Journal of Health Economics 62 (2018) 95–104

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

The limits of simple implementation intentions: Evidence from a field experiment on making plans to exercise

Mariana Carrera[a,*], Heather Royer[b], Mark Stehr[c], Justin Sydnor[d], Dmitry Taubinsky[e]

[a] Case Western Reserve University, United States
[b] University of California at Santa Barbara & IZA & NBER, United States
[c] Drexel University, United States
[d] University of Wisconsin, Madison, NBER, United States
[e] University of California at Berkeley & NBER, United States

ARTICLE INFO

Article history:
Received 9 March 2018
Received in revised form 20 August 2018
Accepted 5 September 2018
Available online 20 September 2018

JEL classification:
C93
I12
D91

Keywords:
Implementation intentions
Planning
Physical activity
Exercise
Health behaviors
Behavioral economics
Nudge

ABSTRACT

Recent large-scale randomized experiments find that helping people form implementation intentions by asking when and where they plan to act increases one-time actions, such as vaccinations, preventative screenings and voting. We investigate the effect of a simple scalable planning intervention on a repeated behavior using a randomized design involving 877 subjects at a private gym. Subjects were randomized into i) a treatment group who selected the days and times they intended to attend the gym over the next two weeks or ii) a control group who instead recorded their days of exercise in the prior two weeks. In contrast to recent studies, we find that the planning intervention did not have a positive effect on behavior. We observe a tightly estimated null effect even though the majority of subjects believed that planning is helpful and despite clear evidence that they engaged with the planning process.

We are grateful for funding from the NIH and the outstanding research assistant work of Jacob Adamcik and Chang Lee. We thank conference participants at Advances with Field Experiments for helpful comments and suggestions.

Anyone who has made a routine visit to a dentist, pediatrician, or personal trainer is accustomed to the "nudge" to set up the next appointment before leaving. "Nudges" are low-cost interventions or manipulations of a choice environment aimed at influencing behavior in a non-coercive way (Thaler and Sunstein, 2008). Since behavioral obstacles like inattention, forgetfulness, and present bias often hamper engagement in optimal health behaviors, there is great interest in the use of nudges to promote public health and help people achieve their desired actions. For example, reminders sent by postcard, email, or SMS have been shown to increase dental check-ups (Altmann and Traxler, 2014), gym attendance (Calzolari and Nardotto, 2016), adherence to antiretroviral medication (Lester et al., 2010; Pop-Eleches et al., 2011), and child vaccinations (Busso et al., 2015). Employees are significantly more likely to get flu vaccinations if the clinic is located along their typical walking path (Beshears et al., 2016). Different types of nudges have targeted healthy eating through menu placement (Downs et al., 2009) and small incentives for healthy options (List and Samek, 2015; Loewenstein et al., 2016).[1]

A prompt to create a concrete plan for action (i.e., a planning prompt or implementation intention) is one particular pervasive nudge that is well-grounded in the psychology literature as a way to reduce the gap between intentions and actions (Gollwitzer and

[1] Of course, not all nudge interventions are effective and it is important to conduct rigorous tests of their effects. Two recent examples of field experiments testing nudge-tactics in health that found little effect of the nudges are Bronchetti, Huffman, and Magenheim on flu vaccination on college campus and Goldzahl et al. (2018) on breast cancer screening rates in France. Bronchetti, Huffman, and Magenheim (2015) find no effect for two low-cost nudges to increase flu vaccination on college campuses, and Goldzahl et al. (2018) find no effect of four behavioral interventions to increase breast cancer screening rates in France.

* Corresponding author.
E-mail address: mariana.carrera@montana.edu (M. Carrera).

https://doi.org/10.1016/j.jhealeco.2018.09.002
0167-6296/

Sheeran, 2006).[2] A planning prompt creates a mental association between engaging in a desired behavior and a specific future moment that, in turn, can help people attain their goals (Gollwitzer et al., 1996; Gollwitzer, 1999). Simple, low-cost planning prompts have shown promise in large-scale field experiments targeting flu vaccination (Milkman et al., 2011) and preventive screenings (Milkman et al., 2013). Thus, the recent Behavioral Interventions to Advance Self-Sufficiency (BIAS) project conducted by the Office of Planning, Research and Evaluation included prompting people for implementation intentions as one of seven highlighted and tested behavioral techniques for nudging change in human services (Richburg-Hayes et al., 2017).

While there is now strong evidence of planning interventions improving behavior for one-time actions, there is little comparable evidence that simple and scalable planning interventions are effective for important repeated actions. It may be more challenging to use implementation intentions to change repeated behaviors than to change one-time behaviors for a variety of reasons. First, people may have pre-established routines for repeated activities, and any concrete plans they make for the future may be mere reflections of those routines, rather than instruments for behavior change. Second, even if a concrete plan to act at a specific time does increase the likelihood of acting at that time, it might do so by reducing the likelihood of acting at an unplanned time. Third, the implementation intention cue may simply be less evocative for repeatable activities because there are many opportunities to perform the behavior in question. Consistent with this last concern, in Milkman et al.'s (2011) study, the effects of a planning prompt for influenza vaccination were present in locations with only one vaccination day but not in locations with multiple opportunities for vaccination.

We address the question of whether simple planning interventions are effective in repeated-action settings for the case of physical activity. Exercise is a natural setting where behavioral interventions may be attractive because the gap between intentions and behavior is often quite large. For example, Royer et al., (2015) document that workers' targeted levels of exercise are 43 percent higher than their actual levels of exercise.[3] We measure the effect of a simple planning intervention on gym visits using a large-scale randomized experiment with 877 members of a fitness facility. The treatment group was asked to make plans for which specific days and times over the following two weeks they would attend the gym. The control group was instead asked to recall the days and times in the prior two weeks when they had used the gym.

Our experimental design overcomes some key limitations found in prior research on planning interventions for repeated behaviors like exercise (Hagger and Luszczynska, 2014). First, our randomized and large-scale design overcomes the design issues inherent in much of the prior literature. Nearly all of the existing research utilizes small-scale studies that are either not randomized or involve tightly controlled but artificial environments. Second, our outcome measure comes from administrative gym check-in records. Administrative data avoid the potential biases associated with self-reports and also prevent problems of attrition that arise in some studies where people are asked to self-report at multiple times. We also are able to evaluate participant engagement with the planning intervention and assess how their plans deviate from their actual behavior - a limitation of the prior literature (Carraro and Gaudreau, 2013; Hagger and Luszczynska, 2014). Third, this study tests the effect of a simple and potentially scalable planning intervention – asking people to simply plan times for using the gym – in a relevant environment.

In contrast to the positive effects for health behaviors like flu shots and preventative screenings, we find that being prompted to make plans did not cause people to attend the gym more often. The planning-treatment group and the control group attended the gym equally often on average over the two-week study period. The treatment group made an average of 2.3 visits over the two-week period, compared to an average of 2.6 for the control group. This is a precise null effect as the 95% confidence interval on our estimated treatment effect excludes an increase in the average number of visits for the treatment group of more than 2%.

Our results lend more clarity to the literature on the effect of planning interventions on physical activity, which has previously found mixed results (Milne et al., 2002; Gollwitzer and Sheeran, 2006; Skar et al., 2011; Handel and Kolstad, 2017). Milne et al. (2002) present some of the strongest evidence of positive effects of planning on physical activity, but notably in their experiment the planning intervention focused subjects on planning to act once. More recent large-scale studies that focused on planning for the possibility of repeated physical activity found no effect of the intervention (Skar et al., 2011; Handel and Kolstad, 2017). The null effects in those studies, though, could be explained by relatively few people engaging with the intervention or, in the case of Handel and Kolstad (2017), by the bundling of the planning prompts with other interventions that could negate their effects. We build on this prior work by providing clear and robust evidence that subjects in our experiment engaged in the planning intervention. For example, we document that subjects actively made plans for significantly more visits than they had made in past weeks and more than they actually attended. Further, we show that the plans made by the subjects in the planning treatment are in fact highly predictive of the days when they actually attended the gym. This rules out the possibility that people haphazardly selected plans with little attention to when they might actually use the gym. Instead, it suggests that they planned to go on days where they were more likely to go, but that the act of planning did not actually increase the number of days they visited. As such, the results here significantly strengthen the evidence that a simple planning intervention for potentially repeated physical activity may have little effect.

This paper also contributes to a broader literature in psychology and economics that tests strategies for motivating healthy behaviors in the domains of exercise, weight loss, and smoking cessation. Controlled experiments have established that modest financial incentives can be effective at changing behavior while incentives are in place. See for example, Charness and Gneezy (2009); Acland and Levy (2015); Royer et al. (2015), and Carrera et al. (2017) for exercise incentives; Volpp et al. (2008) and John et al. (2011) for weight loss; and Halpern et al. (2015) and Volpp et al. (2009) for smoking cessation. However, this literature has also highlighted limitations of financial-incentive approaches. For example, effects are often short-lived and diminish over time (e.g., Cawley and Price, 2013; Carrera et al., 2017). Furthermore, these programs are often costly and not tailored to individual goals or needs. For example, in the case of exercise, it can be difficult to target only new exercise that would not have happened without the incentive. Ultimately this means that much of the incentive budget goes towards those already engaged in the activity (e.g., Royer et al., 2015). Some people are also averse to the very idea of paying people to engage in healthy behaviors (e.g., Carroll, 2015). Interventions like planning prompts have potential to address some of these issues because they are low-cost, easily scalable, and can be personalized. Our study reveals, however, that while this approach may be useful in some settings, the effect of simple planning prompts is certainly context dependent, and they are unlikely to be a broadly effective strategy for increasing exercise.

[2] For a review, see Rogers et al. (2016).

[3] Exercise targets may differ from exercise intentions, but finding physical activity studies measuring intentions and actual behavior are difficult to find.

Method

The primary research question for this experiment was pre-registered through the AEA RCT registry (ID AEARCTR-0001214).[4] We attest here that we report results for all treatment arms in and measures collected in the experiment. This study was approved by the institutional review boards of Case Western Reserve University and of the University of California-Santa Barbara.

We recruited subjects who were members of a private gym in a large Midwestern city and had participated in a prior survey on exercise behavior. The gym is affiliated with a nearby private university but is open to the public and is separate from the university's main student fitness facility.

The recruitment pool for this study consisted of 1210 members of the gym who had previously consented to be contacted about research studies. These members were pre-randomized in even proportions to receive a "planning" (treatment) or "non-planning" (control) version of an online survey directly embedded within an email. Subjects were informed in the email that they would receive a check for $20 for completing the survey.

Participants in the planning group were asked to *"check off the time of your workout [..] or select "no workout" if you don't plan to work out that day"* for each day in the 13 day period starting the following Tuesday (the Monday was the Memorial Day holiday and the gym was closed), May 31st through June 12th, 2016. The form showed a matrix of bubbles in which each day corresponded to a separate row, and columns represented every hour in which the gym would be open each day (6am to 11 pm on weekdays and 8am to 9 pm on weekends) as well as a choice for "No Workout" (See Appendix Fig. 1A). Subjects were told that the information would be used to create calendar invitations for each day/time that they planned to visit the gym, which they could click on in the follow-up email to add to their online calendar (iCalendar, Outlook and Google calendar were supported). They could also select a box to opt out of receiving this follow-up email with calendar invites. We observe whether individuals opted out of receiving the follow-up email, but do not observe whether the emailed calendar appointment invites were accepted by participants.

Participants in the non-planning (i.e., control) group saw a similar matrix of bubbles, but for dates in the preceding two weeks. They were asked to *"check off the time, as best you can remember, that you worked out [. . .] Select "no workout" if you didn't work out that day."* This ensured that the control group engaged in a similar activity focused on personal use of the gym, but without an explicit prompt to plan future visits. For both online form surveys, subjects were required to fill in a bubble for each day (indicating either a specific workout time or no workout) in order to successfully submit the form. Both versions of the survey also contained the same questions about personal experience with scheduling gym visits, use of calendars, and the number of days the participant expected to be out of town in the next two weeks.

The survey remained open for one week. Two weeks after the survey closing date, we obtained visit data from the fitness center's computerized log-in system, from five weeks prior to the survey week to the weeks in which the treatment group members recorded their planned visit days. We also obtained demographic data from the gym's member database.

Data availability

All of the data and code reported in this manuscript are available at the following link: https://figshare.com/s/4ce64b8431efa53fe4fa. The data and coding are in STATA 13 statistical software format. Those interested in replicating the analysis should use begin with the "READ ME" text file included in the file set.

Descriptive statistics

A total of 877 members completed the survey, yielding a response rate of 72%. Because the treatment and control surveys were visible in the recruitment emails there is a potential concern about selection into participation. In designing the experiment, we felt that the simplicity of the embedded survey would increase participation in the study and did not anticipate selection problems. The observable characteristics of participants suggest that differential selection across treatment arms was indeed not a problem. The response rates were nearly identical in the treatment group (438/605) and the control group (439/605). Table 1a displays summary statistics among participants for several pre-treatment characteristics. Consistent with random assignment and a lack of differential selection into participation, the means are similar across the treatment and control groups. In both groups, approximately 60% of participants are female and 56–60% are students. The average number of visits in the two-week pre-intervention period is close to 3, i.e., an average of 1.5 visits per week, and the average participant expects to be out of town approximately 4 days during the treatment period. Table 1b shows the summary statistics we have available for the full recruitment sample, including those who participated and those who did not.[5] Consistent with randomization, the means for control and treatment groups are similar. The only significant difference between participants and non-participants is that the non-participants have lower average pre-survey gym attendance, but this is similar between treatment and control assignment.

A priori, there is potential for a planning intervention to be effective, as Table 1b shows that over 60 percent of the participants reported some agreement with the statement *"I don't go to the gym as much as I would like because I don't set aside time for it in my schedule; then my schedule fills up and I no longer have time to go to the gym."* and a similar fraction reported believing that planning might help them attend the gym more often.[6]

In order for planning to be effective, subjects must take it seriously. We can gauge this by measuring the extent to which plans were associated with behavior. If subjects were filling out our form as quickly as possible to earn the $20 incentive (i.e., at random), and little actual planning were taking place, then their plans should not predict behavior. If, in contrast, subjects were taking time to think about when they might like to go to the gym, then plans should predict behavior.

To this end, in Table 2a, we simply regress the number of visits on the number of planned visits, with and without controlling for visit frequency in the pre-intervention period. Specifically, the regression model in Column 1 relates the number of planned visits over the 2-week intervention period to the number of actual visits during that same period. Overall, there is a significant asso-

[4] https://www.socialscienceregistry.org/trials/1214

[5] We have information from the gym records on basic demographics and prior gym use for this sample, but not answers to the survey questions embedded in the study that are shown in Table 1a. The sample size for Table 1b is 1,186. Twenty-four members who were originally in the recruitment sample were no longer present in the fitness center's records when we analyzed the data for this study. This attrition is quite similar across treatment and control assignment and likely reflects membership cancellations.

[6] Of the 60% of subjects who "Somewhat" or "Very much" agreed with the first statement, 75% answered "Yes" or "Maybe" to the question "Do you think you would go to the gym more often if you planned one or two weeks ahead about what days/times you would go to the gym?"

Table 1a
Summary Statistics for Participants.

	No Planning (Control) Mean	Planning (Treatment) Mean	Difference in Means	P-value Testing Difference = 0
Male	0.39	0.40	0.01	0.66
Age[a]	34.22	35.13	0.92	0.37
University Affiliated	0.66	0.61	−0.05	0.15
Student	0.60	0.56	−0.04	0.26
Secondary on Account	0.06	0.07	0.02	0.33
Gym Visits in Two Week Pre-Intervention Period	2.98 [3.14]	2.83 [3.01]	−0.15	0.47
Expected Days Out of Town in 2 Week Treatment Period	3.83 [4.31]	4.03 [4.38]	0.20	0.50
I Don't Plan Out Gym Attendance				
Doesn't Apply to Me	0.40	0.38	−0.02	0.51
Applies Somewhat	0.41	0.41	0.00	0.91
Applies to Me	0.19	0.21	0.03	0.35
Planning Would Help Me				
No	0.22	0.23	0.01	0.73
Maybe	0.30	0.30	0.00	0.98
Yes	0.23	0.24	0.01	0.86
I Already Plan	0.25	0.24	−0.02	0.60
Number of Observations	439	438		

"Secondary on account" designates people who were added as partners of existing members at a discounted rate. The p-value is for a test of equality of means between treatment and control group. For the non-dichotomous variables, the numbers in brackets represent the standard deviations. Subjects were sent an email with an embedded survey. Sample includes only those subjects who filled out and submitted the survey.

[a] Age is missing for one member of the treatment group and one member of the control group. *I Don't Plan Out Gym Attendance* was generated by asking subjects, ̈To what extent does the following statement apply to you? I don't go to the gym as much as I would like because I don't set aside time for it in my schedule; then my schedule fills up and I no longer have time to go to the gym.¨ *Planning Would Help Me* was generated by asking subjects, D̈o you think you would go to the gym more often if you planned one or two weeks ahead about what days/times you would go to the gym?¨

Table 1b
Summary Statistics for All Subjects.

	No Planning (Control) Mean	Planning (Treatment) Mean	Difference in Means	P-value Testing Difference = 0
Male	0.39	0.4	0.01	0.78
Age‡	34.12	35.21	1.09	0.21
University Affiliated	0.65	0.61	−0.04	0.15
Student	0.59	0.56	−0.03	0.32
Secondary on Account	0.06	0.08	0.02	0.21
Gym Visits in Two Week Pre-Intervention Period	2.75 [3.03]	2.62 [2.94]	−0.13	0.46
Number of Observations	594	592		

Subjects were sent an email with an embedded survey. Sample includes both those who did and did not submit the survey. See notes to Table 1a for variable definitions.

Table 2a
Association Between Planning and Behavior–Total Number of Visits.

	Dependent Variable: Number of Visits During 2-Week Intervention Period		
Independent Variables	(1)	(2)	(3)
Number of planned visits for 2 week intervention period	0.28***		0.19***
	(0.04)		(0.03)
Number of visits in 2 weeks before intervention		0.58***	0.53***
		(0.05)	(0.05)
Observations	438	438	438
Adjusted R-squared	0.17	0.38	0.43
Mean of dependent variable	2.33	2.33	2.33

*** p < 0.01, ** p < 0.05, * p < 0.1. Heteroskedasticity robust standard errors are in parentheses. Each column represents a separate regression. All regressions include indicator variables for age, age missing, gender, university affiliation, student, and membership type whose coefficients are not reported.

ciation between actual attendance and planned attendance. An extra planned visit during the intervention period translates into an increase of 0.28 of a visit. While this estimate is positive and statistically significant, it is well below 1, meaning that not all plans are fulfilled. Column 2 shows that recent pre-intervention gym attendance is in fact more predictive than planned visits for behavior during the 2-week intervention period as both the estimated coefficient and the R-squared increase considerably. This result raises the possibility that planned visits may be predictive of intervention period visits only because they are associated with past visits. Column 3 shows that this is not the case. When both variables are included, pre-intervention attendance is more predictive of intervention visits than planned visits, but the coefficient on an extra planned visit is still statistically significant and large.[7]

In Columns 1 and 2 of Table 2b, we examine the timing of visits more closely by studying whether plans to go at a particular day or time are predictive of whether subjects go on that specific day or time. The estimate in Column 1 implies that a plan to visit on a day is correlated with a 21 percentage point increase in the probability of attending the gym on that day. Given a mean attendance

[7] For these regressions and subsequent regressions, the results are robust to controlling for past attendance beyond 2 weeks prior to the intervention. Also, for the regressions in Table 2a, we can replace the number of visits in the 2 weeks before intervention variable with indicator variables for each possible number of visits (e.g., an indicator for 0 visits, an indicator for 1 visit, etc.) and the results are very similar.

Table 2b
Association Between Planning and Behavior–Time of Visits.

	Dependent Variable: Indicator for Visited on that Day (Columns (1) and (2)) or that Day and Hour (Columns (3) and (4))			
Independent Variables	(1)	(2)	(3)	(4)
Planned visit on that day indicator	0.21***	0.16***		
	(0.02)	(0.01)		
Fraction of days visiting gym on that same day during 2 week pre-intervention period		0.34***		
		(0.03)		
Planned visit on that day and hour indicator			0.09***	0.06***
			(0.01)	(0.01)
Fraction of days visiting gym on that same day and hour during 2 week pre-intervention period				0.27***
				(0.02)
Level of observation	Person x day	Person x day	Person x hour x day	Person x hour x day
Observations	5,694	5,694	101,178	101,178
Adjusted R-squared	0.13	0.20	0.02	0.07
Mean of dependent variable	0.18	0.18	0.01	0.01

*** p < 0.01, ** p < 0.05, * p < 0.1. Standard errors are clustered by individual. Each column represents a separate regression. Planned visit on that day indicator equals one if subject planned to go to the gym on that date and 0 otherwise. Planned visit on that day and hour indicator equals one if subject planned to go to the gym on that date at that hour and 0 otherwise. Fraction of days visiting gym on that same day in 2 week pre-intervention period takes on values of 0 (none), 0.5 (half), or 1 (all). Fraction of days visiting gym on that same day and hour in 2 week pre-intervention period takes on values of 0 (none), 0.5 (half), or 1 (all). All regressions include indicator variables for age, age missing, gender, university affiliation, student, and membership type. Person by day regressions include observations on 438 individuals over 13 days; person by hour regressions include 438 individuals over 231 h.

frequency of 0.18, this means that subjects are more than twice as likely to attend the gym on planned days than on unplanned days. In Column 2 we add controls for the fraction of days that the subject visited the gym on the same day during the two weeks prior to the intervention. The inclusion of this past attendance variable controls for the possibility that planning and attendance are correlated only because those who plan for more visits are more frequent attendees of the gym. As in Table 2a, while past attendance patterns are associated with gym visits over the 2 week intervention period, having a planned visit on a particular day is still predictive of actual attendance in Column 2. Columns 3 and 4 provide an even more granular analysis by studying the correlation between plans and behavior for a particular one-hour time slot. The association is again statistically significant and large, even when accounting for past attendance patterns.

Evaluation of the effectiveness of the implementation intentions

Regression model

We estimate treatment effects comparing participants who were randomly assigned to receive the planning treatment with those who were not, irrespective of whether they made plans to go to the gym. In addition, we also estimate "intent-to-treat" differences by comparing the visit rates for the full sample of recruited subjects, regardless of their participation in the survey, across treatment assignments. For our main results (Table 3), we estimate the following OLS regression:

$$Visits_i = \alpha + \beta PlanningTreatment_i + X_i'\theta + \varepsilon_i,$$

where $Visits_i$ is the total number of days a participant visited the gym during the 2-week (13 day) intervention period, *Planning Treatment* is a binary indicator for the planning group, and X_i' includes the control variables age, gender, university affiliation, student, membership type (regular, graduate student, or subsidized through a wellness program), and the number of days the individual visited the gym in the two-week pre-intervention period.

Regression estimates

The prior results show that the treatment group's plans were predictive of the actual visit patterns, strongly suggesting that the treatment group took the planning exercise seriously. In Table 3, we turn to our main results comparing the planning-treatment group to the control group to assess whether the planning intervention had a *causal* effect on the total number of gym visits during the 2-week intervention period. Columns 1 and 3 report a regression of the number of visits on an indicator variable designating whether or not the subject was assigned to the planning treatment without additional controls. Column 1 presents results for the participant sample and Column 3 shows results for the full recruitment sample. Among participants we estimate that the planning treatment resulted in 0.3 fewer visits on average than the non-planning treat-

Table 3
Effect of Planning on Total Gym Visits.

	Dependent Variable: Number of Visits During 2-Week Intervention Period			
Independent Variable	(1)	(2)	(3)	(4)
Planning treatment indicator	−0.30	−0.26	−0.16	−0.11
	(0.21)	(0.16)	(0.17)	(0.13)
Control variables	No	Yes	No	Yes
Number of observations	877	877	1,186	1,186
Adjusted R-squared	0.00	0.43	0.00	0.42
Mean of dependent variable for control group	2.62	2.62	2.36	2.36

*** p < 0.01, ** p < 0.05, * p < 0.1. Heteroskedasticity robust standard errors are in parentheses. Each column represents a separate regression. Columns 1 and 2 are for subjects who submitted the survey. Columns 3 and 4 include additional subjects who received an email with the survey embedded, but did not submit it. Control variables include the number of days subject visited the gym in the two week pre-intervention period and indicator variables for age, age missing, gender, university affiliation, student, and membership type.

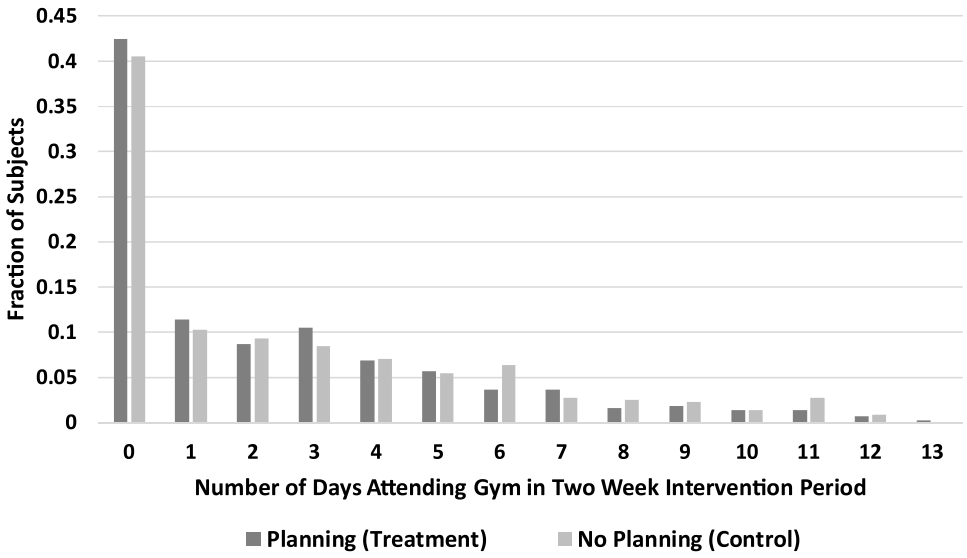

Fig. 1. Distribution of Days Visiting Gym by Treatment Status.

ment. This estimate is small and precisely estimated. The 95 percent confidence interval ranges from −0.71 of a visit to 0.11 of a visit. The upper bound of this confidence interval rules out positive impacts of planning exceeding 4.2 percent. The treatment-effect estimate is similar and smaller for the full recruitment sample. The addition of control variables in Column 2 and 4 does little to change the estimates and improves the statistical precision slightly, such that the upper bound on the 95-percent confidence interval for the participant estimate excludes a positive effect of planning greater than 2 percent.

The effects on the mean number of visits reported in Table 3 could conceivably mask other important distributional impacts, such as a reduction in the fraction making zero visits. To assess this possibility and to complement Table 3, we also display the distribution of visits among participants for the planning (treatment) group and the non-planning (control) group in Fig. 1.[8] We find that the histograms of visits are nearly identical for the treatment and control groups. The Wilcoxon rank sum (Whitney Mann) test p-value is 0.8903, indicating that we cannot reject the null hypothesis of equivalent distributions.

In Appendix Table A1 we report the results of several sensitivity analyses. We test whether the effect of the planning intervention differed based on past frequency of gym visits, survey responses indicating failure to plan gym visits, the belief that planning would help, the number of days during the planning period an individual expects to be out of town, and personal use of an online calendar system that would allow for automated reminders of planned visits. We find no evidence of positive planning effects for these subgroups.

Fig. 2 contrasts the visit patterns of the planning and no planning participant groups throughout the course of the 2-week intervention. Each of the bars denotes the fraction going to the gym on a particular date along with its 95% confidence interval. Consistent with the reported null result, the attendance rates for the treatment and control group are similar for most days, with the largest differences occurring during the weekends during which the attendance rates were somewhat higher for the control group. A Kruskal-Wallis test of the equality of the histograms gives a chi-squared (1) value of 1.22, and a p-value of 0.27, and thus we fail to reject equality. Overall, our analysis lends limited support for the usefulness of planning prompts to influence exercise behavior.

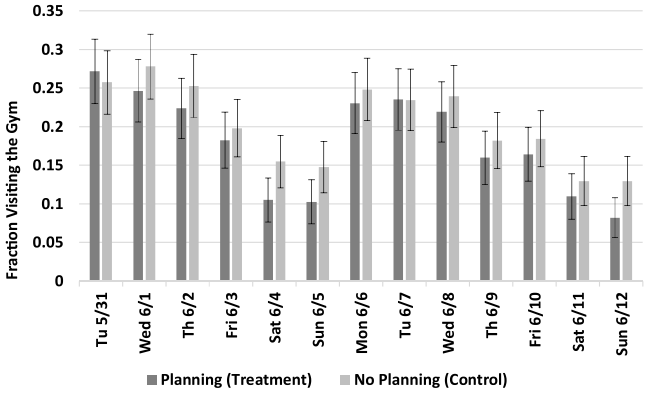

Fig. 2. Fraction Visiting Gym during Each Day of Planning Period.

Discussion

Recent large-scale field experiments have documented a significant effect of simple implementation intention prompts on one-time behaviors such as obtaining influenza vaccines, getting a colonoscopy, or voting. In this paper, we use a similar approach to test for the effect of a simple planning prompt on a key repeated health behavior: exercise. By asking members of a fitness center to plan the date and time of their visits to that center over the next two weeks, our intervention met two key criteria—short time horizons and well-specified intentions—that are thought to increase the impact of implementation intentions (Sheeran and Orbell, 1999). Yet, in contrast to the significant effects reported for one-time actions, we document a tightly estimated null effect, ruling out positive effect sizes greater than 2%.

Our results suggest that encouraging individuals to plan their gym visits over a 2-week period had little, if any, influence on their actual rates of gym attendance. As we demonstrated in Tables 2a and 2b, this is unlikely due to the plans not being meaningful, as plans are indeed predictive of the days, times, and total quantity of gym visits made by participants. Moreover, failure to make plans is not a possible explanation either, as 90% of treatment group subjects made a plan to go on at least one day. Forgetfulness

[8] Note the total number of possible visits during the intervention period is 13 days because one day during the two-week period (Memorial Day) the gym was closed.

about plans is also unlikely to be an explanation, since results were also not positive among the subset of participants using online calendar systems, which allowed them to receive email reminders of scheduled visits (see Appendix Table A1).

One possible factor explaining these findings is our lack of a "pure" control group. Specifically, the control group had to report their prior gym usage (done to keep the control and treatment surveys comparable lengths). It is possible that the act of reporting past attendance affected gym attendance and reduced the difference between the treatment and control groups' behaviors. However, we see no evidence of this type of effect when we compare visit patterns between control group participants and non-participants who were assigned to the control group. Specifically, we analyzed the number of gym visits made between the two weeks preceding and following our survey for both of these groups and see very small and statistically insignificant differences in the trajectory (i.e., difference in differences) of visit rates across the control group and non-respondents. We conclude, therefore that neither participating in the study nor assignment to the planning group had any impact on changes in gym attendance during the study period.

Our results contrast with the more positive results of three recent planning prompt field experiments on voting (Nickerson and Rogers, 2010), flu vaccination (Milkman et al., 2011), and preventive health screenings (Milkman et al., 2013). The different findings between our study and the other studies are not due to power. Each of the prior studies has large sample sizes: the smallest was Milkman et al. (2011) with 3272 subjects and the largest was Nickerson and Rogers (2010) with 287,228 subjects. It could be, however, that the effects of planning prompts are very context specific – a point we discuss in further detail below. The effects may even vary within a study, e.g., the results in Milkman et al. (2011) that planning the date and time changed vaccination rates but planning only the date did not.

One possible explanation for our different results is that, as noted by Milkman et al., "implementation intentions prompts may be most effective at encouraging behaviors when the opportunity for action is fleeting" (Milkman et al., 2011). Repeated behaviors like exercise, however, are very unlikely to produce a feeling of urgency, since many individuals likely have the mindset that they can always exercise "later." Thus, even if the planning prompts succeed at making certain times more salient as opportunities for gym attendance, individuals may nevertheless choose not to act on those opportunities, perhaps because of psychological barriers such as procrastination.[9]

Another possibility is that many people may already have pre-established routines for attending (or not attending) the gym. This would limit the efficacy of implementation intentions prompts if they have a smaller effect on activities guided by strong habits (Webb et al., 2009). And even if individuals do not have regular routines for attending the gym, they may still have difficulty breaking routines for activities that they engage in instead of attending the gym. Although previous work suggests that implementation intentions can "break the link" between past and future behavior (Orbell et al., 1997), our findings do not support this hypothesis. For participants in the planning group, past behavior remained a stronger predictor of future behavior than their implementation intentions. Moreover, the intervention had no effect even within the subsets of participants for whom we would expect to see the strongest effects: *(i)* those who reported that they did not currently plan out their gym attendance and *(ii)* those who thought that planning could help them visit more often (Appendix Table A1).

There is also an important distinction between implementation-intention prompts for one-time actions versus repeated actions. For one-time actions, implementation prompts simply ask individuals to form an intention about *when* and *where* they will take an action that they desire to complete. For some repeated behaviors, such as exercising at a gym, the added question of *how often* is implicitly raised. Consistent with other studies showing that people are overoptimistic about their future exercise patterns (Royer et al., 2015), we observe individuals making considerably overambitious plans to visit the gym, with their total planned visits being three times as large as their actual visits (7.04 vs. 2.33). If people are knowingly making ambitious plans (i.e., planning to visit more often than they think is likely), then the mental association created by each individual visit plan might be weaker than that of a one-time action plan. Potentially consistent with this idea, our subjects are slightly more likely to actually attend on the first day they plan to attend than on subsequent planned days.[10] That pattern may suggest that the mental association created by plans is weaker with more plans or plans made for further in the future. Alternatively, if people are unaware that their plans are overambitious, then the gradual realization that they are falling far short of their visit goals might be de-motivating (Heath et al., 1999). Note that the question of *how often* is less relevant for some repeated behaviors that are supposed to be carried out daily, such as daily vitamin intake, which has been studied in previous work (Sheeran and Orbell, 1999).

Although our study documents that a very simple planning exercise focused on planning *when* to go to the gym was not effective, this evidence does not rule out that more elaborate implementation-intention interventions could be effective. In particular, it may be that approaches that supplement simple *when* plans with plans involving *how* the action will be accomplished and *how* barriers will be overcome could be a way forward for effectively promoting repeated physical activity in practical settings (Hagger and Luszczynska, 2014). Our results imply that the promise of the simplest approach to implementation intentions suggested by the recent field experiments on one-time actions does not extend easily to all domains, but should not discourage future research on the use of implementation intentions for exercise and other health behaviors that require sustained engagement over time.

We hope that by documenting a stark contrast between the effects of simple planning prompts on repeated versus one-time actions, this paper will motivate future research examining the possible key differences between repeated and one-time behaviors that generate this contrast in practical settings. Such research can shed new light on the pathways by which implementation intentions influence human behavior, as well as the factors that mediate the gap between people's intentions and actions. We further hope that these results will support continued research into the use of behaviorally informed interventions to improve health behaviors. Our findings highlight that simple behavioral interventions, while often quite effective, do not work the same in all situations and need to be tested and refined for specific applications in health.

[9] Even if individuals were to visit the gym at the times made more salient by the planning prompts, overall gym attendance may not increase if attendance at those times comes at the expense of attendance at other times.

[10] Among individuals in the planning treatment who made plans, we observe that they actually go to the gym 35% of the time on the first planned day (first calendar day with a plan), while by the 3rd planned day the visit rate is only 27%. There are limits, however, to our ability to analyze the data in these ways since not all participants made plans for the same numbers of days and the number of planned days may be related to the strength of a participant's desire to go to the gym.

Appendix A.

Table A1
Heterogeneity in Effect of Planning on Total Gym Visits.

	Dependent Variable: Number of Visits During 2-Week Intervention Period					
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treated	−0.26	−0.06	−0.43	−0.22	−0.03	−0.67**
	(0.16)	(0.19)	(0.31)	(0.26)	(0.22)	(0.31)
x Above median visits in pre period		−0.41				
		(0.34)				
x Does not currently schedule gym visits			0.29			
			(0.36)			
x Thinks planning might help				−0.08		
				(0.33)		
x Will be out of town 4+ days					−0.49	
					(0.32)	
x Uses online calendar						0.59*
						(0.36)
Above median visits in pre period		−0.39				
		(0.39)				
Does not currently schedule gym visits			−0.69**			
			(0.29)			
Thinks planning might help				−0.33		
				(0.24)		
Will be out of town 4+ days					−0.72***	
					(0.24)	
Uses online calendar						−0.46*
						(0.28)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	877	877	877	877	877	877
Adjusted R-squared	0.43	0.43	0.43	0.43	0.45	0.43
Mean of dependent variable for control group	2.62	2.62	2.62	2.62	2.62	2.62

Heteroskedastic robust standard errors are in parentheses. See notes to Table 3 for a full list of control variables. *Above median visits in pre period* equals 1 if visits in the two week pre period were 3 or more (45% of the sample) and 0 otherwise. *Does not currently schedule gym visits* equals 1 if subject chooses Šomewhatör V̈ery muchẅhen asked the following question, T̈o what extent does the following statement apply to you? I don't go to the gym as much as I would like because I don't set aside time for it in my schedule; then my schedule fills up and I no longer have time to go to the gym.ẗ(61% of the sample) and 0 otherwise. *Thinks planning might* help equals 1 if subject chooses M̈aybeör Ÿeswhen asked, D̈o you think you would go to the gym more often if you planned one or two weeks ahead about what days/times you would go to the gym?ẗ(53% of the sample) and 0 otherwise. *Will be out of town 4+ days* equals 1 if subject chooses four days or more when asked, Ḧow many days do you expect to be out of town in the next two weeks?ẗ(42% of the sample) and 0 otherwise.

*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

6. **Please select your workouts for the week of May 30th:** *

Remember that the gym is closed on Monday, May 30th for Memorial Day.
Mark only one oval per row.

	No Workout	6 AM	7 AM	8 AM	9 AM	10 AM	11 AM	Noon	1 PM	2 PM	3 PM	4 PM	5 PM	6 PM	7 PM	8 PM	9 PM	10 PM	11 PM
Tuesday, May 31st	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Wednesday, June 1st	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Thursday, June 2nd	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Friday, June 3rd	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

7. *

Mark only one oval per row.

	No Workout	8 AM	9 AM	10 AM	11 AM	Noon	1 PM	2 PM	3 PM	4 PM	5 PM	6 PM	7 PM	8 PM	9 PM
Saturday, June 4th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Sunday, June 5th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

8. **Please select your workouts for the week of June 6th:** *

Mark only one oval per row.

	No Workout	6 AM	7 AM	8 AM	9 AM	10 AM	11 AM	Noon	1 PM	2 PM	3 PM	4 PM	5 PM	6 PM	7 PM	8 PM	9 PM	10 PM	11 PM
Monday, June 6th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Tuesday, June 7th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Wednesday, June 8th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Thursday, June 9th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Friday, June 10th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

9. *

Mark only one oval per row.

	No Workout	8 AM	9 AM	10 AM	11 AM	Noon	1 PM	2 PM	3 PM	4 PM	5 PM	6 PM	7 PM	8 PM	9 PM
Saturday, June 11th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Sunday, June 12th	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

Fig. A1. Planning Worksheet for Planning Treatment.

References

Acland, Dan, Levy, Matthew R., 2015. Naiveté, Projection Bias, and Habit Formation in Gym Attendance. Manage. Sci. 61 (1), 146–160.

Altmann, S., Traxler, C., 2014. Nudges at the dentist. Eur. Econ. Rev. 72, 19–38.

Beshears, J., Choi, J.J., Laibson, D.I., Madrian, B.C., Reynolds, G.I., 2016. Vaccination rates are associated with functional proximity but not base proximity of vaccination clinics. Med. Care 54 (6), 578.

Bronchetti, E.T., Huffman, D.B., Magenheim, E., 2015. Attention, intentions, and follow-through in preventive health behavior: field experimental evidence on flu vaccination. J. Econ. Behav. Org. 116, 270–291.

Busso, M., Cristia, J., Humpage, S., 2015. Did you get your shots? Experimental evidence on the role of reminders. J. Health Econ. 44, 226–237.

Calzolari, G., Nardotto, M., 2016. Effective reminders. Manage. Sci. 63 (9), 2915–2932.

Carraro, N., Gaudreau, P., 2013. Spontaneous and experimentally induced action planning and coping planning for physical activity: a meta-analysis. Psychol. Sport Exerc. 14 (2), 228–248.

Carrera, M., Royer, H., Stehr, M.F., Sydnor, J., 2017. The structure of health incentives: evidence from a field experiment. In: National Bureau of Economics Working Paper No. 23188.

Carroll, A., July 6, 2015 2015. Paying People to Be Healthy Usually Works, If the Public Can Stomach It. New York Times, The Upshot https://nyti.ms/1KKfTlW.

Cawley, John, Price, Joshua A., 2013. A case study of a workplace wellness program that offers financial incentives for weight loss. J. Health Econ. 32 (5), 794–803.

Charness, Gary, Gneezy, Uri, 2009. Incentives to exercise. Econometrica 77 (3), 909–931.

Downs, J.S., Loewenstein, G., Wisdom, J., 2009. Strategies for promoting healthier food choices. Am. Econ. Rev. Papers Proc. 99 (2), 159–164.

Gollwitzer, P.M., 1999. Implementation intensions: strong effects of simple plans. Am. Psychol. 54, 493–503.

Gollwitzer, P.M., Sheeran, P., 2006. Implementation intentions and goal achievement: a meta analysis of effects and processes. Adv. Exp. Soc. Psychol. 38, 69–119.
Gollwitzer, P.M., 1996. The volitional benefits of planning. In: Gollwitzer, P.M., Bargh, J.A. (Eds.), The Psychology of Action: Linking Cognition and Motivation to Behavior. Guilford, New York, pp. 287–312.
Goldzahl, L., Hollard, G., Jusot, F., 2018. Increasing breast-cancer screening uptake: a randomized controlled experiment. J. Health Econ. 58, 228–252.
Hagger, M.S., Luszczynska, A., 2014. Implementation intention and action planning interventions in health contexts: state of the research and proposals for the way forward. Appl. Psychol. Health Well. 6 (1), 1–47.
Halpern, Scott, French, Benjamin, Small, Dylan S., et al., 2015. Randomized trial of four financial-incentive programs for smoking cessation. N. Engl. J. Med. 372 (22), 2108–2117.
Handel, B., Kolstad, J., 2017. Wearable technologies and health behaviors: new data and new methods to understand population health. Am. Econ. Rev. Papers Proc. 107 (5), 481–485.
Heath, C., Larrick, R.P., Wu, G., 1999. Goals as reference points. Cogn. Psychol. 38 (1), 79–109.
John, Leslie K., Loewenstein, George, Troxel, Andrea B., Norton, Laurie, Fassbender, Jennifer E., Volpp, Kevin G., 2011. Financial incentives for extended weight loss: a randomized, controlled trial. J. Gen. Intern. Med. 26 (6), 621–626.
Lester, R.T., Ritvo, P., Mills, E.J., Kariri, A., Karanja, S., Chung, M.H., Marra, C.A., 2010. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 376 (9755), 1838–1845.
List, J.A., Samek, A.S., 2015. The behavioralist as nutritionist: leveraging behavioral economics to improve child food choice and consumption. J. Health Econ. 39, 135–146.
Loewenstein, G., Price, J., Volpp, K., 2016. Habit formation in children: evidence from incentives for healthy eating. J. Health Econ. 45, 47–54.
Milkman, K.L., Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C., 2011. Using implementation intentions prompts to enhance influenza vaccination rates. Proc. Natl. Acad. Sci. 108 (26), 10415–10420.
Milkman, K.L., Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C., 2013. Planning prompts as a means of increasing preventive screening rates. Preventative Med. 56 (1), 92–93.
Milne, S., Orbell, S., Sheeran, P., 2002. Combining motivational and volitional interventions to promote exercise participation: protection motivation theory and implementation intentions. Br. J. Health Psychol. 7 (2), 163–184.
Nickerson, D.W., Rogers, T., 2010. Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making. Psychol. Sci. 21 (2), 194–199.
Orbell, S., Hodgkins, S., Sheeran, P., 1997. Implementation intentions and the theory of planned behavior. Pers. Soc. Psychol. Bull. 23 (9), 945–954.
Pop-Eleches, C., Thirumurthy, H., Habyarimana, J.P., Zivin, J.G., Goldstein, M.P., De Walque, D., Ngare, D., 2011. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. AIDS (Lond., Engl.) 25 (6), 825.
Richburg-Hayes, L., Anzelone, C., Dechausay, N., Landers, P., 2017. Nudging Change in Human Services: Final Report of the Behavioral Interventions to Advance Self-Sufficiency (BIAS) Project, OPRE Report 2017-23.
Rogers, T., Milkman, K.L., John, L.K., Norton, M.I., 2016. Making the best laid plans better: how plan-making increases follow-through. Behav. Sci. Policy.
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. Am. Econ. J. Appl. Econ. 7 (3), 51–84.
Sheeran, P., Orbell, S., 1999. Implementation intentions and repeated behaviour: augmenting the predictive validity of the theory of planned behaviour. Eur. J. Soc. Psychol. 29 (23), 349–369.
Skar, S., Sniehotta, F.F., Gerard, M., Prestwich, A., Araujo-Soares, V., 2011. Do brief online planning interventions increase physical activity amongst university students? A randomized controlled trial. Psychol. Health 26, 399–417.
Thaler, R.H., Sunstein, C.R., 2008. Nudge: Improving Decisions About Health, Wealth and Happiness. Yale Univ Press, New Haven, CT.
Volpp, Kevin G., Leslie, K., John, Andrea B., Troxel, Laurie, Norton, Jennifer Fassbender, Loewenstein, George, 2008. Financial incentive–based approaches for weight loss: a randomized trial. J. Am. Med. Assoc. 300 (22), 2631–2637.
Volpp, Kevin G., Troxel, Andrea B., Pauly, Mark V., Glick, Henry A., Puig, Andrea, Asch, David A., Galvin, Robert, et al., 2009. A randomized, controlled trial of financial incentives for smoking cessation. N. Engl. J. Med. 360 (7), 699–709.
Webb, T.L., Sheeran, P., Luszczynska, A., 2009. Planning to break unwanted habits: habit strength moderates implementation intention effects on behaviour change. Br. J. Soc. Psychol. 48 (3), 507–523.