being incompatible models of precommitment.

The current results have implications for the design of precommitment devices. Precommitment devices are applied in many different domains from the financial sector to psychiatry (Bryan et al., 2010). Our findings suggest that increasing the subjective awareness of one's impulsivity might be an effective means to motivate particularly impulsive agents to voluntarily restrict their access to temptations. Because impulse control deficits characterize disorders such as addiction or obesity (Hasler, 2012; Monterosso, Piray, & Luo, 2012), there might be untapped potential for improving these deficits by advising patients on how to restrict their access to the source of temptation. Yet, it has been noted that even many individuals who would strongly benefit from self-restrictions decide against precommitment (D. Laibson, 2015). Our results suggest that impulsiveness alone without awareness of one's susceptibility to temptations may not be sufficient to create strong demand for precommitment.

Acknowledgements

This work was supported by grant 981 from the Velux Foundation and grant 100014_165884 from the Swiss NSF to PNT. AS received a grant from the Richard-Buechner-Foundation and an Emmy Noether fellowship from the German Research Foundation. We are grateful to Séverine Toussaert for helpful discussions and to Cornelia Schnyder for expert help with data collection. The data that support the findings of this study are available on Open Science Framework (https://osf.io/uvgdw/).

References

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: self-control by precommitment. *Psychol Sci, 13*(3), 219-224.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *The Quarterly Journal of Economics, 121*(2), 635-672.

Baumeister, R. F., & Vohs, K. D. (2016). Chapter Two - Strength Model of Self-Regulation as Limited Resource: Assessment, Controversies, Update. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 54, pp. 67-127): Academic Press.

Blain, B., Hollard, G., & Pessiglione, M. (2016). Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. *Proc Natl Acad Sci U S A, 113*(25), 6967-6972. doi:10.1073/pnas.1520527113

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage, 5*(1), 49-62. doi:10.1006/nimg.1996.0247

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

Burger, J. M. (1989). Negative reactions to increases in perceived personal control. *J Pers Soc Psychol, 56*(2), 246.

Carter, E. C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: has the evidence for ego depletion been overestimated? *Front Psychol, 5*, 823. doi:10.3389/fpsyg.2014.00823

Crockett, M. J., Braams, B. R., Clark, L., Tobler, P. N., Robbins, T. W., & Kalenscher, T. (2013). Restricting temptations: neural mechanisms of precommitment. *Neuron, 79*(2), 391-401. doi:10.1016/j.neuron.2013.05.028

Dang, J., Bjorklund, F., & Backstrom, M. (2017). Self-control depletion impairs goal maintenance: A meta-analysis. *Scand J Psychol, 58*(4), 284-293. doi:10.1111/sjop.12371

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nat Neurosci, 16*(1), 105-110. doi:10.1038/nn.3279

Dugard, P., & Todman, J. (1995). Analysis of Pre-test-Post-test Control Group Designs in Educational Research. *Educational Psychology, 15*(2), 181-198. doi:10.1080/0144341950150207

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature, 40*, 350-401.

Fuller, R. K., & Roth, H. P. (1979). Disulfiram for the treatment of alcoholism. An evaluation in 128 men. *Ann Intern Med, 90*(6), 901-904.

Gul, F., & Pesendorfer, W. (2001). Temptation and self-control. *Econometrica, 69*(6), 1403-1435.

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., . . . Zwienenberg, M. (2016). A Multilab Preregistered Replication of the Ego-Depletion Effect. *Perspect Psychol Sci, 11*(4), 546-573. doi:10.1177/1745691616652873

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2010). Ego depletion and the strength model of self-control: a meta-analysis. *Psychol Bull, 136*(4), 495-525. doi:10.1037/a0019486

Hasler, G. (2012). Can the neuroeconomics revolution revolutionize psychiatry? *Neurosci Biobehav Rev, 36*(1), 64-78. doi:10.1016/j.neubiorev.2011.04.011

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends Cogn Sci, 18*(3), 127-133. doi:10.1016/j.tics.2013.12.009

Jimura, K., Chushak, M. S., Westbrook, A., & Braver, T. S. (2018). Intertemporal Decision-Making Involves Prefrontal Control Mechanisms Associated with Working Memory. *Cereb Cortex, 28*(4), 1105-1116. doi:10.1093/cercor/bhx015

Kurth-Nelson, Z., & Redish, A. D. (2010). A reinforcement learning model of precommitment in decision making. *Front Behav Neurosci, 4*, 184. doi:10.3389/fnbeh.2010.00184

Kurth-Nelson, Z., & Redish, A. D. (2012). Don'T let me do that! - models of precommitment. *Front Neurosci, 6*, 138. doi:10.3389/fnins.2012.00138

Ladouceur, R., Blaszczynski, A., & Lalande, D. R. (2012). Pre-commitment in gambling: a review of the empirical evidence. *International Gambling Studies, 12*(2), 215-230. doi:10.1080/14459795.2012.658078
Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*(2), 443-477. doi:Doi 10.1162/003355397555253
Laibson, D. (2015). Why Don't Present-Biased Agents Make Commitments? *American Economic Review, 105*(5), 267-272. doi:10.1257/aer.p20151084
Laibson, D., Repetto, A., & Tobacman, J. (1998). Self-Control and Saving for Retirement. *Brookings Papers on Economic Activity, 29*(1), 91-196.
Monterosso, J., Piray, P., & Luo, S. (2012). Neuroeconomics and the study of addiction. *Biol Psychiatry, 72*(2), 107-112. doi:10.1016/j.biopsych.2012.03.012
Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: does self-control resemble a muscle? *Psychol Bull, 126*(2), 247-259.
Noor, J. (2007). Commitment and self-control. *Journal of Economic Theory, 135*(1), 1-34. doi:10.1016/j.jet.2006.03.017
O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89*(1), 103-124. doi:DOI 10.1257/aer.89.1.103
Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M. L., & Dreher, J. C. (2010). Separate valuation subsystems for delay and effort decision costs. *J Neurosci, 30*(42), 14080-14090. doi:10.1523/JNEUROSCI.2752-10.2010
Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *J Exp Anal Behav, 17*(1), 15-22.
Saunders, B., Milyavskaya, M., & Inzlicht, M. (2015). Variation in cognitive control as emotion regulation. *Psychol Inq, 26*(1), 108-115.
Schwartz, J., Mochon, D., Wyper, L., Maroba, J., Patel, D., & Ariely, D. (2014). Healthier by precommitment. *Psychol Sci, 25*(2), 538-546. doi:10.1177/0956797613510950
Soutschek, A., & Schubert, T. (2016). The importance of working memory updating in the Prisoner's dilemma. *Psychol Res, 80(2)*, 172-180. doi:10.1007/s00426-015-0651-3
Soutschek, A., Strobach, T., & Schubert, T. (2013). Working memory demands modulate cognitive control in the Stroop paradigm. *Psychol Res, 77*(3), 333-347. doi:10.1007/s00426-012-0429-9
Soutschek, A., Ugazio, G., Crockett, M. J., Ruff, C. C., Kalenscher, T., & Tobler, P. N. (2017). Binding oneself to the mast: stimulating frontopolar cortex enhances precommitment. *Soc Cogn Affect Neurosci, 12*(4), 635-642. doi:10.1093/scan/nsw176
Strotz, R. H. (1956). Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies, 23*(62), 165-180.
Thaler, R. H., & Shefrin, H. M. (1981). An Economic-Theory of Self-Control. *Journal of Political Economy, 89*(2), 392-406. doi:Doi 10.1086/260971
Toussaert, S. (2018). Eliciting Temptation and Self-Control Through Menu Choices: A Lab Experiment. *Econometrica, 86*(3), 859-889.
Venti, S. F., & Wise, D. A. (1990). Have IRAs Increased U. S. Saving?: Evidence from Consumer Expenditure Surveys*. *The Quarterly Journal of Economics, 105*(3), 661-698. doi:10.2307/2937894
Wertenbroch, K. (1998). Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing Science, 17*(4), 317-337.

Tables

Table 1. Mean LL outcomes and precommitment decisions (in percent) in Experiment 1. Standard errors of the mean are in brackets.

			Precommitment condition			Willpower condition
			LL outcomes	precommitment decisions	LL outcomes after not precommitting	LL outcomes
0-back	Pre-test	4 s	86 (4)	79 (5)	30 (7)	82 (5)
		7 s	69 (6)	58 (6)	26 (6)	62 (5)
		10 s	55 (6)	47 (6)	15 (6)	41 (6)
	Post-test	4 s	89 (4)	85 (5)	13 (6)	84 (5)
		7 s	69 (6)	58 (6)	19 (6)	56 (7)
		10 s	48 (7)	42 (7)	14 (5)	46 (7)
3-back	Pre-test	4 s	81 (4)	71 (4)	22 (6)	72 (4)
		7 s	67 (5)	57 (5)	19 (6)	61 (5)
		10 s	55 (6)	46 (6)	17 (5)	47 (6)
	Post-test	4 s	83 (4)	77 (4)	20 (6)	62 (5)
		7 s	67 (6)	63 (6)	15 (5)	51 (6)
		10 s	61 (6)	54 (6)	15 (5)	50 (6)

Table 2. Results of the 2 (Group) × 2 (Condition) × 3 (Delay) ANCOVA on post-test behaviour in the precommitment and willpower conditions in Experiment 1 while controlling for pre-test baseline differences in these conditions.

Effect	**F**	**p-value**	**$eta^2_{partial}$**
Group	0.01	0.93	0.000
Condition	3.50	0.07	0.045
Delay	37.07	<0.001	0.331
Group × Condition	6.44	0.01	0.079
Group × Delay	7.88	0.001	0.095
Condition × Delay	1.97	0.14	0.026
Group × Condition × Delay	0.21	0.81	0.003

Table 3. Results of the 2 (Group) × 3 (Delay) ANCOVA on post-test behaviour in the willpower condition in Experiment 1 while controlling for pre-test baseline differences in this condition.

Effect	**F**	**p-value**	**$eta^2_{partial}$**
Group	2.28	0.14	0.029
Delay	21.68	<0.001	0.222
Group × Delay	6.82	0.001	0.082

Table 4. Results of the 2 (Group) × 3 (Delay) ANCOVA on post-test behaviour in the precommitment condition in Experiment 1 while controlling for pre-test baseline differences in this condition.

Effect	**F**	**p-value**	**$eta^2_{partial}$**
Group	1.89	0.17	0.024
Delay	34.62	<0.001	0.313
Group × Delay	5.22	0.006	0.064

Table 5. Mean LL outcomes (in percent) and decisions in Experiment 2, separately for decision phase, delay phase, and combined decision and delay phase. Standard errors of the mean are in brackets.

Condition	LL/Commitment choices (decision phase)	LL outcomes (delay phase)	LL outcomes (decision + delay phase)
Precommitment	60 (3)	26 (3)	71 (3)
Choice	64 (3)		
Opt-out	62 (3)	90 (2)	
Willpower		60 (3)	

Figure captions

Figure 1. Experimental conditions in the precommitment task. In all conditions, participants chose between an immediately available smaller-sooner (SS) reward and a larger-later (LL) reward delivered after 4, 7, or 10 s. Individualized stimulus sets of erotic images served as rewards. In Experiment 1, participants performed only the precommitment and willpower conditions, whereas in Experiment 2 we administered also the choice and opt-out conditions. (A) In the precommitment condition, participants decided either to have the free choice between SS and LL reward during the delay phase, or to precommit to the LL reward; if participants precommitted, the SS reward was removed from the choice set during the delay, such that they could not reverse their decision. (B) The willpower condition consisted of the delay phase only. Participants could choose the immediately displayed SS reward image anytime while waiting for the LL reward delivery. This condition allowed measuring individual differences in willpower, i.e. the capacity to resist SS rewards while waiting for LL rewards. (C) In the choice condition (which controls for participants' preferences for LL over SS rewards), participants made a decision between SS and LL reward and received the chosen reward after the indicated delay. Thus, they could not reverse their decisions during the delay phase. (D) In the opt-out condition participants also decided between SS and LL reward before the delay. However, if they chose the LL option they could still reverse their decision and switch to the SS image during the delay (non-binding LL reward choices).

Figure 2. Experimental procedure and results of Experiment 1. (A) Participants performed the precommitment and the willpower condition of the precommitment task (Figure 1) in both a pre-test and a post-test experimental session. In-between, participants performed a letter-version of the n-back task in either a 3-back (working memory group) or a 0-back condition (control group). (B) Participants gave significantly more correct responses in the 0-back than in the 3-back condition, suggesting that the latter was cognitively more demanding. (C, D)

Reducing the availability of cognitive control resources (working memory vs. control group) significantly (C) reduced impulse control in the willpower condition at the shortest delay and at the same time (D) increased voluntary self-restriction in the precommitment condition for long delays. Error bars represent standard error of the mean, asterisks indicate significant effects ($p < 0.05$).

Figure 3. Results of Experiment 2. (A) The frequency of LL reward outcomes increased significantly in the precommitment condition relative to all other conditions, providing evidence for the effectiveness of precommitment as self-control device. (B, C, D) Correlations between precommitment choices in the precommitment condition, LL decisions in the choice condition, and LL outcomes in the willpower conditions. To control for the impact of LL reward preference on precommitment decisions, we computed the residuals from regressing precommitment decisions on LL decisions in the choice condition. (E, F) More impulsive individuals were increasingly willing to precommit with better metacognitive awareness of impulsiveness (metacognitive accuracy scores closer to 0). For illustration purpose, we split our sample into more impulsive (E, lowest quartile) and more patient (F, highest quartile) individuals based on the percentage of LL outcomes in the willpower condition.

Figure 4. Experimental paradigm and results of Experiment 3. (A) In the confidence task, participants performed a monetary intertemporal decision task. After each choice, they rated their confidence in whether the choice made was the best decision. As measure of metacognitive accuracy, we computed the interaction term between participants' confidence rating and decision noise. (B) In the precommitment task, participants decided between making a binding choice for a LL reward (e.g., 10 Swiss francs in 38 days) and postponing the choice to the future (e.g., 5 Swiss francs in 28 days or 10 Swiss francs in 38 days). If participants precommitted to the LL reward, they received the monetary amount after the indicated delay without having the

possibility to reverse their choice. If participants postponed the decision, they were re-contacted after 28 days and then had to make a choice between the SS (5 Swiss francs today) and LL rewards (10 Swiss francs in 10 days). (C) On trials with high commitment demand (indicating increased risk of preference reversals), individuals with high metacognitive accuracy made more binding LL reward choices than individuals who were less aware of their time preferences. As for Experiment 2, we split our sample into individuals with low (lowest quartiles) and high metacognitive accuracy (highest quartile) for illustration purpose.

Figure 1

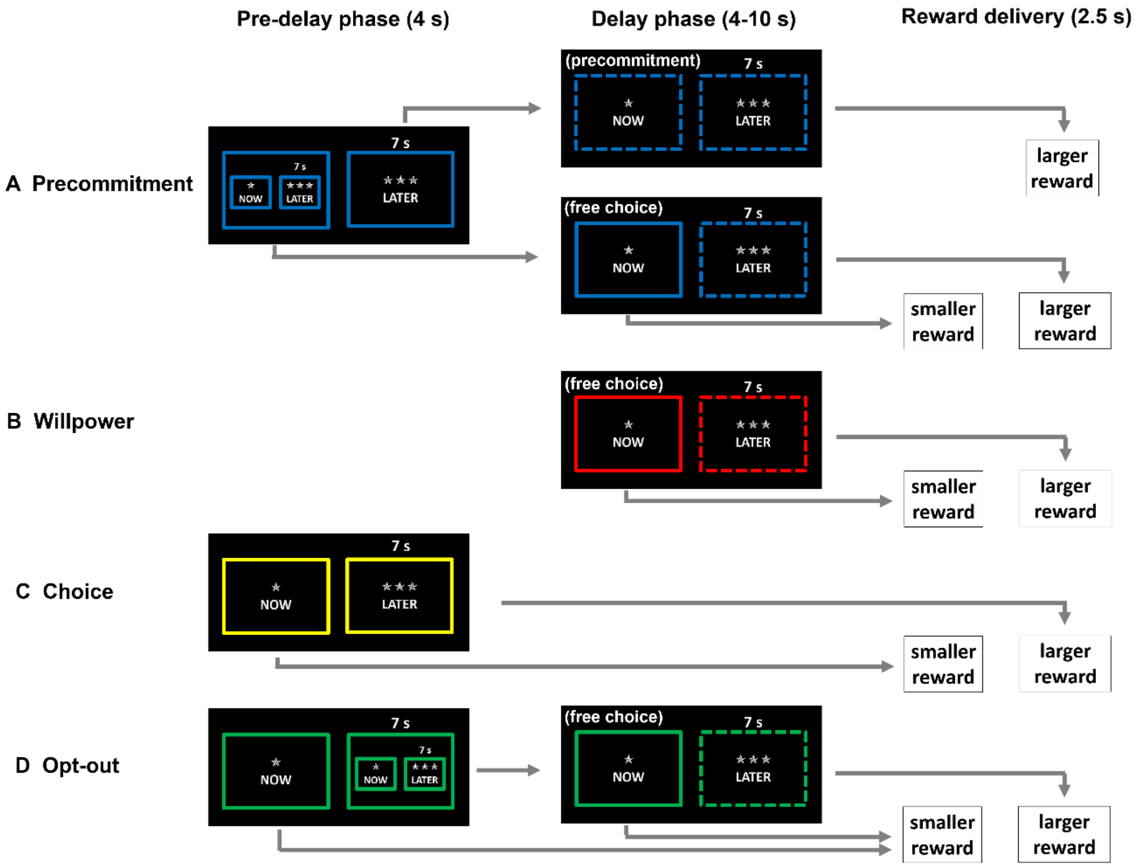

Pre-delay phase (4 s)
Delay phase (4-10 s)
Reward delivery (2.5 s)
A Precommitment
7 s
NOW
LATER
(precommitment)
(free choice)
larger reward
smaller reward
B Willpower
C Choice
D Opt-out

Figure 2

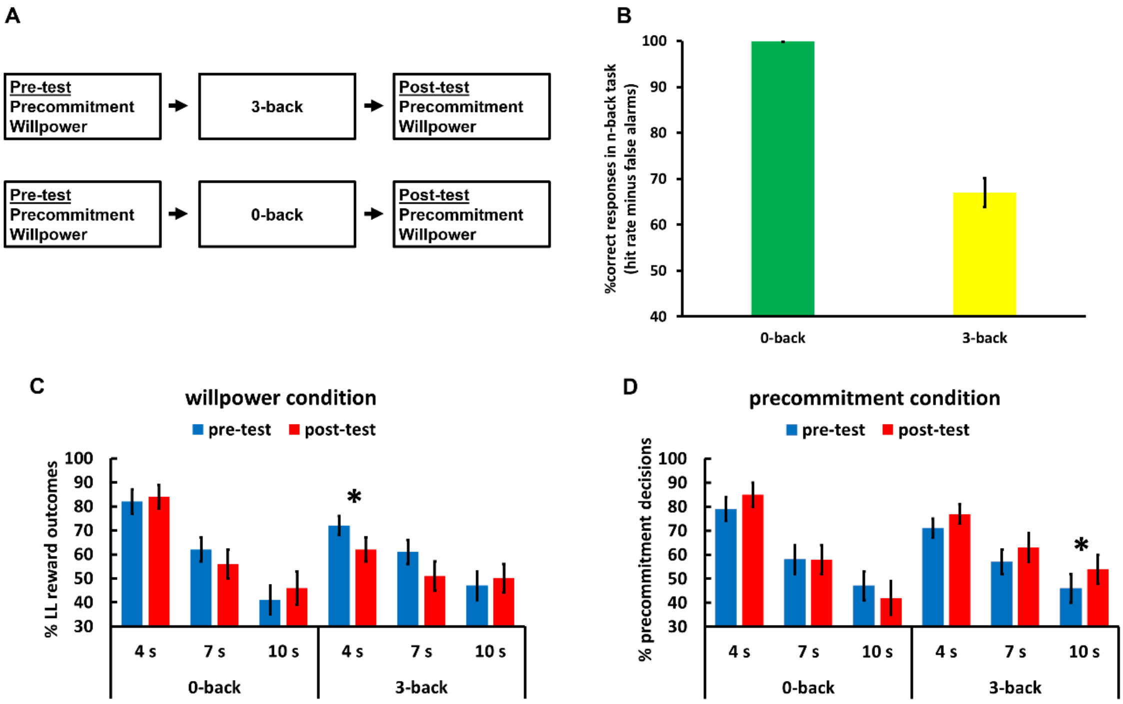

A
Pre-test
Precommitment
Willpower
3-back
Post-test
Precommitment
Willpower
Pre-test
Precommitment
Willpower
0-back
Post-test
Precommitment
Willpower
B
%correct responses in n-back task (hit rate minus false alarms)
0-back
3-back
C
willpower condition
pre-test
post-test
% LL reward outcomes
4 s
7 s
10 s
0-back
3-back
D
precommitment condition
% precommitment decisions

Figure 3

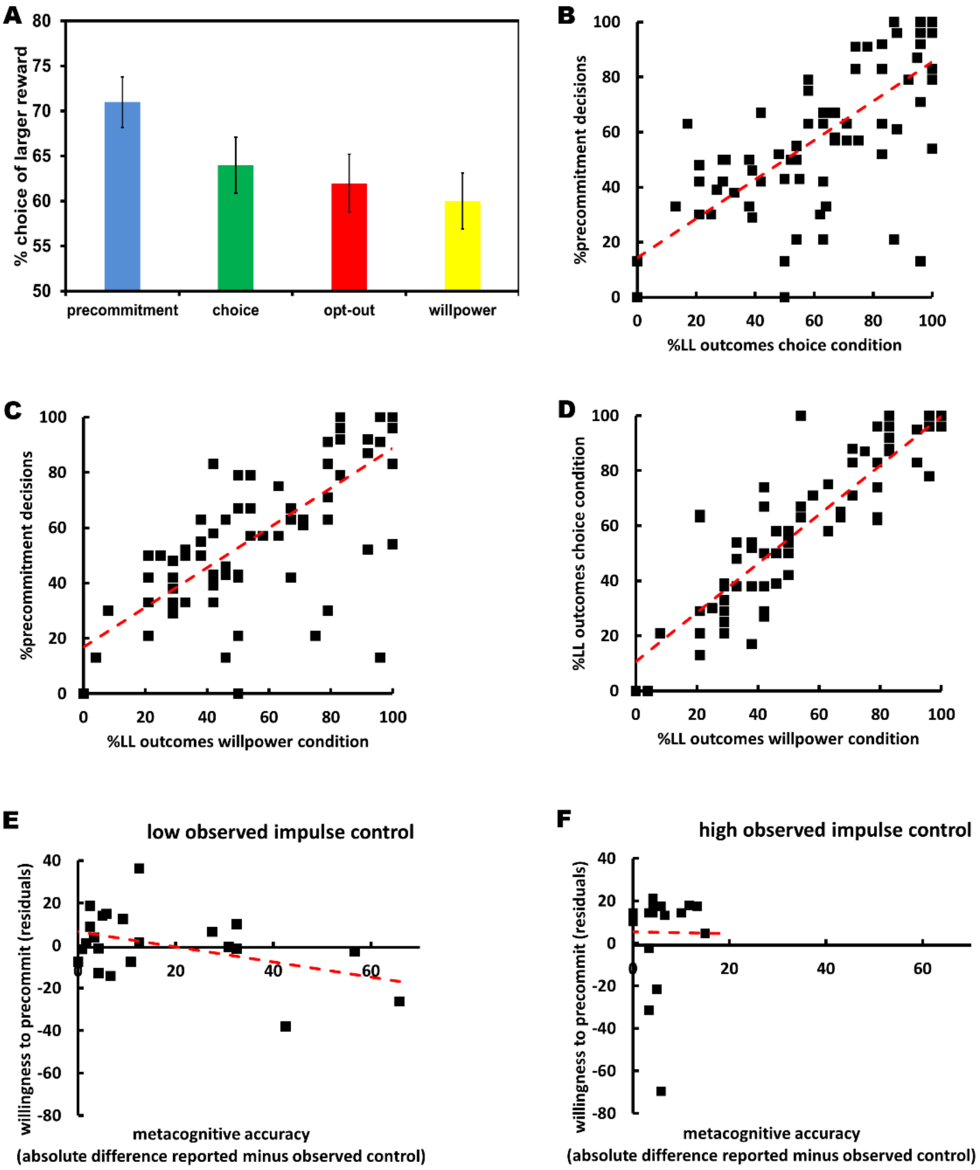
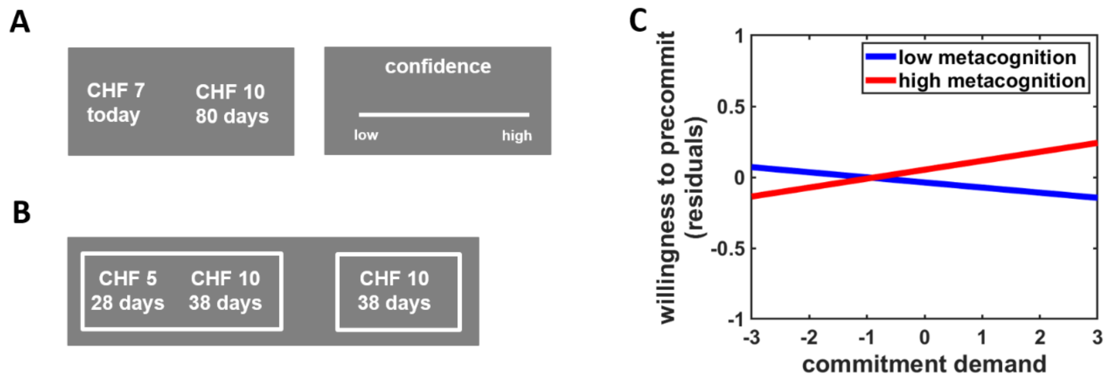

A
% choice of larger reward
80
75
70
65
60
55
50
precommitment
choice
opt-out
willpower
B
%precommitment decisions
%LL outcomes choice condition
C
%precommitment decisions
%LL outcomes willpower condition
D
%LL outcomes choice condition
%LL outcomes willpower condition
E
low observed impulse control
willingness to precommit (residuals)
metacognitive accuracy
(absolute difference reported minus observed control)
F
high observed impulse control
willingness to precommit (residuals)
metacognitive accuracy
(absolute difference reported minus observed control)

Figure 4


A
CHF 7
today
CHF 10
80 days
confidence
low
high
B
CHF 5
28 days
CHF 10
38 days
CHF 10
38 days
C
low metacognition
high metacognition
willingness to precommit
(residuals)
1
0.5
0
-0.5
-1
-3
-2
-1
0
1
2
3
commitment demand