

Online Appendices

[A-1: Field Experiment Pre-Study]

We recruited 87 undergraduate students (at the same university) who were not part of the StudentHealth population through a variety of list-serves commonly used by undergraduates to share information on campus events and activities. In the solicitation, we informed participants that they had to be using some health tracking wearable to participate in the survey. We first asked participants to provide some demographic information about themselves (age, gender, race, etc.) and then asked them to provide some information on their use of wearable fitness devices. Specifically, we asked them to report what type of fitness device they used, how often they wore the device, and their average step count when using their device. We then asked participants to imagine that they had been asked to participate in a 7-day step challenge, in which they would set up a new step goal and try to meet it for those 7 days. Participants then indicated what they would consider to be a reasonable new goal based on their current average step count (10%, 20%, 30%, or a 40% increase); the actual step values were dynamically loaded for participants based on what they indicated was their average step count. In addition, participants were asked to indicate their preference for the gain-only vs. gain-loss incentive scheme, with a varying level of reward associated with the gain-loss incentive scheme. Specifically, we presented participants with nine scenarios, asking them to choose between a gain-only incentive scheme (labeled “Option 1”) that rewarded them \$1.00 for meeting their goal, and a gain-loss incentive scheme (labeled “Option 2”) which would reward them some amount for meeting their goal (the reward started at \$1.00 and was incremented by \$.25 all the way up to \$3.00), but had a \$1.00 penalty if they missed their goal (Table A-1-1).

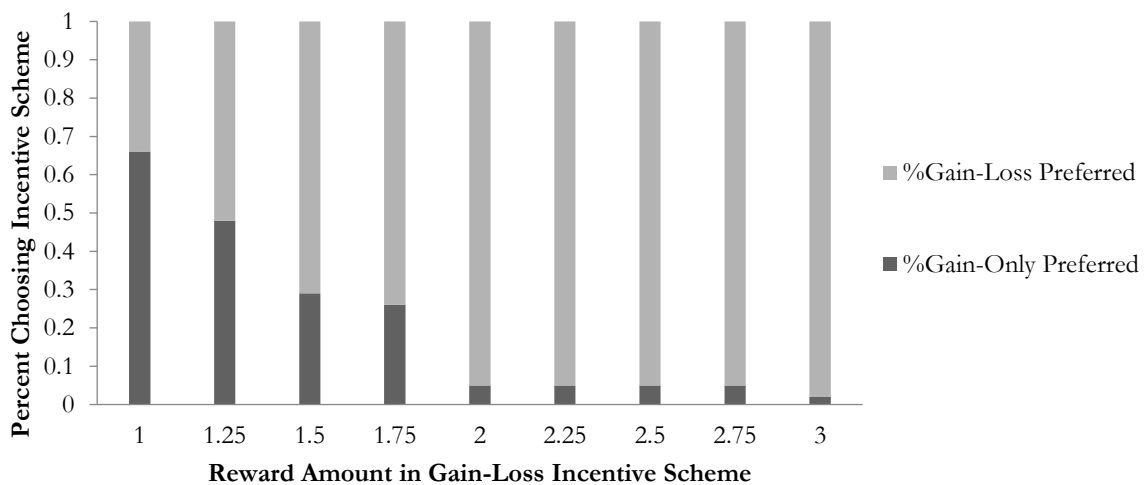
[Table A-1-1: Pre-Study Question]

	Gain-Only Incentive	Gain-Loss Incentives
1	Receive an additional \$1.00 every day you meet your step goal	Receive an additional \$1.00 every day you meet your step goal but have \$1.00 taken away every day you don't meet it
2	Receive an additional \$1.00 every day you meet your step goal	Receive an additional \$1.25 every day you meet your step goal but have \$1.00 taken away every day you don't meet it
...
8	Receive an additional \$1.00 every day you meet your step goal	Receive an additional \$2.75 every day you meet your step goal but have \$1.00 taken away every day you don't meet it

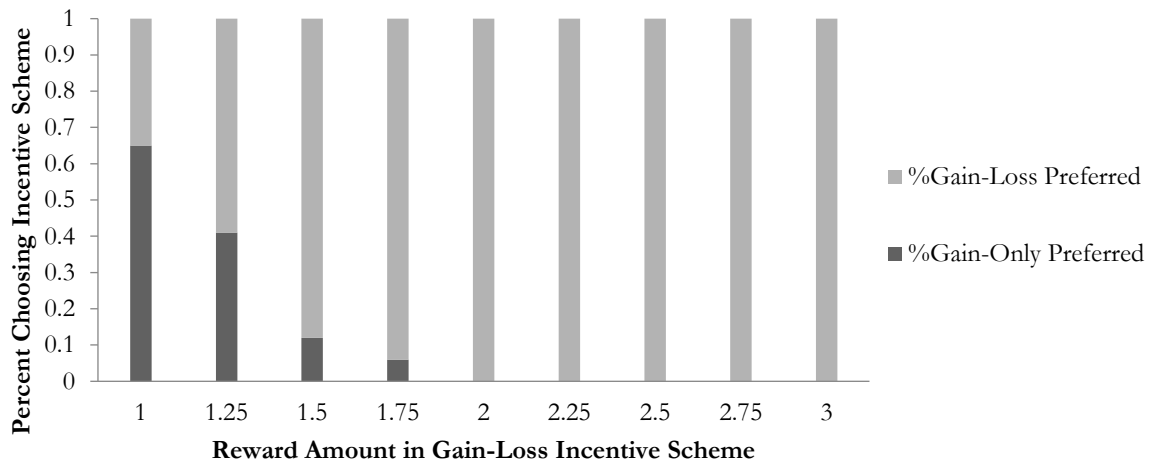
9	Receive an additional \$1.00 every day you meet your step goal	Receive an additional \$3.00 every day you meet your step goal but have \$1.00 taken away every day you don't meet it
---	----------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------

The majority of participants reported that a 20% increase in their current step average was a reasonable but challenging new goal (52% of participants); a 30% increase in their current step average was the next goal that participants indicated would be reasonable (23% of participants). We also found that participants exhibited a strong preference towards the gain-loss incentive scheme. Even when the reward of the gain-loss scheme was identical to that of the gain-only scheme (a case where the gain-only scheme strictly dominates the gain-loss one), 34% of participants indicated they would choose the gain-loss scheme (this is consistent with prior works finding non-zero preference for punishing commitment devices). At \$1.25 reward, a slight majority of participants (52%) indicated they would choose the gain-loss scheme, and at a reward of \$1.50, 71% of participants indicated a preference for the gain-loss scheme (Figure A-1.1). The results are similar when we conditioned on those who chose the 20% increase as a reasonable goal and reported having a Fitbit activity tracker (Figure A-1.2). These results suggest that participants recognize the self-control issues they face when pursuing healthier behavior and that they value (at least in hypothetical terms) gain-loss incentive schemes that help them achieve these goals. Using these results, we set the new goal for the field experiment as a 20% increase on participants' average step counts from the prior month and (if in the treatment conditions) offered participants a choice between a gain-only option that awarded \$1.00 for meeting a step goal (and no penalty), and a gain-loss scheme which awarded participants \$1.25 for meeting their goal but took away \$1.00 if they did not meet it.

[Figure A-1-1: Gain-Only vs. Gain-Loss Incentives]



[Figure A-1-2: Gain-Only vs. Gain-Loss Incentives - Fitbit @ 20%]



[A-2: Recruitment Text Messages]

Initial Recruitment:

Exciting News! All [University Name] students that are part of the StudentHealth study are invited to participate in a “Step Challenge” exclusive to [University Name].

It takes 2 minutes to sign up (no new apps or long surveys required) and you can do it on your phone. Get paid \$10 to participate! [link](#)

Second Recruitment:

Time is running out to sign up for the [University Name] Step Challenge! It takes 2 minutes to sign up (no new apps or long surveys required) and pays \$10! Join today and start the semester on the right foot! [link](#)

Third Recruitment:

Sign-ups for the [University Name] Step Challenge end this week and our records indicate you have yet to sign up. It takes 2 minutes to sign up (no new apps or long surveys required) and pays \$10! Do not hesitate, sign up now! [link](#)

[A-3: Study Text Messages]

Text: Day Before Challenge Starts

Control:

The [University Name] step challenge starts tomorrow Monday, [Date] and runs for 7 days. The last day of the challenge is Sunday, [Date]. Your new step goal during the challenge is **[New Step Goal]**. **[Take a minute now to change your step goal in the Fitbit App]**. You will be compensated \$10 dollars for participating in the challenge. Good luck!

Gain-Loss Incentives:

The [University Name] step challenge starts tomorrow Monday, [Date] and runs for 7 days. The last day of the challenge is Sunday, [Date]. Your new step goal during the challenge is **[New Step Goal]**. **[Take a minute now to change your step goal in the Fitbit App]**. You will be compensated \$10 dollars for participating in the challenge. Also, you **[chose / were randomly assigned]** to receive an additional \$1.25 every day you meet your new step goal but lose \$1.00 every day you don't meet your new step goal. Good luck!

Gain-Only Incentives:

The [University Name] step challenge starts tomorrow Monday, [Date]. and runs for 7 days. The last day of the challenge is Sunday, [Date].. Your new step goal during the challenge is **[New Step Goal]**. **[Take a minute now to change your step goal in the Fitbit App]**. You will be compensated \$10 dollars for participating in the challenge. Also, you **[chose / were randomly assigned]** to receive an additional \$1.00 every day you meet your step goal. Good luck!

Text Before Mid-update

You are halfway through the [University Name] step challenge. We plan to send you a status message summarizing your progress so far. In order for this message to be up to date, please sync your Fitbit before midnight.

Update Text:

Control:

Step Challenge update: Sept 5: You **(met, did not meet)** your goal for this day. Sept 6: You **(met, did not meet)** your goal for this day. Sept 7: You **(met, did not meet)** your goal for this day. Sept 8: Data not yet available. There are only three days left in the challenge, so make sure to finish strong!

Gain-Only Incentives:

Step Challenge update: [Date]: You **(met, did not meet)** your goal for this day, and **(earned, did not earn)** a \$1 bonus payment. [Date]: You **(met, did not meet)** your goal for this day, and **(earned, did not earn)** a \$1 bonus payment. [Date]: You **(met, did not meet)** your goal for this day, and **(earned, did not earn)** a \$1 bonus payment. [Date]: Data not yet available. There are only three days left in the challenge, so make sure to finish strong!

Gain-Loss Incentives:

Step Challenge update: [Date] You **(met, did not meet)** your goal for this day, **(earned, did not earn)** a \$1.25 bonus payment, and **(avoided, incurred)** a \$1 loss. [Date]: You **(met, did not meet)** your goal for

this day, **(earned, did not earn)** a \$1.25 bonus payment, and **(avoided, incurred)** a \$1 loss. [Date]: You **(met, did not meet)** your goal for this day, **(earned, did not earn)** a \$1.25 bonus payment, and **(avoided, incurred)** a \$1 loss. [Date]: Data not yet available. There are only three days left in the challenge, so make sure to finish strong!

Final Text:

Congratulations, you have completed the [University Name] Step Challenge! You will be sent a summary of your performance and also paid for your participation when all participant data is in. Great job and good luck with the rest of the semester!

[A-4: Pre-Trend Analysis and Sample Balance]

We evaluate the balance of our descriptive variables across the treatment conditions and find evidence that our randomization provides well balanced groups. We conduct 54 pairwise comparisons (9 variables * 6 comparisons) and find that 50 of these comparisons identify insignificant differences between the conditions. We identify 4 significant differences between these groups. With an alpha = .1, we would expect that 5 comparisons are significant by random chance. Nonetheless, we now control for these variables in our analysis throughout the manuscript and continue to identify consistent results.

[Table A-4-1: Comparison of Demographics between Incentive Conditions]

Variable	Control	Choice	G-L	G-O	1 vs. 2	1 vs. 3	1 vs. 4	2 vs. 3	2 vs. 4	3 vs. 4
	(1)	(2)	Assign	Assign	(5)	(6)	(7)	(8)	(9)	(10)
New Goal	12,760	12,756	13,422	12,817	0.99	0.43	0.95	0.47	0.95	0.63
Male	0.45	0.48	0.40	0.55	0.67	0.52	0.31	0.32	0.51	0.15
Body Mass Index	22.92	22.76	22.95	22.96	0.72	0.96	0.94	0.72	0.71	0.98
Age	18.68	18.60	18.59	18.49	0.14	0.15	0.01	0.89	0.14	0.18
Extraversion	3.23	3.17	3.16	3.19	0.62	0.60	0.78	0.93	0.90	0.84
Agreeableness	3.77	3.65	3.78	3.80	0.12	0.92	0.82	0.13	0.15	0.89
Conscientiousness	3.76	3.61	3.67	3.66	0.07	0.34	0.34	0.51	0.59	0.95
Neuroticism	2.82	2.93	2.92	2.97	0.25	0.40	0.25	0.89	0.81	0.73
Openness	3.30	3.29	3.35	3.51	0.93	0.61	0.04	0.55	0.02	0.13
N	110	100	55	44	210	165	154	155	144	99

We also conducted a similar set of validation and robustness checks for the individuals provided the prompt to update their Fitbit targets. We start with the set of individual characteristics and again find that demographic characteristics are well balanced between groups. We do identify one difference in that those in the “No Fitbit Target Prompt” condition are more likely to be male. Again, we expect that one of these comparisons will be significant by chance with an alpha = .1.

[Table A-4-2: Demographics by Fitbit Target Prompt]

Variable	No Fitbit Target Prompt	Fitbit Target Prompt	(1) vs. (2)
----------	-------------------------	----------------------	-------------

	(1)	(2)	(3)
New Goal	12418.43	12490.51	0.89
Male	0.52	0.40	0.03
Body Mass Index	23.18	22.62	0.11
Age	18.60	18.61	0.85
Extraversion	3.24	3.17	0.48
Agreeableness	3.74	3.73	0.86
Conscientiousness	3.70	3.65	0.44
Neuroticism	2.84	2.97	0.11
Openness	3.36	3.32	0.48

Despite finding balance between the experimental groups in the pre-treatment period, we sought to further address breaks in randomization by conditioning on individual characteristics and pre-treatment physical activity levels. We continue to find significant and nearly identical results to our main analysis on incentives and our analysis of Fitbit health tools. All incentive approaches have an initial impact on steps but those assigned to the gain-loss condition have the largest effects. More so, only those who are assigned to the gain-loss condition see the most persistent effects on steps walked and goal met in the intermediate and longer term (Table A-4-3).

[Table A-4-3: Main Effect of Incentives]

VARIABLES	(1) Challenge Week	(2) Short Term	(3) Intermediate Term	(4) Long Term	(5) Goal-Challenge Week
Choice	658.8* (395.7)	647.4* (378.2)	368.2 (320.7)	418.7 (300.7)	0.104*** (0.0380)
Gain-Only Assigned	1,071** (467.2)	423.2 (435.9)	-99.11 (493.0)	395.7 (448.4)	0.145*** (0.0409)
Gain-Loss Assigned	1,688*** (570.3)	1,293*** (464.8)	949.5** (419.4)	1,056*** (364.1)	0.183*** (0.0460)
Age	432.7 (468.1)	250.2 (402.7)	371.0 (395.8)	122.3 (328.6)	0.0622 (0.0400)
Male	-741.1* (395.9)	-948.3*** (339.4)	-983.6*** (307.2)	-1,205*** (292.0)	-0.0839** (0.0342)
Body Mass Index	50.33 (56.95)	8.517 (52.64)	38.54 (48.75)	31.56 (43.15)	0.00355 (0.00497)
Extraversion	-63.09 (243.6)	195.9 (224.9)	-9.807 (208.6)	177.1 (196.8)	-0.0129 (0.0204)
Agreeableness	1,111*** (341.3)	840.2*** (306.9)	555.7* (284.3)	351.8 (240.3)	0.0630** (0.0300)
Conscientiousness	228.2 (357.4)	218.3 (295.7)	93.95 (270.9)	357.2 (245.0)	0.0175 (0.0291)
Neuroticism	444.2 (292.9)	336.0 (249.2)	14.35 (238.5)	-38.98 (224.0)	0.00357 (0.0267)
Openness	-190.2 (294.2)	70.39 (267.3)	211.0 (249.1)	-184.4 (231.8)	-0.00786 (0.0252)

Pre-Treat Steps	0.480*** (0.0821)	0.443*** (0.0598)	0.416*** (0.0634)	0.468*** (0.0669)	-5.01e-05*** (6.78e-06)
Pre-Treat Compliance	14.36 (8.780)	21.29** (8.641)	18.57** (7.582)	2.441 (6.687)	0.00353*** (0.000772)
Pre-Treat Goal Met	-5,633*** (1,395)	-3,915*** (1,083)	-4,265*** (1,068)	-4,475*** (1,013)	0.787*** (0.119)
Constant	-5,920 (8,896)	-2,787 (7,680)	-3,120 (7,107)	3,305 (5,904)	-0.895 (0.743)
Individuals	285	281	276	262	285
Observations	1,962	7,634	13,152	19,133	1,962

Robust standard errors in parentheses; ** p<0.01, * p<0.05, * p<0.1; Time fixed effects in all columns

We also evaluate the impact of the Fitbit target prompt with the addition of pre-treatment controls and identify highly consistent results with our main analysis (A-4-4). Specifically, we find significant and comparable effects of providing the target prompt for those assigned to the gain-loss incentives (Columns 2–4). The interaction term falls out of significance in the main study period with a p-value of .13 but the coefficient estimate is comparable to the estimate in the original analysis. The effect in the post-treatment period is significant and consistent with our main analysis.

[Table A-4-4: Effect of Incentives and Fitbit Target Changes]

VARIABLES	(1) Step Challenge	(2) Short Term	(3) Intermediate Term	(4) Long Term
Choice	402.3 (650.2)	314.1 (628.1)	-205.2 (553.9)	-562.3 (489.5)
Gain-Only Assigned	746.3 (721.8)	-223.8 (781.8)	-1,198 (759.0)	-442.5 (640.5)
Gain-Loss Assigned	948.7 (745.3)	582.1 (655.9)	347.0 (585.9)	126.3 (495.0)
Fitbit Target Prompt	-760.8 (609.8)	-705.6 (575.8)	-892.5* (516.1)	-1,395*** (466.7)
Choice*TargetPrompt	398.0 (894.3)	496.9 (860.5)	893.1 (755.4)	1,406* (743.6)
Gain-Only Assign * TargetPrompt	553.1 (936.0)	983.2 (1,012)	1,877* (980.2)	1,008 (931.6)
Gain-Loss Assign * TargetPrompt	2,217 (1,361)	2,123* (1,127)	1,825* (1,070)	2,484** (1,047)
Individuals	285	283	278	264
Observations	1,962	7,691	13,245	19,276

Robust standard errors in parentheses; ** p<0.01, * p<0.05, * p<0.1; Time fixed effects in all columns

[A-5: Non-linear Models for Goal Achievement Estimates and Impact on Compensation]

[Table A-5-1: Logit / Probit Estimation Results]

VARIABLES	(1) Goal-Challenge Week - Probit	(2) Goal-Challenge Week - Logit	(3) Goal-Long Term-Probit	(4) Goal-Long Term-Logit
Choice	0.310** (0.125)	1.655** (0.337)	0.130 (0.103)	1.243 (0.214)
Gain-Only Assigned	0.278* (0.158)	1.570* (0.404)	0.185 (0.128)	1.376 (0.294)
Gain-Loss Assigned	0.509*** (0.143)	2.303*** (0.544)	0.189* (0.112)	1.381* (0.260)
Constant	0.0948 (0.0931)	1.159 (0.174)	-0.0981 (0.0963)	0.846 (0.132)
Individuals	309	309	277	277
Observations	2,128	2,128	20,013	20,013

Robust standard errors in parentheses; ** p<0.01, * p<0.05, * p<0.1; Time fixed effects in all columns

[Table A-5-2: Differences in Compensation Across Incentive Manipulations]

VARIABLES	Daily Compensation (1)
Choice	0.0329 (0.065)
Gain-Only Assigned	-- --
Gain-Loss Assigned	-0.0416 (0.103)
Constant	0.713*** (0.0517)
Observations	1,392
Individuals	199

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;
Time fixed effects included in all columns

[A-6: Goal Change and IV Analysis]

The analysis of Fitbit health tools currently focuses on the randomized Fitbit target prompt and provides a conservative, intent to treat, estimate of the impact of utilizing the Fitbit health tool (only a subset of our participants change their target in response to our treatment). We augment this analysis with evaluation of how actual changes to Fitbit step targets impacted outcomes and how they interacted with incentive schemes

(Columns 1 and 3). We find that these effects are highly consistent with the results focused on the randomized prompt. This is likely because most of the variation in Fitbit health tool usage was driven by our randomized prompt. However, a challenge with estimating the impact on steps of changes to Fitbit step targets is that these decisions continue to be endogenously made in our empirical context. To address this issue, we utilize our randomized prompt as an instrument for each randomly assigned condition.¹ We employ a 2SLS estimation approach and find large and significant effects for the Gain-Loss condition when individuals change their Fitbit step targets in the short term (Column 2) and in the long term (Column 4).

[Table A-6-1: Impact of Goal Change]

VARIABLES	(1) Challenge Period	(2) IV – Challenge Period	(3) Long Term	(4) IV – Long Term
Choice	512.4 (541.8)	348.0 (785.3)	-445.1 (436.0)	-655.3 (642.5)
Gain-Only Assigned	998.5 (711.8)	583.0 (930.6)	-170.1 (577.2)	-774.5 (766.6)
Gain-Loss Assigned	1,348** (676.0)	528.1 (957.5)	472.7 (542.5)	-412.3 (797.0)
TargetChange	-436.4 (768.7)	-2,168 (1,881)	-1,288** (630.4)	-3,667** (1,630)
Choice*TargetChange	1,103 (1,120)	1,850 (2,639)	2,393*** (906.0)	3,505 (2,171)
Gain-Only Assigned * TargetChange	-102.5 (1,393)	1,631 (2,703)	606.1 (1,157)	2,351 (2,305)
Gain-Loss Assigned*TargetChange	2,319* (1,249)	5,299** (2,670)	2,649*** (990.1)	5,788*** (2,214)
Constant	12,455*** (426.6)	13,176*** (555.7)	12,045*** (401.7)	10,691*** (456.8)
Individual	309	309	277	277
Observations	2,128	2,128	20,013	20,013

Robust standard errors in parentheses; ** p<0.01, * p<0.05, * p<0.1; Time fixed effects in all columns

[A-7: Sample Breakdown and Extended Analysis by Incentive Preference]

In the field experiment, participants tended to prefer the gain-only incentives over the gain-loss incentives. Thus, we provide the sample breakdown by incentive preference. We start by examining the breakdown of the sample for the individuals who prefer the gain-only incentives. We find that there is good spread of participants across those provided their choice and those assigned to an incentive scheme (77 vs. 70

¹ For each interaction term, we instrument using the interaction between the exogenously assigned prompt and the exogenously assigned incentive condition.

individuals). More so, within those assigned to an incentive scheme, there is also good spread across gain-only and gain-loss incentives (34 vs. 36). Further, the panel data structure results in ample observations to conduct our analysis.

[Table A-7-1: Sample Breakdown by Condition for Gain-Only Incentives Preferred]

	Safe Preferred (N)	Gain-Only Preferred (Individuals)
Choice	514	77
Gain-Only Assigned	250	34
Gain-Loss Assigned	234	36
Total	998	147

With fewer participants preferring the gain-loss incentives, the analysis conditioning on Gain-Loss preferred relies on relatively few individuals (see Table A-7-1). More so, while the Choice and Gain-Loss conditions have equal number of individuals, there are relatively few individuals assigned to the Gain-Only incentives. However, we still have reasonable sample size due to the panel structure of our data.

[Table A-7-2: Sample Breakdown by Condition for Gain-Loss Preferred]

	Gain-Loss Preferred (Individuals)	Gain-Loss Preferred (N)
Choice	23	161
Gain-Only Assigned	10	70
Gain-Loss Assigned	19	133

*Note that the control group did not indicate a preference for gain-only vs. gain-loss

We evaluate individual demographics across conditions for the individuals who preferred the gain-only or gain-loss incentives in Tables A-7-3 and A-7-4. We find good balance across these experimental groups with a handful of significant differences which is well within the threshold of significance by random chance for an alpha of .1.

[Table A-7-3: Individual Characteristics by Condition for Gain-Loss Preferred]

	Choice (1)	RNC (2)	SNC (3)	(1) vs. (2) (4)	(1) vs. (3) (5)	(2) vs. (3) (6)
New Goal	10142.26	10297.00	11563.80	0.90	0.34	0.48
Male	0.45	0.37	0.60	0.59	0.46	0.25
Body Mass Index	23.38	23.24	23.09	0.91	0.84	0.91
Age	18.75	18.54	18.37	0.14	0.03	0.33
Extraversion	3.08	3.03	3.38	0.84	0.37	0.23
Agreeableness	3.71	3.81	3.98	0.51	0.21	0.36
Conscientiousness	3.63	3.52	3.67	0.57	0.85	0.45

Neuroticism	2.84	2.79	2.69	0.83	0.60	0.71
Openness	3.45	3.37	3.82	0.71	0.11	0.09

[Table A-7-4: Individual Characteristics by Condition for Gain-Only Incentives Preferred]

	Choice	G-L Assign	Stand. Assign	(1) vs. (2)	(1) vs. (3)	(2) vs. (3)
	(1)	(2)	(3)	(4)	(5)	(6)
New Goal	12857.34	13843.61	13141.88	0.28	0.77	0.60
Male	0.49	0.42	0.53	0.46	0.73	0.35
Body Mass Index	22.58	22.79	22.92	0.69	0.54	0.85
Age	18.55	18.61	18.52	0.40	0.74	0.26
Extraversion	3.20	3.23	3.14	0.86	0.68	0.59
Agreeableness	3.63	3.77	3.75	0.21	0.33	0.85
Conscientiousness	3.60	3.75	3.66	0.21	0.61	0.47
Neuroticism	2.96	2.99	3.05	0.87	0.55	0.71
Openness	3.25	3.34	3.42	0.36	0.09	0.45

Finally, we evaluate the effects while conditioning on pre-treatment variables and demographic characteristics (Table A-7-5). We find consistent results. Of those who prefer the gain-loss incentive, there are minimal differences between incentive conditions. However, for those who preferred the gain-only incentives, only those assigned the gain-loss incentives outperformed the control group and this group continued to do so well after the challenge week ended (Columns 5–8).

[Table A-7-5: Analysis Conditional on Incentives Preferred]

VARIABLES	Gain-Loss Preferred				Gain-Only Preferred			
	(1) Challenge Week	(2) Short Term	(3) Intermediate Term	(4) Long Term	(5) Challenge Week	(6) Short Term	(7) Intermediate Term	(8) Long Term
Choice	741.8 (972.0)	1,610** (716.3)	-93.56 (1,129)	-131.2 (1,164)	-244.3 (540.8)	426.8 (575.2)	879.2 (580.4)	-87.33 (563.3)
Gain-Loss Assigned	969.1 (1,250)	1,218 (1,035)	510.1 (1,218)	755.4 (1,314)	1,347* (801.8)	1,712** (707.0)	1,883** (735.4)	1,337* (683.8)
Constant	-54.89 (17,009)	14,424 (12,291)	4,337 (14,513)	-6,005 (12,456)	-2,405 (15,869)	10,092 (13,877)	13,796 (12,498)	17,646 (12,441)
Individuals	50	50	48	46	145	142	139	129
Observations	350	1,333	2,184	3,248	984	3,836	6,612	9,351
Dem+Pers Controls	YES	YES	YES	YES	YES	YES	YES	YES
Pre-Treat Controls	YES	YES	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses; ** p<0.01, * p<0.05, * p<0.1; Time fixed effects in all columns

[A-8: Field Experiment Extensions and Robustness]

We conducted several robustness checks to verify our main results. For the sake of exposition, we only present the analysis for the long-term post period (results are generally consistent for the short and

intermediate term). We evaluate our effects with the addition of a control for daily $Compliance_{it}$ (measured as the percentage of each day t that participant i wore their Fitbit device) to account for differences in the wearing of the Fitbit device by participants (Columns 1 and 2). We also evaluate our result using negative Poisson regression (Columns 3 and 4). Finally, we evaluate the effect of our manipulations on the log of steps, which addresses potential non-normality in our data (Columns 5 and 6). We find consistent results across these robustness checks.

[Table A-8-1: Extensions and Robustness]

VARIABLES	(1) Challenge Week - Compliance	(2) Long Term - Compliance	(3) Poisson Challenge Week	(4) Poisson Long Term	(5) Log Steps - Challenge Week	(6) Log Steps - Long Term	(7) Goal Failure
Choice	801.7* (469.6)	74.72 (365.5)	0.0577 (0.0368)	0.00159 (0.0400)	0.0859 (0.0532)	-0.0106 (0.0553)	647.9 (562.8)
Gain-Only Assigned	841.8* (502.6)	-19.40 (451.7)	0.0719 (0.0475)	-0.0107 (0.0522)	0.137*** (0.0521)	-0.0448 (0.0714)	1,173* (655.9)
Gain-Loss Assigned	1,641** (689.4)	946.5* (509.8)	0.151*** (0.0439)	0.118** (0.0471)	0.186*** (0.0564)	0.125** (0.0589)	1,650** (769.5)
Goal Fail							-183.6 (357.2)
Goal Fail*G-L Assigned							773.7 (617.4)
Goal Fail*G-O Assigned							-435.8 (658.2)
Goal Fail*Choice							343.0 (537.9)
Compliance	68.72*** (7.805)	69.41*** (3.050)					
Constant	6,426*** (820.2)	6,156*** (383.9)	9.423*** (0.0254)	9.367*** (0.0275)	9.325*** (0.0408)	9.201*** (0.0673)	12,627*** (434.3)
Observations	2,128	20,012	2,128	20,012	2,128	20,013	1,732
Time f.e.	YES	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses; ** $p < 0.01$, *** $p < 0.001$, * $p < 0.05$

[A-9: AMT Experiment Pre-studies and Sample Breakdown]

[Table A-9-1: AMT Pre-Study 1 Results]

Attempted Grids	14.4
Correct Grids	13.74
Total Time	394 seconds (27 seconds per grid attempted)

We recruited 101 participants from AMT for the first pre-study. Participants in this pre-study were provided a protocol identical to the main experiment except that they were not given a new goal, or the incentive schemes associated with it. For example, participants in this pre-study underwent the training, went through

the practice set, and had the option to quit or continue after each grid they solved. The purpose of this pre-study was to identify a baseline for how many grids participants would solve with just the basic incentives in place (i.e., \$.02 cents per grid solved). We found that participants attempted 14.4 grids on average, got 13.7 of those grids correct, and spent 27 seconds on each grid (see Table A-9-1).

To calibrate the incentive amounts to use in the gain-only vs. gain-loss incentive schemes, we recruited an additional 250 AMT participants for a second pre-study. The beginning of the study was identical to the main experiment, with the same introduction, training, and practice set. Unlike the main study, however, participants were not asked to perform the main grid solving task. Rather, we provided participants a grid goal that is 10% greater than the average solved in the first pre-study (15 grids) and asked them to indicate (in hypothetical terms) their preference for the gain-only vs. gain-loss incentive scheme.² Specifically, participants were presented five scenarios where they were asked to choose between a gain-only incentive scheme (labeled “Option 1”) that always rewarded them \$0.25 for meeting their goal (with no penalty), and a gain-loss incentive scheme (labeled “Option 2”) which would reward them a larger amount for meeting the goal but also had a penalty if they missed their goal (see Table A-1-2). We varied the reward associated with the gain-loss incentive scheme between subjects. Because many of the manipulated amounts split participants somewhat evenly, we opted for the more cost-effective option with a \$0.40 penalty (\$0.65 if they met their goal and -\$0.40 if they did not meet the goal – see Table A-1-2).

[Table A-9-2: AMT Pre-Study 2 Results]

Gain-Only Incentive	Gain-Loss Incentive	Gain-Loss Preferred
\$0.25 / \$0	\$0.50/\$-0.30	0.39
	\$0.60/\$-0.30	0.44
	\$0.65/\$-0.40	0.41
	\$0.75/\$-0.40	0.38
	\$0.80/\$-0.40	0.60

Finally, we provide the sample breakdown for the various incentive schemes as well as the breakdown conditional on a particular incentive preference.

[Table A-9-3: Sample Breakdown by Incentive Preference, Lab Experiment]

	Gain-Only Preferred	Gain-Loss Preferred	Total
Control	98	145	243
Gain-Only Assigned	116	148	264
Choice	104	128	232
Gain-Loss Assigned	102	158	260

² We had considered asking participants to indicate a goal for the task but worried that AMT participants may provide a goal strategically (i.e., anticipate that they may be asked to meet that goal in the future and provide a low goal). Also, participants on AMT may not be sufficiently familiar with the grid solving task to appropriately gauge what they considered a challenging goal. Thus, we simply chose a 10% increase from the pre-study baseline.