**International Journal of Applied Behavioral Economics**
Volume 7 • Issue 4 • October-December 2018

# The Effects of a Commitment Device on Health Outcomes:
## Reputational Commitment and Weight Loss in an Online Experiment

Manu Manthri Savani, University College London, London, UK

## ABSTRACT

This article tests the effectiveness of commitment devices on weight loss. It can be applied as a health 'nudge' that locks in future behaviours to achieve a desired health outcome, by staking money or reputation on achieving that goal. A field experiment randomly assigned 364 clients of an online weight loss service to either a reputational commitment device or an upfront refund on the monthly fee. The reputational commitment intervention was expected to combat time inconsistency and promote greater weight loss. Weight outcomes were measured at 12 weeks. The results showed that all experimental groups lost weight on average, but the group experiencing the reputational commitment reported end weight outcomes 1.5 kg higher than the comparison group, indicating slower weight loss (± 0.05, effect size -0.32). One possible explanation for the findings is 'commitment overload'. The study nuances the understanding of commitment devices and how best to deploy them in health programmes including e-heath platforms.

## KEYWORDS

Behavioural Economics, Commitment Device, Dual Self, Field, Experiment, Health Behaviour, Nudge, Obesity, Weight Loss

## INTRODUCTION

Obesity remains a top public health priority, with 27% of adults in England and 40% of adults in the US now obese (Fuller, Mindell, & Prior, 2015; Hales, Carroll, Fryar, & Ogden, 2017). Public health discourse widely applies the term 'epidemic' in describing obesity trends (Butland et al., 2007, p. 17). Medical treatments are advancing, policy interventions increasingly target the 'obesogenic environment', and effective prevention strategies are known to combine an active lifestyle with a diet; but for any of these approaches to deliver sustained weight management, individual behaviour change is fundamental.

Like other preventative health behaviours, maintaining a healthy exercise and dietary regimen involves intertemporal choices: tradeoffs between satisfying our current desires and achieving our longer term objectives (Rogers, Milkman, & Volpp, 2014). Left to manage ourselves, too often we

DOI: 10.4018/IJABE.2018100101

privilege the near-term payoff over future wellbeing (O' Donoghue & Rabin, 1999). Such 'present bias' can explain the failure to follow up on good intentions with concrete action (Liu, Wisdom, Roberto, Liu, & Ubel, 2014), and has been linked empirically to weight management issues (Fan & Jin, 2013). A potential solution is pre-commitment strategies, or commitment devices, that bind our future selves to certain behaviours (Strotz, 1955; Thaler & Shefrin, 1981). Such strategies have been shown to promote smoking cessation (Gine, Karlan, & Zinman, 2010; Halpern et al., 2015), exercise (Prestwich et al., 2012; Royer, Stehr, & Sydnor, 2015) and a switching to safer water sources (Inauen, Tobias, & Mosler, 2014).

Where commitment devices have been tested as weight management aids they have shown some promising results.[1] Financial commitment devices in the form of a deposit contract stake money on achieving a weight loss outcome have been shown to have positive short-term effects (Volpp et al., 2008). An alternative form of commitment device relies on reputational rather than monetary stakes, and these too have been shown to promote weight loss: Nyer and Dellande (2010) report improved weight loss when goals are posted to a public gym noticeboard; Prestwich et al (2012) report greater weight loss when exercise is planned with a partner rather than as a solo activity. However, other studies have shown more mixed results from planning and commitment strategies (Chapman, Campbell, & Wilson, 2015; Verhoeven, Adriaanse, Ridder, Vet, & Fennis, 2013). Questions remain around the impact of reputational commitment devices.

What is striking is that none of these studies are set in the context of digital interventions, defined as programmes that promote health via a digital platform such as a website or smartphone (Alkhaldi, Hamilton, & Murray, 2016, p. 1). Recent studies highlight the considerable potential of such tools to improve health, and their increasing application to a wide range of health issues including weight management (Alkhaldi et al., 2016; Murray et al., 2016). A gap exists however in our understanding of how commitment devices may promote weight loss attempts supported by digital interventions.

This paper contributes to the scholarly debate with a novel test of how a reputational commitment device affects weight outcomes in a digital health setting. A randomised control trial introduced an additional reputational commitment for some online users of a weight management service. The comparison group comprised fee-paying clients. A second group of participants were asked to nominate an independent referee (a 'coach') who could verify their progress, thereby adding a reputational commitment element to their weight loss endeavours. A third group were offered a refund on their monthly subscription fee.

The results challenge the idea that increasing commitment boosts health outcomes: participants in the reputational commitment group reported lower weight loss over a 12-week period. The discussion considers plausible explanations for the results, and speculates that the negative effects of the reputational commitment device may indicate some form of commitment 'overload' that reduced weight loss performance. Avenues for future research around this potential explanation are raised.

## BACKGROUND

### A Dual-Self Framework for Health Behaviour Change

A commitment device is broadly defined as "an arrangement entered into by an individual with the aim of helping fulfil a plan for future behaviour that would otherwise be difficult owing to intrapersonal conflict" (Bryan, Karlan, & Nelson, 2010, p. 672). Such intrapersonal conflict might arise from a lack of self-control and competing internal preferences, for example those modelled by Thaler and Shefrin (1981) in their planner-doer framework. Each individual is understood to be composed of dual sub-selves, a far-sighted planner and a myopic doer, who are locked in an internal tussle over how to maximise utility. The doer sub-self cares only for current utility, while the planner sub-self aims to maximise life time utility. Considering the individual through a dual-self lens sets up a plausible battle between immediate and delayed gratification.[2]

Thaler and Shefrin applied their planner-doer framework to time inconsistent behaviour around economic choices. This paper extends their framework to preventative health behaviours. Like savings behaviours, preventative health behaviours require some investment today – time, money, effort, leisure or pleasure foregone – in order to reap the benefits many years from now. People make plans accordingly, but when the moment comes to adopt that preventative behaviour, people find themselves deviating from their plan, because in that moment the costs of doing so loom large relative to the benefits. A gap opens up between intentions and actions (Rogers et al 2015), with adverse consequences for long run health and wellbeing.

The outcome of these internal battles is manifest in health outcomes, for example weight or body mass index (BMI) status, which reflect whether the individual behaves in a way that promotes long term health benefits or succumbs to short term desires. Put differently, health outcomes can reflect whether the doer or planner has dominated decision making, since where a doer dominates, the individual is more likely to exhibit present bias, by over emphasising immediate gains over longer term payoffs (Loewenstein, Asch, Friedman, Melichar, & Volpp, 2012; O' Donoghue & Rabin, 1999). Health behaviours thus provide fertile ground for investigating time inconsistency and commitment devices, because they implicitly rely on intertemporal choices.

The remainder of this section briefly formalises two propositions arising from the planner-doer model. They draw on the intuition of the original model but apply it for the first time to health behaviours for weight loss.

**Proposition 1:** A short-sighted doer sub-self will over-consume, and the planner will seek commitment strategies.

If left to his own devices, the doer sub-self will choose to consume as much as possible within a given budget constraint up to some self-limiting satiation point (Equation 1), yielding actual consumption of $C_D$. This may exceed the ideal consumption identified by the planner sub-self, who takes account of current utility but seeks to maximise long run utility (Equation 2). The planner's benchmark may rely on some external norm, such as the recommended daily net intake of 2000/2500 kcal for women and men respectively which produces an optimal consumption level $C_P^*$.[3] Given the simple utility functions for the planner and doer, actual consumption exceeds the planner's preferred consumption $\left(C_P^* < C_D\right)$, this implies $V < 0$ and therefore at any given time $U_P^* < U_D$. The planner sub-self is unsatisfied with the status quo. Over time, if $C_P^* < C_D$ on a sustained basis, the individual is very likely to have a higher weight than the planner would prefer:[4]

$$U_D = f\left(C_t\right) \tag{1}$$

$$U_P = U_D + V \tag{2}$$

$$V = \sum_{n=t}^{T} \frac{\left(C_P^* - C_D^*\right)}{\left(1+r\right)^n} \tag{3}$$

The divergence between the doer's actual consumption $C_D$ and the planner's preferred consumption $C_P^*$ captures precisely the problem of intra-personal conflict. The planner would then want to change the doer's behaviour and achieve an optimal consumption level that maximises long-

term utility $U_P$. This is what predicts the planner's demand for a commitment device to lock in the doer's actions on a preferred consumption path.

**Proposition 2:** A commitment device acts as a tax on consumption and curbs the doer's over-consumption.

A commitment device generates some influence on the doer's actions, which can be represented by $\theta$: a "preference modification parameter" (Thaler & Shefrin, 1981, p. 395). This parameter is incorporated in Benabou and Pycia (2002) as a "reduced form representation of more concrete incentives (rewards, punishments) or rules put in place by the planner" (p. 422), but there is little beyond these broad definitions to explain how $\theta$ arises and its psychological underpinnings. I contend that $\theta$ represents the commitment device as a 'tax' applied by the planner on the doer.

In the case of a deposit contract this is a straightforward monetary penalty if the health outcome is not secured. Alternatively, a reputational commitment device applies a "psychological tax" (Miller & Prentice, 2013, p. 303) that affects self-respect, self-esteem or public image. As with a monetary tax, if the individual does not act in a way that is consistent with their goal, the costs of excess consumption increase.[5] In both cases, $\theta$ generates penalties for excess future consumption and in this way can alter the doer's future choices. Ideally, $\theta$ will reduce the doer's consumption such that it aligns with the planner's preferred benchmark, so that $C_D = C_P^*$. This behaviour change, in turn, yields the desired health outcome of weight loss towards the target weight. In essence, the model predicts that a commitment device will exert positive treatment effects on health behaviours and weight loss.

What is the basis for explaining $\theta$ as a psychological tax that arises from reputational concerns? Commitment has been defined as the "pledging or binding of the individual to behavioural acts" (Kiesler & Sakumura, 1966, p. 349). The commitment makes an act less changeable, because the individual wants to be consistent with what he has declared to others, and to avoid the personal and social disapproval that accompanies inconsistency. Parrott et al believe this lens helps explain "why the use of written and verbal pledges, promises, and contracts has increased compliance with various health care routines" (1998, p. 392), and find that the act of making a public commitment as part of a skin cancer campaign led to more people undertaking prevention and detection behaviours.

The negative psychological effects – the reputational costs – are what underpin such reputational commitment devices. While these may be characterised as "soft" penalties relative to the "hard" penalties of a financial commitment device (Bryan et al., 2010), they nonetheless can change incentives to the point where planner's and doer's interests align and positive behaviour change arises (Benabou & Tirole, 2004). In the weight loss sector, public pledges are used in various guises. Public weigh-ins at a weight loss group serve to hold the individual to account against their stated target; attendance at an exercise club might be encouraged through a promise to a team; and pledge boards are a common feature at gyms. Further, pledges do not have to be very public to be effective. Even brief dialogue or written correspondence with a general practitioner can encourage greater participation in NHS weight loss programmes, linked to a sense of commitment between patient and doctor (Allen, Cohn, & Ahern, 2015; Aveyard et al., 2016). Making the commitment an external one – even if it to just one other person – has the power to inspire a sense of accountability and create powerfully-felt reputational stakes that spur behaviour change.

In summary, the planner-doer framework is more than mere "metaphor" (Frederick, Loewenstein, & O'Donoghue, 2002). The challenge of making the right health choices in the context of intertemporal trade-offs and behavioural biases have been shown to have empirical significance (Fan & Jin, 2013; Ruhm, 2012). Many individuals understand the importance of losing weight and achieving a healthy body mass index (BMI), intend to take action to achieve their desired BMI, yet struggle with staying on track with their goal. This is precisely where a commitment strategy is theorised to support

behaviour change by going "beyond good intentions" (Rogers, Milkman, John, & Norton, 2015). Taking up a commitment device generates short run costs to deviating from the health goal, essentially a "psychological tax" (Miller & Prentice, 2013, p. 303), and in this way the planner sub-self is able to rein in the wayward doer. The intertemporal shifting of benefits and costs strengthens the incentives to do the right thing at the present time. Theory thus implies that a commitment device will promote behaviour change and weight loss.

As highlighted earlier, while there is some evidence that bears out this prediction for financial commitment devices, in the context of non-digital health settings, there is little testing of theory in the increasingly important eHealth context using reputational commitment devices.

## STUDY SETTING AND INTERVENTIONS[6]

The field experiment was nested within an online weight management service in the UK that charged clients an average £5 monthly subscription.[7] The service included access to a calorie counter tool, a food journal, online discussion forums, healthy eating advice, and a self-reported weight tracker. Based on initial weight, ideal weight and the timeframe for achieving it, users calculated their appropriate net energy intake, which served as a personal daily target. The website and app allowed users to keep track of their food intake and exercise, and benchmark their net energy intake against their personal target. The registration process ensured all users gained the same basic information of what levels of consumption would be in line with their stated health goals. Signing up to the service implied a degree of initial motivation to improve health status: as well as the effort of completing the brief registration, users were paying for a calorie counting service when free alternatives were available and easily accessible through a web or app search.

Eligible participants were paying clients of the service.[8] In total, 364 participants were recruited over July to November 2013.[9] The trial was marketed through emails, social media and an online post visible to members who had logged in to the website. Taking part in the survey and trial was entirely voluntary, not linked to any aspect of the service, and no financial incentives were offered. All participants were requested not to discuss the study with other members.[10] After completing a baseline survey, participants were randomly allocated to one of the three experimental groups (see Figure 1).[11]

### Experimental Groups

The comparison group completed the baseline survey and continued with the Food Monitor service as usual (n=145), including paying the monthly fee.

The first intervention asked users to nominate a 'coach', a friend or family member who was aware of and supported their weight management goals, and who could be contacted at the end of the four week period to verify weight loss progress (n=118, 41% compliance). Naming a coach was expected to create an element of reputational commitment in order to stay on track with their weight loss goals. By making a weight loss goal publicly known to another person, the treatment was designed to create the 'psychological tax' to reneging on that goal. In this way, the planner sub-self locks in 'good' behaviours such as self-monitoring, improved diet and physical activity, that the doer sub-self might otherwise abandon. Treatment design is also in line with existing research on 'supportive accountability', where a trusted coach can improve adherence to behaviour change goals in the context of e-health interventions (Mohr, Cuijpers, & Lehman, 2011). The experiment therefore allowed for testing of the hypothesis that a reputational commitment device would promote weight loss.

**Hypothesis 1:** The coach group will experience higher weight loss than the comparison group.

A second intervention offered a refund of one month's subscription fee (n=101, 95% compliance). The fee refunds ranged from £0 to £7.99, with a mean of £5.21.[12] The aim was to assess

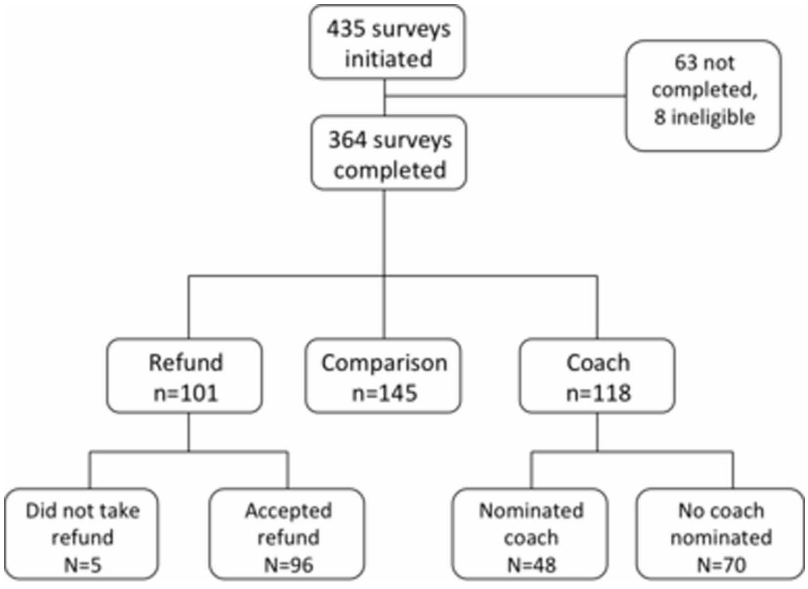
the impact of a temporary reprieve from the fee on weight loss outcomes. This intervention provided an opportunity to test whether the monthly fee serves as a financial commitment device. Setting up an upfront payment for a service for which free alternatives are available may be an example of a strategic behaviour to influence one's future choices, in other words a form of commitment device. However, unlike the more conventional deposit contract design, there is no money at stake. The study therefore aims to test whether the monthly fee exerts a similar influence as a financial commitment device. If temporarily removing it makes a difference to weight loss outcomes, this would indicate it is a commitment device.

**Hypothesis 2:** The refund group will experience lower weight loss than the comparison group.

Participants were invited to sign up at a time of their choosing, so it was a recognised limitation that the full sample was not available to stratify or randomise in advance, and it was anticipated that there might not be perfect balance in numbers or baseline characteristics across the experimental groups. The trial was designed around three experimental groups, with a quota imposed on the number of refunds. Once this quota (of 100 refunds) was reached, random allocation of participants to the refund group was stopped. Recruitment continued for the comparison and coach groups until the sample size target was achieved (see appendix for details of sample size calculations). For the purposes of analysis, the period before and after the quota are referred to as phase 1 and 2, but in practice they ran on consecutively with no break in the trial. Analysis takes full account of the two phases to estimate separate causal effects from the coach and refund treatments. [13]

Participants were not blinded to their own treatment status as the aim of the experiment was to identify whether the change in commitment affected behaviour. Participants were not made aware of the possibility of other treatments. The researcher had to be aware of treatment status in order to trigger the monthly refund for participants assigned to the limited commitment group, and to initiate follow up with coaches. The partner firm was made aware of treatment status for those clients who were due a refund, which was delivered through their administrative systems. As Figure 1 demonstrates, 364 participants were enrolled, of which 95% complied with the refund treatment and 41% with the coach treatment. Take up of treatments is discussed in further detail below.

**Figure 1. CONSORT flowchart**



## Study Sample and Participant Characteristics

Participants' average weight was 83 kg, ranging from 48 kg to 157 kg. Average body mass index (BMI) was 30.2. The majority of participants were either overweight (34%) or obese (35%), with a small minority who were severely obese (9%). Amongst the 97.5% of participants who were aiming to lose weight over the following four weeks, the average weight loss goal was 3.5 kg (just under 1 kg a week). As is typical of weight loss studies, the majority of the sample (90%) were women (Jolly et al., 2011).[14] Most participants (87%) reported taking part in other activities besides using the online service in order to meet their weight loss goals; and cited on average three exercise sessions per week. The majority of participants reported they did not have children at home (61%). Table 1 summarises characteristics across experimental groups.

Baseline covariates were largely well-balanced across experimental groups, however it is common to find some chance imbalances in any experiment where a series of baseline covariates are tested, and particularly in smaller samples (Glennerster & Takavarasha, 2013; Torgerson & Torgerson, 2008). Starting weight and initial BMI vary between the groups due to chance differences in the share of severely obese individuals across groups. The analysis takes account of this in two ways. Firstly, BMI categories are included as control variables in all regressions reported below. Secondly, regression analysis is run on the sample excluding the severely obese (n=33), and reported along with other robustness checks in the appendix to confirm there is no change to the overall findings.

## Take-Up of the Reputational Commitment Device

Five participants failed to initiate the refund process even after prompting. Of the 118 participants offered the coach treatment, 41% nominated a coach who could be contacted after four weeks. Regression analysis investigates the characteristics of those who complied with the treatment (see Table 2). Of the BMI categories, obese participants were more likely to nominate a coach. Women were more likely to nominate a coach than men ($p = 0.004$). Those reporting the most short-term health attitudes were most likely to take up the coach offer ('hedonistic immortals'). Younger participants were less likely to nominate a coach than those aged 30 or over, and participants with children were less likely to nominate a coach. In terms of lifestyle, those reporting higher exercise levels were less likely to nominate a coach. Those joining the trial in August were more likely to nominate a coach. Among background characteristics, participants with an undergraduate degree and the self-employed were more likely to nominate a coach. There were no differences based on weight loss target.

Non-compliers were asked why they were turning down the intervention offer. Often the decision was driven by not wanting to share the coach's details (30% of respondents), which was a necessary part of treatment design; or because people did not want to share their goal with anyone else (29%), raising useful insights into the personal nature of the weight loss endeavours for a significant proportion of respondents. Some could not think of a coach to name (24%), and a small minority felt it was too much effort for them or their coach (3%). The remainder of participants (14%) gave other reasons for declining the coach treatment offer. These broadly included wanting to rely on themselves ("I could ask but prefer to do [it] on my own", "[it's] up to me to get my head round healthy eating [and] healthy lifestyle"), or constraints to getting a coach ("I did try to get someone to support me but they weren't interested enough"). One person reported they already had a coach.

The implications of non-compliance are discussed in further detail below. For analysis purposes, it will be important to examine both the average treatment effect, based on the randomised offer of treatment, and the complier average causal effect, amongst a subset of those who took up the treatment (Gerber & Green, 2012).

## Outcome Data

Weight outcomes at 12 weeks were gathered using self-reported data. Participants were encouraged to use the service as they normally would, meaning that weigh-ins were not mandatory. The trade-off for ensuring a more realistic experimental setting was a reliance on participants returning to the

**Table 1. Summary statistics by experimental group at baseline**

| | Comparison N = 145 (1) | Coach N=118 (2) | Refund N=101 (3) |
|---|---|---|---|
| Starting weight (kg) | 84.9 (1.72) | 83.4 (1.68) | 79.4* (1.58) |
| BMI | 31.1 (0.606) | 30.2 (0.582) | 28.9* (0.543) |
| Overweight | 0.368 (0.040) | 0.291 (0.042) | 0.35 (0.048) |
| Obese | 0.319 (0.039) | 0.402 (0.046) | 0.34 (0.048) |
| Severely obese | 0.139 (0.029) | 0.085 (0.026) | 0.03* (0.017) |
| Weight loss target (%) | 4.13 (0.178) | 4.27 (0.186) | 4.18 (0.155) |
| Female (%) | 0.903 (0.025) | 0.864 (0.032) | 0.921 (0.027) |
| Fruit and veg daily intake | 4.13 (0.186) | 3.85 (0.197) | 3.78 (0.206) |
| Exercise sessions per week | 3.11 (0.226) | 2.95 (0.305) | 3.32 (0.256) |
| Experienced major life changes recently (%) | 0.290 (0.038) | 0.347 (0.044) | 0.376 (0.048) |
| Doing other activities to lose weight (%) | 0.883 (0.27) | 0.872 (0.031) | 0.842 (0.037) |
| Children at home | 0.710 (0.083) | 0.667 (0.089) | 0.663 (0.088) |
| Recruited in phase 1 | 0.731 | 0.72 | 01.00*** |
| Health attitudes (modal) | Unconfident fatalist | Unconfident fatalist | Unconfident fatalist |
| Age (modal) | 40-49 years | 50-59 years | 40-49 years |
| Education (modal) | Bachelors | Bachelors | Bachelors |
| Income (modal) | Up to £19k | £40k - £49k | £50k - £59k |
| Job status (modal) | Paid employment | Paid employment | Paid employment |

Notes: Mean values reported with standard errors in parentheses for all continuous and binary variables. Modal categories reported for age, education, income and job status categorical variables. N=361 for starting weight and BMI variables; n=363 for other activities, children and job; n=362 for education; n=357 for age; n=354 for weight loss target; n=345 for income; n=364 for all other variables. Columns 2 and 3 indicate if treatment group has any significant difference in mean values compared with comparison group, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Where significant differences are detected (column 3), further checks using Bonferroni and Benjamini-Hochberg correction techniques for multiple hypothesis testing indicate these differences are no longer statistically significant.

website to share outcome data and the accompanying risk of attrition. It seemed reasonable to assume that clients would log back on to provide weight readings, given they were voluntarily paying for the service. Data on 3,224 weight entries was gathered during the trial, but with uneven coverage: 98 participants did not enter an eligible weight reading, while one participant recorded their weight 34 times. Weigh-ins were recorded if they fell within 14 days before or after the target date, with the closest date match preferred.[15] The dataset recorded 12-week weight readings for 162 participants, implying 55% attrition. Strategies for addressing attrition are discussed below.

**Table 2. Take up of the reputational commitment device**

| | |
|---|---|
| Starting weight (kg) | -0.005 (0.020) |
| Overweight | 1.773** (0.671) |
| Obese | 0.435 (0.767) |
| Severely obese | 0.574 (1.451) |
| Weight loss target (%) | -0.020 (0.118) |
| Female | 1.783** (0.617) |
| Age: 30s | 2.049** (0.758) |
| Age: 40s | 2.409** (0.808) |
| Age: 50s | 2.662** (0.864) |
| Age: 60s | 3.840*** (0.984) |
| Fruit and vegetable daily intake | 0.063 (0.081) |
| Exercise sessions per week | -0.173** (0.065) |
| Experienced major life changes recently (%) | -0.721 (0.503) |
| Doing other activities to lose weight (%) | 0.610 (0.724) |
| Number of children at home | -0.528* (0.265) |
| Health attitudes: 'live for today' | -1.769* (0.769) |
| Health attitudes: 'unconfident fatalist' | -2.296** (0.747) |
| Health attitudes: 'health conscious realist' | -2.050** (0.758) |
| Health attitudes: 'balanced compensator' | -2.140* (0.965) |
| Start month: August | 1.047* (0.495) |
| Start month: September | -0.149 (0.654) |
| Start month: October | 0.131 (0.644) |
| Start month: November | 0.795 (0.844) |
| Education: GCSE | 0.951 (0.795) |
| Education: A level | 1.044 (0.840) |
| Education: Bachelors degree | 2.789*** (0.823) |
| Education: Masters degree | 0.287 (0.780) |
| Education: Doctoral degree | 0.497 (1.115) |
| Income: £20,000 to £29,999 | 0.639 (0.685) |
| Income: £30,000 to £39,999 | 0.000 (.) |
| Income: £40,000 to £49,999 | -0.635 (0.839) |
| Income: £50,000 to £59,999 | -0.335 (0.946) |
| Income: £60,000 to £69,999 | -0.466 (0.933) |
| Income: £70,000 to £99,999 | 0.514 (0.820) |
| Income: £100,000 to £149,999 | 0.325 (0.780) |
| Income: over £150,000 | -2.376 (1.358) |
| Job: self-employed | 1.841** (0.684) |
| Job: student | 1.796 (1.268) |
| Job: looking for a job | 0.000 (.) |
| Job: not employed and not looking | 0.372 (0.651) |
| Observations | 92 |

Notes: robust standard errors in parentheses clustered at individual level, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, health attitudes base category 'hedonistic immortal', start month base category 'July', education base category 'no qualifications', income base category less than £19k, job base category 'paid employment'.

## Statistical Model

Regression analysis aimed to uncover the average treatment effect, or intent-to-treat estimate, and includes all participants with end weight readings. The statistical model included a set of explanatory variables with robust standard errors clustered at the individual level and estimated treatment effects for the two interventions separately.[16] Equation 1 estimated the causal impact of the coach treatment, and Equation 2 the causal impact of the refund treatment, on end weight outcomes:

$$Y_i = \alpha + \beta^1 \cdot Coach + \sum_{j=1}^{J} \gamma_j \cdot W_{ij} + \varepsilon_i \quad (1)$$

$$Y_i = \alpha + \beta^2 \cdot Refund + \sum_{j=1}^{J} \gamma_j \cdot W_{ij} + \varepsilon_i \quad (2)$$

*Y* measured the outcome variable post-test weight. Treatment status was captured by dummy variables 'coach' and 'refund'. The OLS estimator for $\beta^1$ was expected to recover the average treatment effects for the reputational commitment device, and in line with the hypothesised relationship was expected to have a positive sign. W comprised a series of baseline covariates J, with coefficients $\gamma_j$, including variables to control for (1) initial health status, captured by starting weight and BMI category, and weight loss target; (2) usage of the online service during the four-week treatment period; (3) individual traits, covering gender, age, health attitudes and outlook, with the latter captured through a segmentation model creating five sub-groups (Williams, McVey, Davies, & MacGregor, 2011); lifestyle, based on exercise sessions per week, undertaking other activities to lose weight, number of children at home, if any major life changes were experienced recently); and (5) seasonal effects, proxied by the month in which the participant joined the trial. Regressions were run on a sample of participants aiming to lose (not maintain) weight.[17]

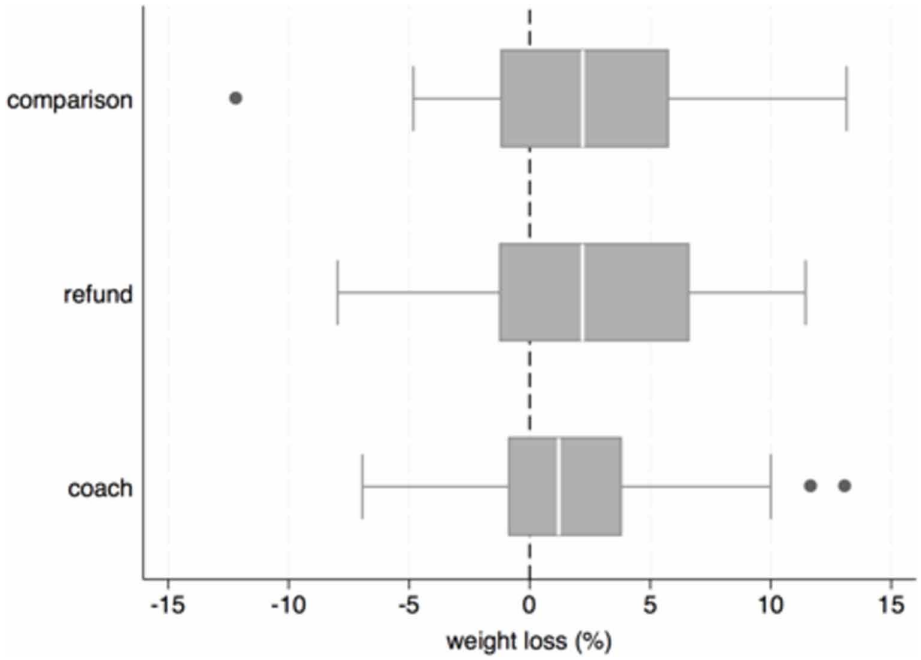
Treatment status was not significantly associated with attrition (see appendix). Not being observed at 12 weeks was positively associated with initial level of reported exercise ($p < 0.01$) and usage of the service over the first four weeks of the trial ($p < 0.001$). These variables may signal pre-existing motivation to use the service, and explain why these clients were less likely to attrite. Age was a milder predictor of attrition, with those in 40s more likely to attrite by 12 weeks ($p < 0.05$).

While there was no evidence of differential attrition based on treatment status, attrition remained an issue given the loss of sample size and potential for unobserved heterogeneity across those who were observed and those who were not. Robustness checks therefore incorporated inverse proportionality weights (IPW) to model the probability of being observed at the end of the trial using covariates. This approach assumes that data is missing independent of potential outcomes and conditional on covariates (Gerber & Green, 2012, p. 222). Results were also triangulated with the non-parametric technique of placing bounds around the treatment effect (Augurzky, Bauer, Arndt, Schmidt, & Tauchmann, 2012; Tauchmann, 2014); the Lee Bounds estimator offered an interval of values based on a scenario of "worst case sample selection biases" (Lee, 2002, p. 12).[18, 19]

## RESULTS

All experimental groups lost weight over 12 weeks on average. Figure 2 examines weight loss as a percentage of initial weight. The refund did not appear to affect performance: the refund group lost on average 2.4% of initial body weight compared to the 2.2% average weight loss reported in the comparison group. The coach group, surprisingly, reports the slowest rate of weight loss with an average of 1.1%. Regression analysis confirms these findings (see Table 3).

**Figure 2. Weight loss (%) over 12 weeks by experimental group**




The evidence refutes the hypothesis that the reputational commitment device would promote weight loss (see Table 3 panel A). The table reports the average treatment effect (intent-to-treat) of a coach were weight readings 1.5 kg higher than the comparison group (p=0.032, effect size -0.32). Alternative model specifications and robustness checks corroborate the conclusion that the reputational commitment device exerted a negative treatment effect on weight loss including applying inverse proportionality weighting to address attrition, and testing a binary dependent variable 'weight loss target met' (see appendix).[20] The Lee Bounds approach identifies that the average treatment effect of the reputational commitment on percentage weight loss lies in an interval containing negative values (see appendix), further corroborating the finding of a negative treatment effect.

The evidence also refutes hypothesis 2, as the refund made little difference to weight loss performance (see Table 3 panel B). Excluding the five participants who did not take up their refund does not alter the findings, and other robustness checks (see appendix) corroborate the lack of treatment effect from the refund.

Covariates reported in Table 3 do not allow for causal inference, nevertheless they offer some interesting insights. Starting weight is a strong predictor of end weight reading, as is usage of the online service during the first four weeks of the trial. Having children at home is associated with lower end weight readings. Some seasonal effects are detected in panel A, with participants who started the trial in August and November more likely to report lower end weight, which may reflect increased challenges with weight management during the festive season (the time period when participants who started in September and October were reporting their 12-week weight outcomes).

## DISCUSSION

### Why Does Increased Commitment Exert a Negative Effect on Weight Management on Average?

Participants in the coach group reported end weight outcomes 1.5 kg higher than the comparison group (p=0.032, effect size -0.32), indicating slower weight loss. I speculate on three possible reasons for this outcome. Firstly, low treatment intensity may have led to underestimating the potential positive

**Table 3. Average treatment effects of commitment devices on weight outcomes**

| | |
|---|---|
| **Panel A: The effect of the reputational commitment device on end weight (kg)** | |
| Coach | 1.530* (0.704) |
| Starting weight in kg | 0.978*** (0.041) |
| Weight loss target (%) | -0.180 (0.177) |
| Overweight | 0.452 (1.019) |
| Obese | -1.433 (1.527) |
| Severely obese | -0.837 (2.580) |
| Female | 1.665 (1.472) |
| Age 30-39 | 0.776 (1.092) |
| Age 40-49 | 1.488 (1.152) |
| Age 50-59 | 0.171 (1.211) |
| Age 60-65 | 0.579 (1.521) |
| Health attitudes: 'live for today' | 1.044 (1.680) |
| Health attitudes: 'unconfident fatalist' | 2.160 (1.199) |
| Health attitudes: 'health conscious realist' | 0.488 (1.292) |
| Health attitudes: 'balanced compensator' | 2.105 (1.795) |
| Usage of service over 4 weeks | -0.132*** (0.025) |
| Exercise sessions per week | 0.195 (0.120) |
| Recent life changes | 0.704 (0.826) |
| Other activities pursued to lose weight | 0.873 (1.120) |
| Children at home | -0.764* (0.375) |
| Starting month August | -1.855 (1.144) |
| Starting month September | -1.560 (1.208) |
| Starting month October | -2.003 (1.356) |
| Starting month November | -2.556 (1.437) |
| Observations | 104 |
| $R^2$ | 0.978 |
| **Panel B: The effect of a refund on end weight (kg)** | |
| Refund | 0.252 (0.804) |
| Coach | 1.063 (0.911) |
| Starting weight in kg | 0.979*** (0.047) |
| Weight loss target (%) | -0.270 (0.177) |
| Overweight | 1.049 (0.981) |
| Obese | -0.630 (1.465) |
| Severely obese | 0.316 (2.658) |
| Female | 2.951 (1.513) |
| Age 30-39 | 1.053 (1.019) |

*continued on following page*

**Table 3. Continued**

| | |
|---|---|
| Age 40-49 | -0.488 (1.147) |
| Age 50-59 | -0.766 (1.053) |
| Age 60-65 | -1.332 (1.417) |
| Health attitudes: 'live for today' | 1.721 (1.659) |
| Health attitudes: 'unconfident fatalist' | 2.752* (1.131) |
| Health attitudes: 'health conscious realist' | 2.642* (1.227) |
| Health attitudes: 'balanced compensator' | 3.325* (1.513) |
| Usage of service over 4 weeks | -0.093*** (0.027) |
| Exercise sessions per week | 0.166 (0.137) |
| Recent life changes | 0.507 (0.709) |
| Other activities pursued to lose weight | 1.403 (0.975) |
| Children at home | -0.720* (0.335) |
| Starting month August | -0.472 (0.967) |
| Starting month September | -0.729 (1.174) |
| Observations | 122 |
| $R^2$ | 0.972 |

Notes: OLS regression on sample of complete cases of participants who aimed to lose weight. Panel A uses full length of trial to recover ATE on reputational plus financial commitment device, comparing coach and comparison groups. Panel B uses phase 1 sample only and recovers ATE on limited commitment treatment, comparing refund and comparison groups. Robust standard errors in parentheses, clustered at individual level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

effects of a reputational commitment device amongst active compliers. Secondly, the involvement of a coach may have entailed a substitution of accountability and self-monitoring practice away from the digital tools, undermining weight loss efforts. Thirdly, the take up of additional, reputational commitment to achieve already challenging health goals may have created a sense of 'overload', with participants feeling less motivated and less willing to absorb the short-term trade-offs involved in achieving significant weight loss. Each explanation is discussed in turn.

One way to account for this surprising finding is that the treatment was not applied as intended, and that a combination of weak adherence and low compliance has underestimated the theorised impact of a reputational commitment device. Qualitative follow-up with participants indicates moderate to low adherence to the treatment and variation in the way coaches took on the role of 'supportive accountability': just over half of coaches reported they had been actively involved in the participant's weight loss efforts, and three were unaware they had been nominated. [21] Amongst the coaches who reported being actively involved almost half were not sure whether the individual had met their weight loss target, again indicating that the extent of follow-up and accountability varied across the coaches.[22] Perhaps if more coaches had been actively involved, this may have strengthened the overall intensity of treatment, generated a greater 'psychological tax' from the commitment device, and leveraged stronger treatment effects.

Take-up of the reputational commitment device, at 41%, may reflect a mismatch between the more public nature of the intervention (involving another individual in the weight management goal) and the participant's revealed preference for online weight loss tools (which, in comparison, are relatively private). Low compliance might itself be explained by the design mismatch of the online digital health study setting and the offline reputational commitment intervention.

However, if weak treatment intensity was the nub of the issue, we would expect the complier average causal effect (CACE) to be positive and significant for the sub-set of people

who accepted the reputational commitment treatment.[23] In fact CACE estimates reveal a more negative average treatment effect: those who nominated the coach recorded an end weight 4.4 kg higher (p=0.028). In other words, compliers fared worse than those who were offered the treatment but declined. While low compliance and adherence may have broadly diluted the effects of the commitment device, it does not explain the negative and statistically significant effects on weight loss (see appendix).

A second explanation is that participants who nominated a coach were more likely to take their self-monitoring behaviour offline, and this transfer of weight loss accountability from the digital tools to a human coach may have eroded weight loss effectiveness. However, no differences are detected in usage of the digital self-monitoring tools based on whether a coach was nominated in the reputational commitment device group. Both compliers (who named a coach) and non-compliers used the calorie counter tool 15 times on average over the 4 week coach accountability period, and approximately 33 times during the 12 week trial overall (p = 0.9 in both cases). This evidence suggests that substitution effects away from the digital weight management tools towards the coach are not an adequate explanation of the negative treatment effect.

An alternative explanation might be that participants experienced some form of commitment 'saturation' or 'overload', implying some form of threshold beyond which more commitment has either no or adverse effects. The literature, and the model presented in this paper, do not examine the possibility or implications of threshold effects, so this discussion is necessarily speculative. However, in the absence of other explanations for why the coach group under-performed, this is one possible explanation that needs further investigation. It raises new research questions around the optimal level of commitment that motivates action; and how commitment overload generates adverse consequences.

Although the active reputational commitment period, when the coach was required to be involved, was the first four weeks of the trial, the divergence in weight management trajectories was evident in the 12-week outcomes. If commitment overload was driving the weaker weight management performance, this would indicate the effects were felt after the nomination and accountability period ended. This raises additional questions, for example, to what extent commitment devices interact with subjective wellbeing to create either a positive spiral of progress, or a negative spiral of demotivation; and is overall wellbeing linked to the possibility of experience commitment overload? The welfare effects of commitment devices are under-studied, and theory gives no clear indication of the impacts (Bryan et al., 2010, pp. 679, 693). The empirical evidence here raises the importance of wellbeing and commitment for future research.

## What Does the Null Effect on the Refund Treatment Tell Us About Financial Commitment Devices?

Hypothesis 2 tested whether the act of paying a fee for a service that could be otherwise obtained for free was in itself a form of commitment device. The payment, however modest, was hypothesized to be a strategic interaction with oneself, creating an incentive to use the weight loss service in order to make the most of having already paid for it. Under this logic, a refund would represent a (temporary) dismantling of the financial commitment, and would therefore lead to lower weight loss. The results refute the hypothesis, finding instead no difference in weight loss at either 4 or 12 weeks between those who continue to pay as usual and those who received the refund. Robustness checks examine whether the size of the refund matters, and any impact of the five participants who did not take up a refund (see appendix); in both cases the results are unchanged. The findings undermine the idea of membership fees as a strategic interaction with oneself, and argue against it being a financial commitment device. Rather, the willingness to pay is simply a reflection of initial motivation rather than a strategy to rein in one's own time inconsistency.[24]

## What Does the Study Imply for the Use of Commitment Devices in Digital Health Interventions?

Most studies of commitment devices for weight loss take place in contexts where people who aim to lose weight interact face to face with others such as health professionals. For example, Nyer and Dellande's (2011) study is based at a health centre in India, while Volpp et al (2008) involve frequent contact between participants and staff at a US medical research centre. As noted earlier, the context for weight management aids is rapidly evolving with growing use of digital health tools to support lifestyle change, and this study raises two issues of note. Firstly, the impact of these digital interventions is determined partly by user engagement and adherence (Alkhaldi et al., 2016; Johnson & Wardle, 2011), yet attrition and non-use of interventions is frequently identified as a challenge (Arem & Irwin, 2010). Staying committed to digital health tools is a valid concern and a promising area for commitment devices to make a difference, but the results of the study reported here suggest that offline commitment devices may not be the most effective way to ensure adherence to online interventions: there is no significant difference in usage of the calorie-counter tool between the coach group and other participants during the experiment[25]. Secondly, the study also raises questions about the feasibility of combining online and offline weight management support, with a sizeable minority suggesting they did not want to share their goal with anyone else or they could not thinking of someone to provide that coaching support (31% of participants offered a coach).

## Study Limitations

External validity is a frequent critique of field experiments (Deaton and Cartwright, 2016), and it is worth noting the sample here is drawn from a pool of people who are sufficiently self-aware of their need for external support that they sign up voluntarily for a financial commitment device; and one that is digital in nature, offering relative anonymity and privacy on their weight loss journey. Further work is needed to establish whether reputational commitment devices can be effective in improving adherence in different settings, such as face-to-face weight loss groups and free weight management services. The trial is not unusual amongst weight loss studies in reporting attrition (Elobeid et al., 2009). While this has not affected the validity of the statistical analysis, the attrition rates reported in this trial underscore the need for additional mitigation strategies when relying on self-reported weight readings, for example through second round sampling (Gerber and Green, 2012). A further limitation of the study is the reliance on self-reported weight.

Field experiment methodology is valued for its ability to generate high internal validity and causal inference. The research design employed here is, however, limited when it comes to asking questions about causal mechanisms and explaining results that are contrary to the hypotheses (Lewin, Glenton, and Oxman, 2009; Hesse-Biber, 2013). A mixed methods strategy in future would allow for further probing of the statistical result – for example through follow-up data from participants on how they interpreted the treatments – and emerging propositions around commitment overload in order to refine and build on planner-doer theory and dual-self modelling.

Embedding qualitative analysis more fully in a field experiment design would allow for a deeper exploration of causal mechanisms to better understand how commitment devices can leverage stronger health outcomes, and to understand the circumstances in which commitment devices might have unintended negative consequences. Future research could usefully frame new questions around threshold effects and the optimal level of commitment that motivates action; the causal mechanisms potentially underpinning commitment overload; and other ways to encourage commitment to weight loss goals that are being pursued through digital health interventions. Qualitative evidence may offer novel opportunities to test the value of the planner-doer model in theorising the behaviour and internal tussles of those striving to lose weight; the interpretation of and adherence to the treatment; and an understanding of lived experiences for those who took up the coach intervention.

## CONCLUSION

Participants in the coach group reported end weight outcomes 1.5 kg higher than the comparison group ($p < 0.05$), indicating slower weight loss with an effect size of -0.32. The negative treatment effect was especially pronounced amongst those who took up the reputational commitment: compliers reported on average 4.4 kg lower weight loss than the comparison group ($p < 0.05$). The evidence raises the possibility of 'commitment overload' arising from the reputational commitment device. Those who accepted the treatment offer may have experienced this overload most intensely, which explains why the extent of weight loss was lowest in the group that complied with the treatment. Participants offered a refund on their monthly subscription to the online service showed no significant differences in their weight loss performance compared to those who continued with their monthly fees ($p > 0.05$). This finding argues against the idea that the act of monthly pre-payment for the online tools is a financial commitment device, it is more likely a reflection of initial motivation which is unaffected by a temporary reduction in payment for the service.

This paper presented an application of Thaler and Shefrin's (1981) planner-doer framework to health behaviours. While behavioural economics theory makes the case for reputational commitment devices having the influence to leverage behaviour change, the results reported here do not bear out this expectation. The findings contrast with those reported in other empirical tests of reputational commitment devices (Inauen et al., 2014; Nyer & Dellande, 2010; Prestwich et al., 2012); but chime with findings reported by Verhoeven et al (2013), who highlight the potential negative effects from an overload of planning: in a study where multiple plans to curb unhealthy snacking were less effective than a single plan with the same objective. In a follow up study, the authors suggest that neither initial motivation nor the act of making multiple plans explained the lack of effect on health behaviour change, but the process of enacting multiple plans became problematic. The results reported by Verhoeven et al. (2013) lends further credence to the idea of 'overload' applying to the enactment of commitment devices for health behaviours.

The concept of commitment overload is speculative at this stage but raises important questions for future research both in terms of theory development and empirical testing. For example, how might the planner-doer framework accommodate the possibility of commitment overload? What design features of a reputational commitment device are particularly important to bringing about behaviour change? How important are the characteristics of the 'coach', and what effect does their relationship with the individual have for the overall effectiveness of the reputational commitment strategy?

The findings of this article also contribute to digital health scholarship research. Speaking broadly about digital health tools, Kelly highlights that "the ways that people engage with the technologies, the mechanism of action, the human responses and the outcomes will be heterogeneous" (2016, p. 861); the same can be said of commitment devices and the combination of offline and online interventions as tested here. The negative effects from the offer of involving an offline coach in an otherwise online weight management effort somewhat challenge the ideas put forward by Mohr et al. (2011) that "the effectiveness of and adherence to eHealth interventions is enhanced by human support". Mohr et al call for empirical testing of their theoretical model, and the work presented here arguably offers relevant insights: firstly, the prospect of negative consequences for weight outcomes, and secondly the lack of effect on adherence to the digital health tool.

## ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. My sincere thanks to Peter John, Roland Kappe, David Torgerson and Albert Weale for their helpful comments and advice, and to three anonymous reviewers for suggestions on an earlier draft. Thanks are due to UCL School of Public Policy for supporting the research through a PhD studentship.

## REFERENCES

Alkhaldi, G., Hamilton, F. L., & Murray, E. (2016). The Effectiveness of Prompts to Promote Engagement With Digital Interventions: A Systematic Review. *Journal of Medical Internet Research*, *18*(1), e6. doi:10.2196/jmir.4790 PMID:26747176

Allen, J. T., Cohn, S. R., & Ahern, A. L. (2015). Experiences of a commercial weight-loss programme after primary care referral. *The British Journal of General Practice*, *65*(April), 248–255. doi:10.3399/bjgp15X684409 PMID:25824185

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Oxford: Princeton University Press.

Arem, H., & Irwin, M. (2010). A review of web-based weight loss interventions in adults. *Obesity Reviews*, *12*. doi:10.1111/j.1467-789X.2010.00787.x PMID:20804523

Augurzky, B., Bauer, T. K., Arndt, R. R., Schmidt, C. M., & Tauchmann, H. (2012). Does Money Burn Fat? *Ruhr Economic Papers*.

Aveyard, P., Lewis, A., Tearne, S., Hood, K., Christian-brown, A., Adab, P., & Jebb, S. A. et al. (2016). Screening and brief intervention for obesity in primary care: A parallel, two-arm, randomised trial. *Lancet*, *388*(10059), 2492–2500. doi:10.1016/S0140-6736(16)31893-1 PMID:27789061

Benabou, R., & Tirole, J. (2004). Willpower and Personal Rules. *Journal of Political Economy*, *112*(4), 848–886. doi:10.1086/421167

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics*, *2*(1), 671–698. doi:10.1146/annurev.economics.102308.124324

Butland, B., Jebb, S., Kopelman, P., McPherson, K., Thomas, S., Mardell, J., & Parry, V. (2007). Tackling Obesities: Future Choices – Project report. Government Office for Science.

Chapman, J., Campbell, M., & Wilson, C. (2015). Insights for Exercise Adherence From a Minimal Planning Intervention to Increase Physical Activity. *Health Education & Behavior*, *42*(6), 730–735. doi:10.1177/1090198115577374 PMID:25842384

Fan, M., & Jin, Y. (2013). Obesity and Self-control: Food Consumption, Physical Activity, and Weight-loss Intention. *Applied Economic Perspectives and Policy*, *36*(1), 125–145. doi:10.1093/aepp/ppt034

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, *40*(2), 351–401. doi:10.1257/jel.40.2.351

Fuller, E., Mindell, J., & Prior, G. (2015). Health Survey for England 2015: Health, social care and lifestyles.

Gerber, A. S., & Green, D. P. (2012). *Field Experiments: Design, Analysis and Interpretation*. Norton.

Gine, X., Karlan, D., & Zinman, J. (2010). Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation. *American Economic Journal. Applied Economics*, *2*(October), 213–235. doi:10.1257/app.2.4.213

Glennerster, R., & Takavarasha, K. (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton University Press.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*(1), 549–576. doi:10.1146/annurev.psych.58.110405.085530 PMID:18652544

Hales, C. M., Carroll, M. D., Fryar, C. D., & Ogden, C. L. (2017). *Prevalence of Obesity Among Adults and Youth* (pp. 2015–2016). United States.

Halpern, S. D., French, B., Small, D. S., Saulsgiver, K., Harhay, M. O., Audrain-McGovern, J., & Volpp, K. G. et al. (2015). Randomized Trial of Four Financial-Incentive Programs for Smoking Cessation. *The New England Journal of Medicine*, *372*(22), 2108–2117. doi:10.1056/NEJMoa1414293 PMID:25970009

Inauen, J., Tobias, R., & Mosler, H.-J. (2014). The role of commitment strength in enhancing safe water consumption. *British Journal of Health Psychology*, *19*(4), 701–719. doi:10.1111/bjhp.12068 PMID:24112306

Johnson, F., & Wardle, J. (2011). The association between weight loss and engagement with a web-based food and exercise diary in a commercial weight loss programme. *The International Journal of Behavioral Nutrition and Physical Activity*, *8*(1), 83. doi:10.1186/1479-5868-8-83 PMID:21810222

Jolly, K., Lewis, A., Beach, J., Denley, J., Adab, P., Deeks, J. J., & Aveyard, P. et al. (2011). Comparison of range of commercial or primary care led weight reduction programmes with minimal intervention control for weight loss in obesity. *British Medical Journal*, *343*, 1–16. doi:10.1136/bmj.d6500 PMID:22053315

Kelly, M. P. (2016). Digital Technologies and Disease Prevention. *American Journal of Preventive Medicine*, *51*(5), 861–863. doi:10.1016/j.amepre.2016.06.012 PMID:27745686

Kiesler, C. A., & Sakumura, J. (1966). A Test of a Model for Commitment. *Journal of Personality and Social Psychology*, *3*(3), 349–353. doi:10.1037/h0022943 PMID:5906339

Lee, D. S. (2002). Trimming for Bounds on Treatment Effects with Missing Outcomes. *NBER Technical Working Paper Series*, (June), 1–18.

Liu, P. J., Wisdom, J., Roberto, C., Liu, L. J., & Ubel, P. (2014). Using behavioral economics to design more effective food policies to address obesity. *Applied Economic Perspectives and Policy*, *36*(1), 6–24. doi:10.1093/aepp/ppt027

Loewenstein, G., Asch, D. A., Friedman, J. Y., Melichar, L. A., & Volpp, K. G. (2012). Can behavioural economics make us healthier? *BMJ (Clinical Research Ed.)*, *344*(May), e3482. doi:10.1136/bmj.e3482 PMID:22623635

Miller, D. T., & Prentice, D. A. (2013). Psychological Levers of Behaviour Change. In E. Shafir (Ed.), *The Behavioral Foundations of Public Policy* (pp. 301–309). Princeton University Press.

Mohr, D. C., Cuijpers, P., & Lehman, K. (2011). Supportive Accountability: A Model for Providing Human Support to Enhance Adherence to eHealth Interventions. *Journal of Medical Internet Research*, *13*(1), e30. doi:10.2196/jmir.1602 PMID:21393123

Murray, E., Edin, F., Hekler, E. B., Andersson, G., Collins, L. M., Doherty, A., & Wyatt, J. C. et al. (2016). Evaluating Digital Health Interventions. *American Journal of Preventive Medicine*, *51*(5), 843–851. doi:10.1016/j.amepre.2016.06.008 PMID:27745684

Nyer, P. U., & Dellande, S. (2010). Public Commitment as a Motivator for Weight Loss. *Psychology and Marketing*, *27*(1), 1–12. doi:10.1002/mar.20316

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

Parrott, R. L., Monahan, J., Ainsworth, S., & Steiner, C. (1998). Communicating to Farmers About Skin Cancer. The Behavior Adaptation Model. *Human Communication Research*, *24*(3), 386–409. doi:10.1111/j.1468-2958.1998.tb00422.x

Prestwich, A., Conner, M. T., Lawton, R. J., Ward, J. K., Ayres, K., & McEachan, R. R. C. (2012). Randomized controlled trial of collaborative implementation intentions targeting working adults' physical activity. *Health Psychology*, *31*(4), 486–495. doi:10.1037/a0027672 PMID:22468716

Rogers, T., Milkman, K. L., John, L. K., & Norton, M. I. (2015). Beyond good intentions: Prompting people to make plans improves follow-through on important tasks. *Behavioral Science and Policy*, *1*(2), 33–41. doi:10.1353/bsp.2015.0011

Rogers, T., Milkman, K. L., & Volpp, K. G. (2014). Commitment devices: Using initiatives to change behavior. *Journal of the American Medical Association*, *311*(20), 2065–2066. doi:10.1001/jama.2014.3485 PMID:24777472

Royer, H., Stehr, M., & Sydnor, J. (2015). Incentives, Commitments and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company. *American Economic Journal. Applied Economics*, *7*(3), 51–84. doi:10.1257/app.20130327

Ruhm, C. J. (2012). Understanding overeating and obesity. *Journal of Health Economics*, *31*(6), 781–796. doi:10.1016/j.jhealeco.2012.07.004 PMID:22935355

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, *23*(3), 165. doi:10.2307/2295722

Tauchmann, H. (2014). The Stata Journal. *The Stata Journal*, *14*(4), 884–894.

Thaler, R. H., & Shefrin, H. M. (1981). An Economic Theory of Self-Control. *Journal of Political Economy*, *89*(2), 392–406. doi:10.1086/260971

Torgerson, D. J., & Torgerson, C. J. (2008). *Designing Randomised Trials*. Basingstoke: Palgrave Macmillan.

Verhoeven, A. A. C., Adriaanse, M. A., Ridder, D. T. D. D. E., Vet, E. D. E., & Fennis, B. O. B. M. (2013). Less is more : The effect of multiple implementation intentions targeting unhealthy snacking habits. *European Journal of Social Psychology*, *43*(August), 344–354. doi:10.1002/ejsp.1963

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial Incentive–Based Approaches for Weight Loss. *Journal of the American Medical Association*, *300*(22), 2631–2637. doi:10.1001/jama.2008.804 PMID:19066383

Williams, B., McVey, D., Davies, L., & MacGregor, E. (2011). The Healthy Foundations Lifestage Segmentation.

## ENDNOTES

1 These examples demonstrate that commitment devices represent a range of different mental strategies, not physical or digital devices as such. The paper uses the terms commitment device as defined by Bryan et al (2010), interchangeably with the concept of pre-commitment strategy as originally coined by Strotz (1955).

2 see Ruhm 2012; Fudenberg and Levine 2006; and Gul and Pesendorfer 2001 for a discussion on related dual-self and dual-system theories.

3 https://www.nhs.uk/Livewell/Goodfood/Pages/eat-less.aspx

4 Ruhm (2012) demonstrates through an alternative dual-system formalisation that food consumption would exceed the optimal level due to the potential for self-control problems in such a decision-making context, compared with a simpler case where the planner sub-self faces no resistance from the doer's actions. Empirical evidence supports this prediction, with data from the US nutrition and health survey suggesting that overweight and obese people are more likely to prefer to weigh less than they actually do (Ruhm, 2012, p. 789).

5 Bryan et al (2010) use the dichotomy of 'hard' versus 'soft' commitment devices to draw a distinction between those that stake money and those that rely on psychological costs to generate behaviour change.

6 The study was approved by the UCL ethics committee (project ID 4518/002) and registered with the American Economic Association's Social Sciences Registry (http://www.socialscienceregistry.org/trials/942).

7 In line with a non-disclosure agreement with the company, their brand name is not mentioned here.

8 At the time of recruitment, participants were asked to confirm they had at least four weeks' membership remaining to the service. Invitations to the study were through online adverts visible only to clients once they had logged in to the system. Click through led to a request for informed consent and then a baseline survey.

9 Following pre-specified sample size calculations. On 21 November 2013 the desired sample size of 364 was reached and the survey was closed.

10 Monitoring of the online discussion forums revealed no evidence of cross-talk amongst participants.

11 Randomisation took place through the online survey, built on the Qualtrics platform.

12 Refunds varied because clients were on different monthly subscriptions based on the going price was when they signed up, and if they moved from an introductory offer to the 'normal' fee.

13 The partner firm imposed a limit of 100 refunds. 80% of all participants were recruited in phase 1 before the quota was reached. No significant differences were observed between participants recruited earlier and later in the trial.

14 Characteristics across male and female participants were broadly similar, excepting age: female participants cover the full age range from 18 to 65, male participants were concentrated in the 40s and 50s age range (none were younger than 30).

15 Later observations were recorded where two equally distant time stamps were found.

16 To take account of the change in randomisation allocation rule after the refund quota was reached, $\beta^{LC}$ is estimated across the sample recruited in phase 1 while $\beta^{RC}$ is estimated across the full sample since it was available as a treatment throughout.

[17] This excludes 9 participants in total who sought weight maintenance; 4 from the comparison group, 3 from the refund group and 2 from the coach group. Model specification including these participants does not alter the results.

[18] The Lee Bounds approach is premised on the data having two characteristics: firstly that treatments are as good as randomly assigned, as confirmed by Table 1 above; and secondly, that those who are observed (or missing) would be observed (or missing) regardless of treatment status. The latter assumption is justified through the attrition investigations

[19] Other strategies were considered but rejected: statistical modelling to fill in missing data, selection models (Heckman, 1976) or multiple imputation (Graham, 2009). These approaches required much stronger assumptions to be made on the nature of the missing data mechanism; assumptions which may be "tenuous and untestable" (Little and Rubin, 2000, p. 137).

[20] With inverse proportionality weighting, the ATE remains negative ($p = 0.028$) Including weight maintainers unsurprisingly lowered the weight loss differential ($p = 0.05$).

[21] A brief email survey was sent to 30 coaches and 16 were successfully reached, giving an overall response rate of 33% and a 53% response rate amongst those contacted.

[22] 5 of the 11 coaches who were 'actively' or 'very actively' involved reported not knowing or being unsure of whether the individual's weight loss target had been met at 4 weeks.

[23] Synonymous with the local average treatment effect (LATE) (Angrist & Pischke, 2009; Gerber & Green, 2012).

[24] The finding also has potential implications for our understanding of financial incentives for health, suggesting that such incentives, particularly if they are in the form of refunds, may be less effective where an individual is already accustomed to the idea of paying for something.

[25] Measured by the number of times the calorie counter tool was used over 12 weeks.

*Manu Savani teaches Public Policy at University College London. Her doctoral research evaluated the impact of commitment devices on health behaviours linked to obesity. Her research interests include the effects of present bias on decisions, dual-self theories, and the effectiveness of policy interventions that seek to influence choices that require a trade-off between benefits now and later.*