

**Online Weight Loss Program:
Incentivize Weight Loss or Health Education?**

TECK-HUA HO
CATHERINE YEUNG
NOAH LIM
ROB M. VAN DAM
RYOKO SATO
KWANG-WEI THAM
HONG-CHANG TAN

This is a working paper. Please do not cite or distribute without permission.

Teck-Hua Ho (teck@nus.edu.sg) is the senior deputy president and provost, National University of Singapore. Catherine Yeung (cyeung@cuhk.edu.sg) is associate professor of marketing, Chinese University of Hong Kong. Noah Lim (nlim@bus.wisc.edu) is the John P. Morgridge Distinguished Chair in Business and professor of marketing, Wisconsin School of Business, University of Wisconsin-Madison, and the director of the NUS Global Asia Institute, National University of Singapore. Rob M. van Dam (rob.van.dam@nus.edu.sg) is professor of epidemiology, Saw Swee Hock School of Public Health, National University of Singapore. Ryoko Sato (ryokos1226@gmail.com) is a junior professional officer at the World Bank (Poverty and Equity Global Practice). Kwang-Wei Tham (tham.kwang.wei@singhealth.com.sg) is a senior consultant, Department of Endocrinology, and the director of the LIFE Center and the Obesity and Metabolic Unit, Singapore General Hospital. Hong-Chang Tan (tan.hong.chang@singhealth.com.sg) is a consultant, Department of Endocrinology, Singapore General Hospital. Address correspondence to Teck Ho and Catherine Yeung.

Teck Ho and Catherine Yeung developed the study concept. Teck Ho, Catherine Yeung, Rob van Dam, Kwang-Wei Tham, and Hong-Chang Tan contributed to the study design. Teck Ho, Catherine Yeung, Rob van Dam, and Hong-Chang Tan contributed to the production of the online weight loss program. Data collection was performed by Ryoko Sato, Teck Ho, and Catherine Yeung. Catherine Yeung and Noah Lim performed the data analysis and interpretation. Catherine Yeung and Noah Lim wrote the manuscript and Teck Ho provided critical revisions. All authors provided comments and approved the manuscript for submission.

We thank the psychologists, physiotherapists, and dieticians of the Singapore General Hospital LIFE Centre; Richard Lau, Cindy Tay, and the staff and student helpers at the NUS Centre for Behavioral Economics; and Rehan Ali for his assistance in the production of the online weight loss program and editorial help.

ABSTRACT

In a randomized field experiment, we examined whether an incentive for weight loss worked better than an incentive for education in promoting weight loss among overweight individuals. 710 overweight individuals participated in a 3-month self-administered online weight loss program. We offered either a \$150 “incentive for weight loss” (IW) for achieving 5% weight loss, or a \$150 “incentive for education” (IE) for passing quizzes on health and nutrition literacy based on lessons taught throughout the program. IE resulted in improved health literacy but did not result in more people losing weight than in a no-incentive control condition. In contrast, IW more than tripled the proportion of males who met the weight loss target compared to the control condition. Weight loss for participants in IW persisted during a 3-month post-intervention period when no incentive was offered. Interestingly, the IW incentive had no such effect on females. This is the first RCT to compare the effects of IE and IW on weight loss and the first to demonstrate gender differences in response to weight loss incentives. Recognizing these differences is crucial considering the increasing use of financial incentives in commercial and public weight loss programs.

Keywords: financial incentives, health literacy, behavioral change, weight loss, scalable obesity prevention intervention

Over a third of the world's population is either overweight or obese today (Ng et al. 2014; Stevens et al. 2012). By 2030, an estimated 58% of the world's adult population will be overweight or obese (Kelly et al. 2008). Obesity greatly increases the risks of getting many chronic non-communicable diseases, including Type II diabetes, cardiovascular disease, hypertension, kidney disease, and some types of cancer. Because the increasing prevalence of obesity will lead to substantial disease burdens on many societies, governments are increasingly looking for innovative approaches to combat obesity at the community level. Thus, interventions that can be scaled up to promote weight loss for large numbers of people are of particular interest to policymakers and companies that run weight loss programs.

This research investigates whether participants in a fully self-administered online weight loss program can be induced to achieve a weight loss goal using financial incentives. We conducted a 3-condition randomized field experiment (also known as randomized controlled trial, RCT) stratified by gender to examine how overweight individuals respond to two different types of incentives. The first type of incentive, IW (incentive for weight loss), was a cash award of US\$150 given for achieving 5% weight loss over 12 weeks. The second type of incentive, IE (incentive for education), involved a US\$150 award for passing quizzes that tested participants' literacy about the risks of being overweight and strategies for weight reduction. We compare the effectiveness of IW to IE in promoting weight loss in overweight individuals and investigate whether the two incentives work equally well for men and women.

RESEARCH BACKGROUND

Weight loss programs that have the capacity to reach a large number of overweight individuals must be fundamentally self-directed (i.e., involve little or no healthcare provider participation). The challenge therefore is to motivate actual weight reduction in the context of a self-help program. Governments, not-for-profit organizations, and companies have explored the use of financial incentives to motivate weight loss in such programs. For example, commercial programs such as Dietbet (www.dietbet.com) and Fatbet (fetbet.net) motivate their clients by getting them to bet on meeting certain weight loss goals; the clients get their money back and make additional money if they achieve their goals, but they lose their money if they fail. This type of incentive has received strong empirical support from RCTs reported in Volpp et al. (2008) and John et al., (2011). In the two studies, participants who placed a 1:1 bet on meeting the target of losing 1 pound a week (for 16 and 24 weeks, respectively) lost more weight than those in a control condition who were not offered the opportunity to bet. Nevertheless, because betting requires people to put their own money at stake, concerns about its use as an incentive for weight loss increase when the audience is diverse in financial status and ethnicity (Consensus Statement of the Health Enhancement Research Organization, 2012).

A direct positive reward is another type of weight loss incentive, often used in interventions targeting a wide community. For example, in the “Weigh and Win” program currently run by Kaiser Permanente, and the “Pounds for Pounds” pilot program funded by the UK National Health Service from 2009 to 2010, cash rewards are (were) tied directly to each individual participant’s weight loss result (e.g., win \$X upon meeting a specific weight loss target during a specific period). Participants in the Pounds-for-Pounds program lost an average of 3.9% of their baseline weight in 12 months, but the findings must be interpreted with the caveat that there was no control group where participants did not receive any financial incentives (Relton, Strong and Li 2011).

To date, three RCTs have been conducted to evaluate the clinical effectiveness of direct positive rewards on weight loss, but the results are mixed. In Finkelstein et al. (2007), participants who were offered a reward of \$14 for each percentage point of weight lost achieved an average of 2.3% weight loss in three months, but the effect of the incentive was not evaluated with respect to a no-incentive control condition. In Kullgren et al. (2013), participants who were offered \$100 per person per month to lose 1.6 kg per month (for 6 months) did not lose more weight than those in a no-incentive control condition. In Finkelstein et al. (2017), participants who received a reward of up to \$660 (Singapore dollars) to lose weight and exercise lost more weight than their counterparts in the control condition. Nevertheless, the incentive was layered over an out-patient program consisted of intensive medical, diet, and nutrition management; therefore, the findings may not generalize to the context of self-administered weight loss.

Overall, existing research provides some initial evidence that direct positive incentives might work to promote weight loss, but the mixed results strongly indicate that further research is warranted to better understand how such incentives work and when does it work.

THE CURRENT RESEARCH

In our research, we examined whether participants in a fully self-administered online weight loss program could be induced to achieve a weight loss goal using a direct, positive incentive for weight loss. We asked two specific questions that would enhance our understanding of how the incentive works. The first question we ask is whether we can promote weight loss more effectively if we tie positive financial incentives to attaining health literacy rather than to achieving weight loss. Many healthcare leaders believe that one way to save lives is to improve health literacy on obesity prevention (for example, see Carmona 2005). If people are unaware that obesity is a risk factor for diseases they fear (e.g., kidney disease, heart disease), or if they underestimate the benefits associated with weight loss but overestimate the difficulty of weight loss, they will not find a compelling reason to lose weight and will hence adopt a lifestyle that is detrimental to their long-term health. Health literacy education informs people that even

modest weight loss (e.g., 5%) can cut the risk of getting diabetes. It helps participants to overcome the difficulty of weight loss by recommending small steps that are easy to implement and by relying on habit formation to sustain a healthy lifestyle over time. It also educates the use of clinically proven behavioral modification strategies (such as the antecedent-behavior-consequence analysis) to help people resolve behavioral problems that could lead to an unhealthy lifestyle. With this array of knowledge, people may undertake the behavioral modifications that are necessary for weight reduction (Bandura 1986, 2004). Our IE condition was designed specifically to test whether an incentive for improving health literacy can lead to weight loss, and whether it works better than IW in promoting weight loss. To the best of our knowledge, all past studies focused exclusively on directly tying financial incentives to weight loss, and our RCT is the first study to examine whether financial incentives for attaining high health literacy can be more effective than an incentive tied to weight loss in promoting weight reduction.

The second question we ask is whether there is a gender difference in responding to incentives that promote weight loss. Across the world, women live longer than men (CIA 2018). Women are more likely to assume a dominant role in taking care of their family's health (Carpenter 1980; Rhoads and Rhoads 2012). Younger women also exhibit stronger preferences than men to remain trim (Feingold and Mazzella 1998). Given that there are gender differences in attitudes, behavior and outcomes related to health, it is possible that men and women might respond differently to incentives for weight loss.

Past RCTs on incentives for weight loss¹ have focused on either predominantly (i.e., > 75%) male (John et al. 2011; Volpp et al. 2008) or female participants (Finkelstein et al. 2007; Kullgren et al. 2013). Analysis of gender difference is impossible in these studies given their sample sizes. In domains other than weight loss, research has documented that males prefer riskier and competition-based incentives than females (Croson and Gneezy 2009; Eckel and Crossman 2008; Saccardo, Petrisz, and Gneezy 2017), but since both our IE and IW are direct positive rewards that contain neither probabilistic nor competitive elements, it was unclear, *ex ante*, whether the two genders would respond differently to these incentives. Our research is the first systematic study that investigates gender differences in how people respond to such incentives. The likely heterogeneous responses will allow weight management program practitioners and policymakers to craft more effective approaches to motivate healthy behavior.

¹ To date, only five RCTs have been conducted to test the effects of financial incentives on inducing weight loss. They are Kullgren et al. (2013), Volpp et al. (2008), John et al. (2011), Finkelstein et al. (2007), and Finkelstein et al. (2017). The sample sizes were 105, 57, 66, 207, and 161, respectively (all employed a three-condition design). In comparison, our study has a sample size of 710, making it the largest RCT to examine the effect of incentives on weight loss. Studies conducted based on workplace wellness programs tend to have larger sample sizes (e.g., Cawley [2013] with $N = 2,635$ and Misra-Hebert et al. [2016] with $N = 1,092$), but these findings must be interpreted with the caveat that randomized designs were not used in these studies. Other weight loss studies that did not use a randomized design or did not use an intent-to-treat analysis are not discussed further (for further discussion, see Ananthapavan, Peterson, and Sacks 2017).

METHOD

The Online Weight Loss Program

The program was a 12-week, self-administered online weight loss program based on the University of Pittsburgh's Lifestyle Balance Program (The DPP Research Group, 2002). The following topics were covered in the program, in sequence: (1) The risks of being overweight; (2) Be a smart eater; (3) Healthy eating; (4) Move those muscles; (5) Tip the calorie balance; (6) Take charge of what's around you; (7) Problem solving; (8) Healthy eating while out (9) The slippery slope of lifestyle change; (10) Jump start your physical activity plan; (11) Eating and exercising while away; (12) Preparing for self-management.

Since the RCT was conducted in Singapore, the content was adapted for the local culture, diet, and lifestyle; it was also modified to suit bite-size e-learning. The material was developed by the research team and verified by a clinical team consisting of an endocrinologist, a psychologist, a physiotherapist, and two dietitians. Because our primary interest is in cost-effective, scalable weight loss programs, we deliberately excluded features such as online meetings and personalized feedback from dietitians, even though such features may be effective in promoting weight loss (Gold et al. 2007; Tate, Wing, and Winnett 2001).

The program was delivered through 12 weekly sessions; each included educational videos and a health literacy quiz. Each quiz contained 10 questions, and a quiz was considered "passed" when at least seven questions were answered correctly. All participants were enrolled in the same online program, had access to the videos and quizzes, and were tasked to lose 5% of their baseline body weight by the end of the program.

RCT Design and Timeline

There were three experimental conditions in the RCT: 1) Control condition with no financial incentive for weight loss or passing quizzes; 2) IE condition with a \$150 incentive for passing at least nine quizzes by the end of the program; and 3) IW condition with a \$150 incentive for losing at least 5% of baseline weight by the end of the program. In other words, although all participants were given the goal of losing 5% of weight and had access to the online program and its quizzes, only those in the IE condition were incentivized to pass the quizzes, and only those in the IW condition were incentivized to meet the 5% weight loss goal. Figure 1 shows the concept behind the three conditions.

The RCT consisted of two periods: a 12 week intervention, and a 12 week post-intervention (figure 2). Participant weight was measured by trained research assistants at three points: right before the intervention (at Week 0), at the end of the intervention (at Week 13), and at the end of the post-

intervention period (Week 25). All participants received \$65 as a participation fee for attending all 3 weigh-ins. This amount was split into two payments: \$15 was paid at the Week 13 weigh-in and \$50, at the Week 25 weigh-in. The two incentives—IW and IE—were tied to attendance at all 3 weigh-ins, and were paid at the Week 25 weigh-in.

FIGURE 1. Diagram explaining the concept behind the RCT design

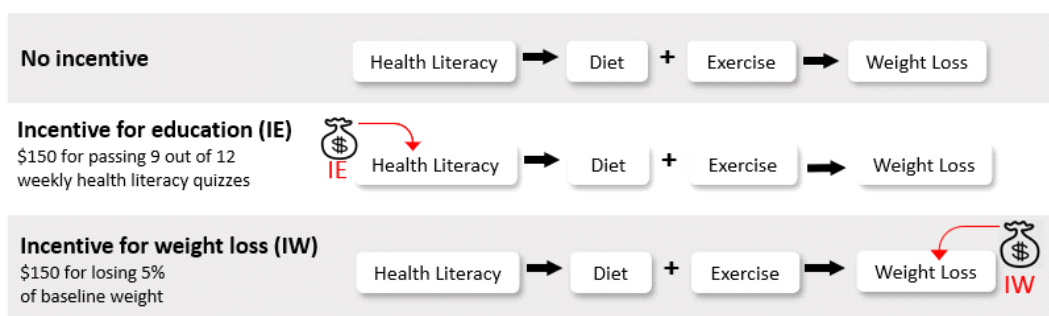
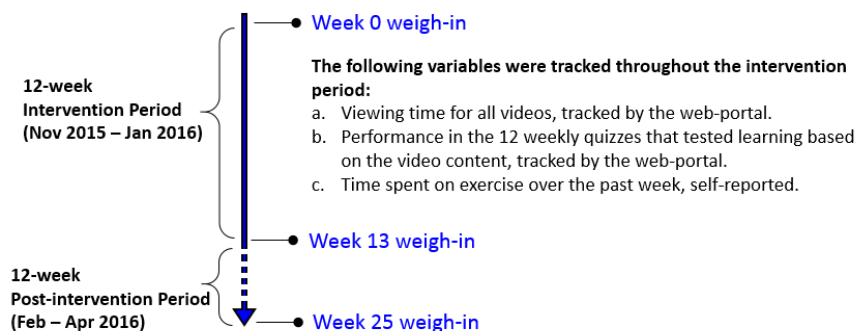


FIGURE 2. Timeline of RCT



Participants, Procedure, and Randomization

Participants were recruited from the public through newspaper advertisements. They had to be between 40 and 60 years old, and overweight or obese. For the complete list of selection criteria, see Appendix A. Potential participants were invited to visit a website for an initial screening. Those who passed the screening were invited to attend a weigh-in session at Week 0, where they had their weight and height measured, age verified, and other baseline measures collected. They were advised to log in to the program's website after the weigh-in.

Randomization took place when the participants first logged on to the website. The randomization sequence followed a stratified block randomization scheme with 8 gender-ethnicity strata [2 (Male,

Female) \times 4 (Chinese, Malay, Indian, Other)]. In other words, the participants were first sorted by gender and ethnicity, and then randomly assigned to one of three conditions. After randomization, information on the additional financial incentives offered in the IW and IE conditions was shown to participants in those conditions via an automated message shown onscreen.

Outcome and Process Measures

The first outcome of interest was weight loss after 12 weeks (measured at Week 13), which reveals whether IW or IE led to greater weight loss. To provide an unbiased comparison of weight loss between genders, we chose percentage weight loss rather than absolute weight loss as the unit of analysis because the former accounts for baseline weight differences between the two genders. Percentage weight loss is defined as the difference in weight between Week 13 and Week 0, divided by weight at Week 0. The second outcome of interest was percentage weight loss at Week 25, which reveals whether the results sustained when the incentives were removed. Both Week 13 and 25 weight loss outcomes were analyzed using the intent-to-treat (ITT) approach: participants who did not attend the respective weigh-in were still included in the analysis and were treated as having the same weight as at Week 0. A supplementary analysis was conducted using the per-protocol (PP) approach, which included only the participants who attended all three weigh-ins.

To provide insight into how the incentives may affect weight loss, we also collected the number of videos viewed, number of quizzes passed, and weekly exercise level. The number of videos viewed and quizzes passed were tracked by the website throughout the 12-week intervention, and were analyzed using the ITT approach. Exercise level was self-reported; each week, the website prompted participants to report the number of minutes they exercised over the previous week. These self-reports were voluntary; a total of 300 participants reported their exercise for each of the 12 weeks, while 29 participants did not provide any reports at all. We reported analysis results of exercise in Week 1 and average weekly exercise for the intervention period. Average weekly exercise was computed for a participant only when three or more self-reports were logged. Because no baseline measure for exercise duration was collected, an ITT analysis of the exercise duration data was not possible.

RESULTS

Attrition and Participant Characteristics

A total of 710 participants were randomly assigned to one of the three experimental conditions (males: Control = 85, IE = 87, IW = 86; females: Control = 150, IE = 151, IW = 151). A total of 574 attended both Week 13 and Week 25 weigh-ins (males: Control = 66, IE = 69, IW = 76; females: Control

= 119, IE = 120, IW = 124). Details on participant flow and an attrition analysis are in Appendix A. To avoid selection bias due to attrition, we conducted our primary analysis using the intent-to-treat (ITT) approach by including all the 710 participants in the analysis.

We recruited a total of 710 participants (258 males). Our male participants were 47.7 years old on average (SD = 5.1), with an average BMI of 27 kg/m² (SD = 2.6). Our female participants were 48.9 years old on average (SD = 5.5), with an average BMI of 26.6 kg/m² (SD = 2.5). Other participant characteristics are shown in Appendix A. There were slight differences between the male and female participants in terms of age, BMI, and education levels, but within each gender, there were no differences in these aspects across the three conditions. To ensure the robustness of our findings, we report sensitivity analyses in which baseline BMI, age, and educational level were included as covariates.

Weight Loss at Week 13

We first evaluated the effects of each incentive type on male and female weight loss using the ITT approach. In table 1, the top panel shows average weight loss in each treatment condition by gender. In figure 3, the left panel shows these outcomes graphically. An ANOVA yielded a significant main effect of incentive type ($F(2, 704) = 3.87, p = .02, \eta_p^2 = .01$), a non-significant main effect of gender ($F(1, 704) = 0.42, p = .52$), and a significant incentive type \times gender interaction ($F(2, 704) = 9.05, p < .001, \eta_p^2 = .03$). Pairwise comparisons showed that for males, average weight loss was higher in the IW ($M = 2.40$) than in the Control ($M = 0.87$), ($t(704) = 3.83, 95\% \text{ CI: } [0.74, 2.32], p < .001, \eta_p^2 = .02$) and the IE ($M = 0.87$) ($t(704) = 3.83, 95\% \text{ CI: } [0.74, 2.31], p < .001, \eta_p^2 = .02$) conditions, indicating that the direct financial incentive promoted greater weight loss among males. Notice that weight loss in the IE and Control conditions was identical at 0.87%, suggesting that IE did not motivate males to lose weight. For females, we did not detect any difference in weight loss across the conditions.

Next, we compare weight loss across gender. In the IW condition, weight loss was higher for males than for females (2.40 vs. 1.03) ($t(704) = 3.81, 95\% \text{ CI: } [0.67, 2.07], p < .001, \eta_p^2 = .02$). Note that in the Control condition, there was no difference in weight loss between genders ($t(704) = 1.58, p = .11$). This suggests that without a financial incentive, males and females in our trial did not differ in their motivation to lose weight.

The regression-adjusted effects of IW and IE on Week 13 weight loss are similar to those reported above without regression adjustment. See Web-Appendix table A1 for the regression-adjusted mean effects (along with their 95% CI). In estimating regression-adjusted effects, we controlled for baseline BMI, age, and educational level since the males and females in our sample differed slightly in these attributes. The following robustness checks are also reported: (1) regression results using the per-protocol approach (in table 1, and Web-Appendix tables A1 and A2) and (2) regression models using

heteroscedasticity-consistent standard errors (Web-Appendix table A3). Importantly, they show that the finding that males lost more weight in the IW condition is robust using both the ITT and per-protocol approaches, and independent of participants' baseline BMI, age, and education level.

TABLE 1. Percentage weight loss at Week 13 and Week 25

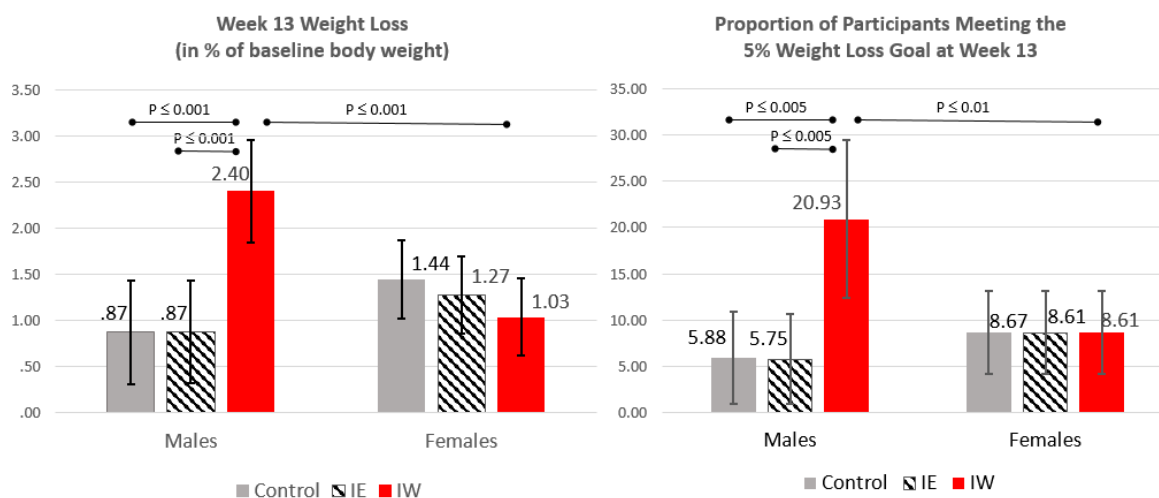
	Males			Females		
	Control	IE	IW	Control	IE	IW
Weight loss, Week 13						
Full Sample						
Mean	0.87	0.87	2.40*** ^{cγ}	1.44	1.27	1.03
95% CI	0.31 – 1.43	0.32 – 1.43	1.84 – 2.96	1.02 – 1.87	0.85 – 1.69	0.61 – 1.45
Per-Protocol Sample						
Mean	1.12	1.00	2.57*** ^{cγ}	1.83	1.53	1.26
95% CI	0.44 – 1.81	0.33 – 1.68	1.93 – 3.21	1.32 – 2.35	1.02 – 2.04	0.76 – 1.76
Weight loss, Week 25						
Full Sample						
Mean	0.88	1.36	2.53*** ^a	1.55	1.46	1.66
95% CI	0.16 – 1.61	0.64 – 2.08	1.80 – 3.25	1.01 – 2.10	0.91 – 2.00	1.11 – 2.20
Per-Protocol Sample						
Mean	1.14	1.69	2.82**	2.14	1.84	2.00
95% CI	0.27 – 2.02	0.83 – 2.55	2.01 – 3.64	1.49 – 2.79	1.19 – 2.49	1.36 – 2.64

Note. The full sample consists of all 710 participants. The per-protocol subsample consists of the 574 participants who attended all the weigh-ins. *, **, *** indicate $p \leq .05$, $p \leq .01$, and $p \leq .005$, respectively, for comparison with the same-sex Control condition.

^{a, b, c} indicate $p \leq .05$, $p \leq .01$, and $p \leq .005$, respectively, for comparison with the same-sex IE condition.

^{α, β, γ} indicate $p \leq .05$, $p \leq .01$ and $p \leq .005$, respectively, for comparison with the same condition in the opposite gender.

FIGURE 3. Weight loss at Week 13



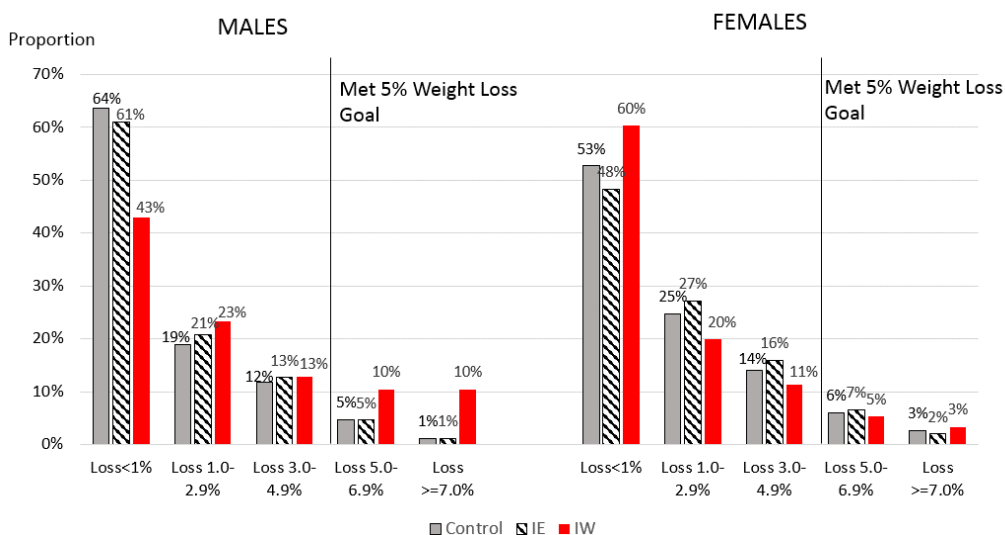
Error bars show 95% CI. $N = 710$.

Weight Loss Target Met at Week 13

Out of the 710 participants, 67 (28 males, 39 females) met the 5% weight loss target. In figure 2, the right panel displays the proportion of participants in each condition who met the target. We fitted a binary logistic regression on success in meeting the target and conducted pairwise comparisons based on the model. We found that 20.93% of males in the IW condition achieved the weight loss target, higher than the 5.88% in the Control and the 5.75% in the IE conditions; both differences were significant (IW vs. Control: $\chi^2(1) = 8.78$, 95% CI: [5.10, 24.99], $p = .003$, OR = 4.24; IW vs. IE: $\chi^2(1) = 9.04$, 95% CI: [5.29, 25.07], $p = .003$, OR = 4.34). For females, there were no differences across the three conditions, with 8.67%, 8.61% and 8.61% of participants meeting the target in the Control, IE and IW conditions, respectively.

We next examined whether the greater weight loss among males in the IW condition was driven only by those who met the weight loss target. See figure 4, which shows the proportion of participants by various weight loss outcomes in each condition. Notice that the proportion of males who recorded weight loss of 1% or less (this includes participants who gained weight) was lower in the IW condition (43.02%) than the Control (63.53%) ($z = 2.69$, $p = .007$) and IE (60.92%) ($z = 2.36$, $p = .02$) conditions. Moreover, an ANOVA conducted based on the subsample of participants who did not meet the 5% target (230 males, 413 females) showed that the incentive type \times gender interaction was still significant ($F(2, 637) = 3.82$, $p = 0.02$, $\eta_p^2 = .01$), and percentage weight loss was still higher among males in the IW ($M = 1.09$) than males in the Control ($M = 0.44$) condition, $t(637) = 2.00$, 95% CI: [0.01 to 1.29], $p = .05$, $\eta_p^2 = .006$. These findings show that the greater weight loss in the IW condition achieved by males was driven not only by participants who met the weight loss target, but also those who did not.

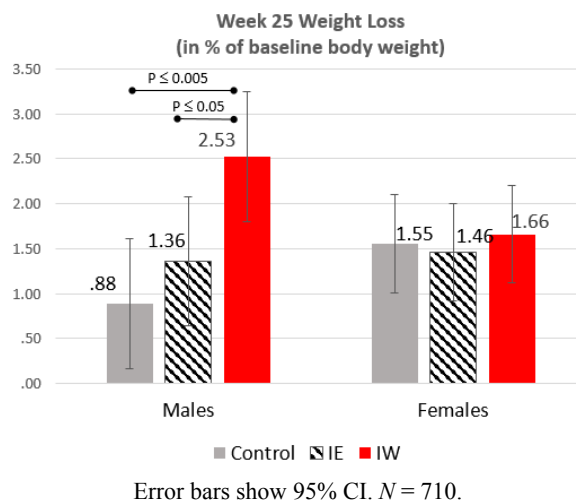
FIGURE 4. Proportion of participants by weight loss outcomes at Week 13



Weight Loss at Week 25 (Post-Intervention)

In table 1, the bottom panel reports average weight loss at Week 25 in each treatment condition by gender. Figure 5 shows these outcomes graphically. The result of an ANOVA showed that the main effect of incentive type was significant ($F(2, 704) = 3.98, p = .02, \eta_p^2 = .01$), the main effect of gender was not significant ($F(1, 704) = 0.02, p = .90$), and the incentive type \times gender interaction was marginally significant ($F(2, 704) = 2.83, p = .06$). Pairwise comparisons showed that weight loss in the IW condition ($M = 2.53$) remained higher than in the Control condition ($M = 0.88$) ($t(704) = 3.15, 95\% \text{ CI: } [0.62, 2.67], p = .002, \eta_p^2 = .01$) and the IE condition ($M = 1.36$) ($t(704) = 2.23, 95\% \text{ CI: } [0.15, 2.18], p = .03, \eta_p^2 = .01$). In other words, for males, the effect of IW sustained for at least three months after the incentive was removed. For females, as in Week 13, there remain no differences in weight loss at Week 25 across the three conditions.

FIGURE 5. Weight loss at Week 25 by treatment and gender



Weight Regain after Loss

To understand how IE and IW might affect weight cycling (the losing and regaining of weight), we next focus on the subsample of participants who lost weight at Week 13 and examine their post-intervention weight regain. Table 2 shows the proportion of males and females in each condition who kept their weight constant, regained weight, and achieved further weight loss during the post-intervention. (Refer to table 2 for details on the computation of these proportions.) Among all the males in the Control condition, 36.47% of them lost at least 1% of weight at Week 13, of which, 54.8% regained weight. Compared to the Control condition, the IE condition had a similar proportion of males who lost weight at

Week 13, but it had a substantially lower proportion of males regaining weight post-intervention (20.59% vs. 54.84%); a pairwise comparison conducted based on a binary logistic regression using weight regain (regained vs. not) as the dependent variable suggested that this difference was statistically significant ($\chi^2(1) = 9.17$, 95% CI [12.08, 56.42], $p = .003$, OR = .21).

We also notice that even though the IW condition had a greater proportion of males who lost weight at Week 13 than the Control condition, the proportion of males who regained weight was not higher (in fact, it was directionally lower) in the IW than the Control condition (IW = 36.73% vs. Control = 54.84%; $\chi^2 = 2.57$, 95% CI: [-4.01, 40.22], $p = .11$). These findings suggest that IW did not increase weight cycling, and that IE could potentially reduce it.

For females, there were no differences across the three conditions in the proportion of participants who regained weight.

TABLE 2. Post-intervention weight change conditional on weight loss

	Males			Females		
	Control (<i>n</i> = 85)	IE (<i>n</i> = 87)	IW (<i>n</i> = 86)	Control (<i>n</i> = 150)	IE (<i>n</i> = 151)	IW (<i>n</i> = 151)
Participants who lost at least 1% of baseline weight at Week 13						
Numbers	31	34	49	71	78	60
%	36.47%	39.08%	56.98%** a	47.33%	51.66%	39.74%
Of those who lost weight at week 13, proportion of participants who:						
Regained weight (Diff > +0.5%) [^]	54.84%	20.59%***	36.73%	40.85%	38.46%	31.67%
Kept weight constant (Diff = 0 ± .5%) [^]	12.90%	26.47%	18.37%	19.72%	20.51%	28.33%
Lost more weight (Diff < -0.5%) [^]	32.26%	52.94%	44.90%	39.44%	41.03%	40.00%

[^]Diff = [(Week 25 weight – Week 13 weight) / Week 13 weight]*100%

*, **, *** indicate $P \leq .05$, $P \leq .01$, and $P \leq .005$, respectively, for comparison with the same-sex Control condition.

^{a, b, c} indicate $P \leq .05$, $P \leq .01$, and $P \leq .005$, respectively, for comparison with the same-sex IE condition.

Process Measures

Video Viewing and Quizzes Passed. Table 3 reports the number of videos watched, classified by the proportion of participants who viewed no videos to completion and the average number of videos completed, conditional on completing at least one video. Note that a substantial number of participants did not view any videos to completion; this ranged from 22% to 42% across the three conditions. Pairwise comparisons conducted based on a logistic regression showed that the proportion of males who completed no videos was lower in the IE (25.29%) than the Control condition (42.35%), $\chi^2 = 5.78$, 95% CI: [3.15 to 30.99], $p = .02$, OR = .46; this is unsurprising given that an incentive was offered in the IE condition to pass the quizzes. Interestingly, this proportion was also lower in the IW (24.42%) than in the Control

condition ($\chi^2 = 6.41$, 95% CI: [4.05, 31.82], $p = .01$, OR = .44), despite no reward being offered for passing quizzes in the IW condition. In fact, no difference was observed in video viewing between the IW and IE conditions (Diff = 0.9%, $p = .90$). Conditional on completing at least one video, there were no differences across the three conditions in the average number of videos completed.

TABLE 3. Video viewing, quizzes passed, and exercise duration

	Males			Females		
	Control	IE	IW	Control	IE	IW
Video Viewing¹						
Proportion of no videos completed						
Mean ²	42.35%	25.29%*	24.42%**	29.33% ^a	21.85%	24.50%
95% CI	32.33–53.04	17.27–35.44	16.50–34.57	22.61–37.10	15.97–29.15	18.30–31.99
Sample size	85	87	86	150	151	151
No. of videos completed, conditional on having completed at least 1 video						
Mean ³	12.37	12.95	13.29	13.42	14.72	13.70
95% CI	9.99–14.74	10.89–15.01	11.23–15.35	11.80–15.03	13.19–16.25	12.15–15.26
Sample size	49	65	65	106	118	114
Weekly Quizzes Passed						
Number of quizzes passed						
Mean ³	5.31	6.94*	6.81	6.24	7.15	6.44
95% CI	4.18–6.43	5.83–8.05	5.70–7.93	5.39–7.09	6.31–8.00	5.60–7.29
Exercise Duration (min)						
First week						
Mean ³	105.70	107.27	223.84** ^{c,γ}	107.79	140.99	119.02
95% CI	43.3–168.1	49.2–165.3	166.2–281.5	62.1–153.5	96.5–185.5	73.4–164.7
Sample size	59	68	69	110	116	110
Average, 12 weeks						
Mean ³	141.63	143.94	204.05* ^a	169.24	145.01	176.43
95% CI	101.5–181.8	105.9–182.0	166.0–242.1	140.4–198.1	115.3–174.7	147.9–204.9
Sample size	71	79	79	138	130	141

Note.

¹ There were 25 videos in total (two videos for each of the 12 weeks, plus one introductory video); for each video, the length of content viewed was tracked and a video was considered “viewed to completion” if a participant watched more than 95% of it. The findings hold when different cutoffs (90%, 97%) were used.

² Statistical tests were conducted based on a binary logistic regression using treatment and gender as independent variables.

³ Statistical tests were conducted based on an ordinary least square regression using treatment and gender as independent variables.

*, **, *** indicate $P \leq .05$, $P \leq .01$, and $P \leq .005$, respectively, for comparison with the same-sex Control condition.

^a, ^b, ^c indicate $P \leq .05$, $P \leq .01$, and $P \leq .005$, respectively, for comparison with the same-sex IE condition.

^α, ^β, ^γ indicate $P \leq .05$, $P \leq .01$ and $P \leq .005$, respectively, for comparison with the same treatment across genders.

Except for weekly exercise, analyses of all the above measures were conducted using the ITT approach.

Table 3 also reports the average number of quizzes passed by participants. For males, the average of 6.94 quizzes passed was higher in the IE condition than the 5.31 quizzes passed in the Control condition ($t(704) = 2.02$, 95% CI: [0.06, 3.22], $p = .04$, $\eta_p^2 = .01$), but again interestingly, no different from the 6.81 quizzes passed by participants in the IW condition ($t(704) = 0.16$, $p = .87$). Thus, while

incentivizing health literacy (IE) leads to improve literacy as captured by the higher number of quizzes passed, directly incentivizing weight loss (IW) yields the same level of literacy, while achieving greater weight loss.

For females, there were no differences across the three conditions for video viewing and quizzes passed.

Exercise Duration. Table 3 shows exercise duration (in minutes) at Week 1 of the intervention and average exercise duration across the 12-week intervention. For males, exercise duration in Week 1 was higher in the IW condition ($M = 223.84$) than in the Control ($M = 105.70$) ($t(526) = 2.73$, 95% CI: [33.22, 203.07]; $p = .006$, $\eta_p^2 = .01$) and IE ($M = 107.27$) ($t(526) = 2.80$, 95% CI: [34.74, 198.41]; $p = .005$, $\eta_p^2 = .02$) conditions. This pattern remained for the entire intervention period: The average weekly exercise duration was higher in the IW ($M = 204.05$) than in the Control ($M = 141.63$) ($t(632) = 2.21$, 95% CI: [7.07, 117.76], $p = .03$, $\eta_p^2 = .01$) and IE ($M = 143.94$) ($t(632) = 2.19$, 95% CI: [6.27, 113.96], $p = .03$, $\eta_p^2 = .01$) conditions. A bootstrapping mediation analysis using 5,000 samples revealed that the 12-week average exercise duration partially mediated the relationship between the IW condition and weight loss (95% CI of indirect effect: [0.0002, 0.0059]). In other words, one reason why there was greater weight loss in the IW condition was because males in that condition exercised more than those in the Control and IE conditions. For females, there were no differences in exercise duration across the three conditions and across the intervention period.

GENERAL DISCUSSION

This paper shows that offering overweight males and females a significant financial incentive to meet a weight loss target (IW) works only for males. Among our male participants, weight reduction occurred for those who met the 5% weight loss target, *and* those who did not. Remarkably, our RCT reported a statistically significant difference in weight loss between the IW and Control conditions at the end of the post-intervention period. (For comparison, see Volpp et al. [2008] and John et al. [2011].)

Incentive for Health Education versus Incentive for Weight Loss

We find that the financial incentive for attaining a high level of health literacy (IE) did not lead to greater weight reduction in either gender. For males, the IE incentive did indeed motivate them to improve their education on how to lose weight, but the acquired knowledge did not translate into actual weight loss. In fact, in the IW condition where there was no incentive for improving health literacy, males attained the same level of health literacy as in the IE condition, while losing more weight. Moreover, the greater weight loss in the IW condition relative to the IE condition persisted 12 weeks after the incentive

for weight loss was removed. These findings suggest incentives linked directly to weight loss targets are more effective than those linked to health literacy.

Our findings are consistent with the perspective that health literacy alone is not sufficient for inducing the behavioral modifications required for weight loss, because people have present-bias preferences. When people are present-biased, the immediate costs of undertaking a weight loss regime are more salient in comparison with its future benefits, so people procrastinate in losing weight (Loewenstein and Prelec 1992; Thaler and Loewenstein 1992). The IW incentive serves to add an immediate reward for weight loss, thereby tipping the cost-benefit tradeoff for behavioral change. This does not mean that improved health literacy is not important however, as there is preliminary evidence that improved health literacy may help to sustain weight loss among those participants who are motivated to lose weight in the absence of financial incentives: Among the males in the IE and Control conditions who lost weight during the intervention period, we found that fewer participants regained weight in the IE condition. While beyond the scope of this paper, the relationship between health literacy and weight loss sustainability warrants further investigation.

Gender Difference in Responses to Incentive for Weight Loss

Our results show that the IW incentive does not induce females to lose more weight. Since our IW contained neither probabilistic nor competitive elements, one would not have expected, *ex ante*, the two genders to respond to it differently. In fact, since IW was tied directly to an individual's weight loss outcome and the two genders in our trial were not different in their baseline motivation to lose weight, it would be reasonable to predict similar responses from the two genders. This was clearly not the case. By providing strong evidence of gender differences towards IW—the most basic form of weight loss incentive—our research highlights to both policymakers and consumer behavior researchers that applying the same type of weight loss incentive to both genders may be problematic, and that it is critical to understand potential differences in what motivates males and females to lose weight.

The gender difference we observed was likely due to multifaceted and intersecting sociocultural factors. We discuss three factors and their possible influences with the objective to stimulate further research on this topic. First, in terms of exercise propensity, evidence from a large body of research shows that men exercise more than women in terms of frequency, duration, and intensity (Cooper et al. 2000; Hawkes and Holm 1993; Jago et al. 2005). The disparity has biological, sociological, and evolutionary underpinnings (Deaner, Balish, and Lombardo 2016); it has been observed not only in adults and adolescents, but also in pre-school children (Deaner et al. 2012; Pate et al. 2004) and the elderly above the age of 70 (Davis et al. 2011). Even among the elderly with chronic diseases, females are still found to be less likely than their male counterparts to achieve the recommended energy expenditure (Lin et al. 2010).

We found in our RCT that the effect of IW on weight loss in males was partially mediated by the males' higher exercise level in the IW condition than the Control condition. Importantly, this difference in exercise level emerged in Week 1 of the online weight loss program, despite the fact that the recommendation to exercise was only mentioned in Week 4. In contrast, the female participants in the IW condition did not exercise more than their control condition counterparts. These findings are consistent with the notion that men have a greater propensity to exercise than women, which might explain why provisioning the IW drove the males (but not the females) to increase exercise as part of their weight loss strategy, thereby contributing to greater success in weight loss among the males.

Second, in terms of weight stigmatization, women are more vulnerable to weight discrimination than men are (Puhl, Andreyeva, and Brownell 2008). While men do not experience notable weight discrimination until their BMI reaches 35, women feel discriminated at the much lower BMI level of 27 (Puhl et al. 2008). Importantly, perceived weight stigmatization leads overweight individuals to avoid exercise in public (Vartanian and Shaprow 2008; Seacat and Mickelson 2009). As such, even though the exercise recommended by our weight loss program—brisk walking—involves a low level of physical exertion, females might still be uninterested in doing it, not just because of their low exercise propensity as discussed above, but also because they want to avoid being negatively judged by others in public.

Third, social conditioning that emphasizes the importance of physical appearance is greater for females than for males (Feingold 1990, 1992). Since other people's judgments of one's body size—along with the implications on romantic relationships, popularity, and employment—are considerations that drive intentions to lose weight, females are far more likely than males to have deliberated on these considerations. Therefore, the \$150 reward—being one reason to lose weight in this study—would be compared down (relative to other motivations) and would appear to be less compelling to females than to males. Hence, the \$150 in the IW condition might not mobilize females to lose weight.

Potential Limitations

We conclude with two limitations. First, we provided a cash incentive of \$150 for achieving 5% weight loss. It is unclear if one would obtain the same pattern of results with an incentive of a substantially different monetary value. Second, we had only one post-intervention weigh-in (12 weeks after the end of the program) and are thus unable to conclude if the greater weight loss among males in the IW condition persists beyond the 24 week timeframe.

APPENDIX A

ADDITIONAL DETAILS ON METHODOLOGY

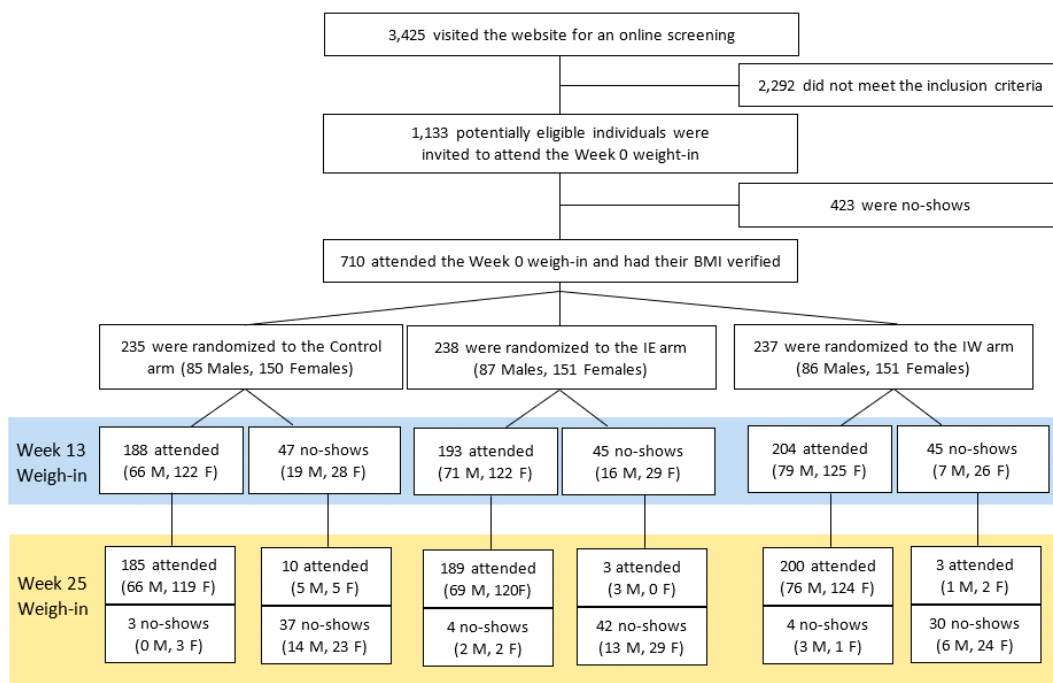
Selection Criteria

Participants had to meet the following eligibility criteria: age between 40 and 60; BMI between 23 and 33 kg/m²; not pregnant or planning to get pregnant; free from chronic diseases that require medical attention, including diabetes, cardiovascular disease, high blood pressure, and lung disease. A BMI of 23 kg/m² was chosen as Asian individuals above this BMI are considered to be overweight and at a greater risk for cardiometabolic complications (WHO Expert Consultation, 2004). The BMI upper limit of 33 kg/m² was chosen to minimize the influence of outliers on the main result of weight loss, as less than 5% of the Singapore population has a BMI of 34 kg/m² and higher according to a National Health Survey conducted by the Singapore Ministry of Health in 2004. The exclusion criteria concerning chronic diseases were imposed to ensure that participants were medically suited to undertake a self-administered weight-loss program. All selection criteria were set before subject recruitment commenced.

Attrition Analysis and Participant Characteristics

Among the males, the Week 13 attrition rates for the Control, IE, and IW conditions were 22.35%, 18.39%, and 8.14%, respectively. Among the females, the Week 13 attrition rates for these conditions were 18.67%, 19.21%, and 17.22%, respectively. Pairwise comparisons conducted based on a binary logistic regression using Week 13 attrition as the dependent variable showed that for males, the attrition rate was lower in the IW condition than in both the Control ($\chi^2(1) = 6.92$, 95% CI: [3.64, 24.79], $p < .01$) and IE conditions ($\chi^2(1) = 4.06$, 95% CI: [0.27, 20.23], $p = .04$). For females, there were no such differences across the three conditions. At Week 25, the overall attrition rate was 16.9%; there were no differences across the 3 conditions, for both males and females. The diagram in figure 6 shows the flow of participants through the RCT. To handle attrition, we report results using the ITT approach by assuming that the participants who did not show up at the weigh-ins remained at their baseline weights.

Figure 6 Diagram of Participant Flow



See table 4 for baseline measures. All baseline measures were gathered at the Week 0 weigh-in (i.e., prior to randomization). Statistical tests were conducted separately for each baseline measure. In terms of age, BMI, and education levels, males had a slightly higher BMI than females (main effect of gender: $F(1, 704) = 4.10, p = .04$), were on average one year younger than females (main effect of gender: $F(1, 704) = 7.69, p = .006$), and had higher levels of education (main effect of gender: $\chi^2(1, N = 710) = 10.23, p = .001$). Within each gender, there were no differences in age, baseline BMI, and educational levels across the three conditions.

We also asked the participants to predict their weight in three months and the number of video sessions they would view. There were no differences in expected weight loss across conditions for each gender. In terms of video viewing, males in the IE condition expected to view about one session less than males in the Control (diff = 1.32, 95% CI: [0.58, 2.06]), males in the IW (diff = 0.86, 95% CI: [0.12, 1.59]), and females in the IE (diff = 0.91, 95% CI: [0.26, 1.57]) conditions. There were no other differences in baseline measures.

TABLE 4 Baseline measures by gender and treatment condition

	MALES				FEMALES			
	Control (n = 85)	IE (n = 87)	IW (n = 86)	Overall (n = 258)	Control (n = 150)	IE (n = 151)	IW (n = 151)	Overall (n = 452)
Age, Mean (SD)	48.3 (4.9)	47.6 (5.7)	47.3 (4.7)	47.7 (5.1)	49.2 (5.5)	48.9 (5.6)	48.6 (5.3)	48.9 (5.5)
BMI, Mean (SD)	27.1 (2.7)	26.9 (2.7)	27.0 (2.3)	27.0 (2.6)	26.7 (2.6)	26.6 (2.5)	26.5 (2.4)	26.6 (2.5)
Weight in kg, mean (SD)	79.9 (9.3)	78.8 (10.1)	78.1 (9.9)	78.9 (9.8)	66.4 (7.9)	66.3 (7.6)	66.7 (8.1)	66.5 (7.9)
Education								
Lower secondary	0.0%	2.3%	4.7%	2.3%	0.7%	0%	1.3%	0.7%
Secondary	11.8%	10.3%	10.5%	10.9%	21.3%	25.2%	22.5%	23.0%
Post-secondary	27.1%	21.8%	20.9%	23.3%	39.3%	33.1%	35.8%	36.1%
University	61.2%	65.5%	64.0%	63.6%	38.7%	41.7%	40.4%	40.3%
Race								
Chinese	83.5%	82.8%	83.7%	83.3%	88.0%	87.4%	87.4%	87.6%
Asian ¹	14.1%	14.9%	11.6%	13.6%	10.7%	11.3%	11.3%	11.1%
Other	2.4%	2.3%	4.7%	3.1%	1.3%	1.3%	1.3%	1.3%
Expected weight loss, mean (SD) ²	4.5 (2.9)	4.3 (2.3)	4.9 (2.9)	4.6 (2.7)	5.1 (3.6)	4.5 (2.4)	4.7 (3.3)	4.8 (3.1)
Expected viewing ³	11.0 (1.6)	9.7 (3.5)	10.6 (2.2)	10.4 (2.6)	10.5 (2.3)	10.6 (2.5)	10.6 (2.4)	10.6 (2.4)

Note.

¹ Includes Indian, Malay, Indonesian, Filipino, Thai, Arab, Sri Lankan, and Burmese.

² Participants were asked to predict their weight (in kg) in 3 months, upon program completion. Expected weight loss in % of initial weight was computed as weight loss divided by baseline weight.

³ Participants were asked to predict the number of weekly sessions (out of the total of 12) they would complete.

REFERENCES

- Ananthapavan, Jaithri, A. Peeterson, and Gary Sacks (2017), "Paying People to Lose Weight: The Effectiveness of Financial Incentives Provided by Health Insurers for The Prevention and Management of Overweight and Obesity—A Systematic Review," *Obesity Reviews*.
- Bandura, Albert (1986), *Social Foundations of Thought and Action: A Social Cognitive Theory*, Englewood Cliffs, US: Prentice-Hall.
- Bandura, Albert (2004), "Health Promotion by Social Cognitive Means," *Health Education & Behavior*, 31(2), 143-64.
- Carmona, Richard H (2005), "Improving Health Literacy: Preventing Obesity with Education," *Journal of the American Dietetic Association*, 105(5), 9-10.
- Carpenter, Eugenia S (1980), "Children's Health Care and the Changing Role of Women," *Medical Care*, 1208-18.
- Cawley, John and Joshua A. Price (2013), "A case study of a workplace wellness program that offers financial incentives for weight loss," *Journal of Health Economics*, 32(5), 794-803.
- Central Intelligence Agency (CIA) (2018) Data from "The World Factbook."
- Consensus Statement of the Health Enhancement Research Organization (2012), "Guidance for A Reasonably Designed, Employer-Sponsored Wellness Program Using Outcomes-Based Incentives," *Journal of Occupational and Environmental Medicine*, 54(7), 889-96.
- Cooper, A. R., A. Page, K. R. Fox, and J. Misson (2000), "Physical Activity Patterns in Normal, Overweight and Obese Individuals Using Minute-By-Minute Accelerometry," *European Journal of Clinical Nutrition*, 54(12), 887.
- Crosnon, Rachel and Uri Gneezy (2009), "Gender Differences in Preferences," *Journal of Economic Literature*, 47(2), 448-74.
- Davis, Mark G., Kenneth R. Fox, Melvin Hillsdon, Debbie J. Sharp, Jo C. Coulson, Janice L. Thompson (2011), "Objectively Measured Physical Activity in a Diverse Sample of Older Urban UK Adults," *Medicine and Science in Sports and Exercise*, 43(4), 647-54.
- Deaner, Robert O., David C. Geary, David A. Puts, Sandra A. Ham, Judy Kruger, Elizabeth Fles, Bo Winegard, and Terry Grandis (2012), "A sex difference in the predisposition for physical competition: Males play sports much more than females even in the contemporary U.S. *PLoS ONE*, 7, e49168.
- Deaner, Robert O., Shea M. Balish, and Michael P. Lombardo (2016), "Sex differences in sports interest and motivation: An evolutionary perspective." *Evolutionary Behavioral Sciences*, 10(2): 73-97.
- Diabetes Prevention Program (DPP) Research Group (2002), "The Diabetes Prevention Program (DPP): description of lifestyle intervention," *Diabetes Care*, 25(12), 2165-71.
- Eckel, Catherine C. and Philip J. Grossman (2008), "Men, Women and Risk Aversion: Experimental Evidence," *Handbook of Experimental Economics Results*, 1, 1061-73.
- Feingold, Alan (1990), "Gender Differences in Effects of Physical Attractiveness on Romantic Attraction: A Comparison Across Five Research Paradigms," *Journal of Personality and Social Psychology*, 59(5), 981.

- Feingold, Alan (1992), "Gender Differences in Mate Selection Preferences: A Test of The Parental Investment Model," *Psychological Bulletin*, 112(1), 125.
- Feingold, Alan and Ronald Mazzella (1998), "Gender Differences in Body Image Are Increasing," *Psychological Science*, 9(3), 190-95.
- Finkelstein, Eric A., Laura A. Linnan, Deborah F. Tate, and Ben E. Birken (2007), "A Pilot Study Testing the Effect of Different Levels of Financial Incentives on Weight Loss Among Overweight Employees," *Journal of Occupational and Environmental Medicine*, 49(9), 981-89.
- Finkelstein, Eric A., Kwang-Wei Tham, Benjamin A. Haaland, and Aarti Sahasranaman (2017), "Applying Economic Incentives to Increase Effectiveness of An Outpatient Weight Loss Program (TRIO)—A Randomized Controlled Trial," *Social Science & Medicine*, 185, 63-70.
- Gold, Beth Casey, Susan Burke, Stephen Pintauro, Paul Buzzell, and Jean Harvey-Berino (2007), "Weight Loss on The Web: A Pilot Study Comparing a Structured Behavioral Intervention to A Commercial Program," *Obesity*, 15(1), 155.
- Hawkes, Jill M. and Karyn Holm (1993), "Gender Differences in Exercise Determinants," *Nursing Research*, 42(3), 166-72.
- Jago, Russell, Cheryl B. Anderson, Tom Baranowski, and Kathy Watson (2005), "Adolescent Patterns of Physical Activity: Differences by Gender, Day, and Time of Day," *American Journal of Preventive Medicine*, 28(5), 447-52.
- John, Leslie K., George Loewenstein, Andrea B. Troxel, Laurie Norton, Jennifer E. Fassbender, and Kevin G. Volpp (2011), "Financial Incentives for Extended Weight Loss: A Randomized, Controlled Trial," *Journal of General Internal Medicine*, 26(6), 621-26.
- Kelly, Tanika, Wenjie Yang, Chung-Shiuan Chen, Kristi Reynolds, and Jiang He (2008), "Global Burden of Obesity in 2005 and Projections to 2030," *International Journal Of Obesity*, 32(9), 1431.
- Kullgren, Jeffrey T., Andrea B. Troxel, George Loewenstein, David A. Asch, Laurie A. Norton, Lisa Wesby, Yuanyuan Tao, Jingsan Zhu, and Kevin G. Volpp (2013), "Individual-Versus Group-Based Financial Incentives for Weight Loss: A Randomized, Controlled Trial," *Annals of Internal Medicine*, 158(7), 505-14.
- Lin, Yen-Chun, Mei Chang Yeh, Ya-Mei Chen, and Lian-Hua Huang (2010), "Physical Activity Status and Gender Differences in Community-Dwelling Older Adults with Chronic Diseases," *Journal of Nursing Research*, 18(2), 88-97.
- Loewenstein, George and Drazen Prelec (1992), "Anomalies in Intertemporal Choice: Evidence and An Interpretation," *The Quarterly Journal of Economics*, 107(2), 573-97.
- Misra-Hebert, Anita D., Bo Hu, Glen Taksler, Robert Zimmerman, and Michael B. Rothberg (2016), "Financial Incentives and Diabetes Disease Control in Employees: A Retrospective Cohort Analysis," *Journal of General Internal Medicine*, 31(8), 871-77.

- Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., and Abraham, J. P. (2014), "Global, Regional, And National Prevalence of Overweight and Obesity in Children and Adults During 1980–2013: A Systematic Analysis for The Global Burden of Disease Study 2013," *The Lancet*, 384(9945), 766-81.
- Pate, Russell R., Karin A. Pfeiffer, Stewart G. Trost, Paula Ziegler, and Marsha Dowda (2004), "Physical Activity Among Children Attending Preschools," *Pediatrics*, 114(5), 1258-63.
- Puhl, Rebecca M., Tatiana Andreyeva, and Kelly D. Brownell (2008), "Perceptions of Weight Discrimination: Prevalence and Comparison to Race and Gender Discrimination in America," *International Journal of Obesity*, 32(6), 992.
- Relton, Clare, Mark Strong, and Jessica Li (2011), "The 'Pounds for Pounds' Weight Loss Financial Incentive Scheme: An Evaluation of a Pilot in NHS Eastern and Coastal Kent," *Journal of Public Health*, 33(4): 536-42.
- Rhoads, Steven E. and Christopher H. Rhoads (2012), "Gender Roles and Infant/Toddler Care: Male and Female Professors on The Tenure Track," *Journal of Social, Evolutionary, and Cultural Psychology*, 6(1), 13.
- Saccardo, Silvia, Aniela Pietrasz, and Uri Gneezy (2017), "On the Size of The Gender Difference in Competitiveness," *Management Science*.
- Seacat, Jason D. and Kristin D. Mickelson (2009), "Stereotype Threat and The Exercise/Dietary Health Intentions of Overweight Women," *Journal of Health Psychology*, 14(4), 556-67.
- Stevens, Gretchen A., Gitanjali M. Singh, Yuan Lu, Goodarz Danaei, John K. Lin, Mariel M. Finucane, Adil N. Bahalim et al (2012), "National, Regional, And Global Trends in Adult Overweight and Obesity Prevalences," *Population Health Metrics*, 10(1), 22.
- Tate, Deborah F., Rena R. Wing, and Richard A. Winett (2001), "Using Internet Technology to Deliver a Behavioral Weight Loss Program," *Journal of the American Medical Association*, 285(9), 1172-77.
- Thaler, R. and G. Loewenstein (1992), "Intertemporal Choice," in *The Winners Curse: Paradoxes & Anomalies of Economic Life*, ed. R. Thaler, New York, Free Press, 92-106.
- Volpp, Kevin G., Leslie K. John, Andrea B. Troxel, Laurie Norton, Jennifer Fassbender, and George Loewenstein (2008), "Financial Incentive–Based Approaches for Weight Loss: A Randomized Trial," *Journal of the American Medical Association*, 300(22), 2631-37.
- Vartanian, Lenny R. and Jacqueline G. Shaprow (2008), "Effects of Weight Stigma on Exercise Motivation and Behavior: A Preliminary Investigation among College-Aged Females," *Journal of Health Psychology*, 13(1), 131-38.
- WHO, Expert Consultation (2004), "Appropriate Body-Mass Index for Asian Populations and Its Implications for Policy and Intervention Strategies," *Lancet*, 363(9403), 157.

WEB-APPENDIX

Online Weight Loss Program: Incentivize Weight Loss or Health Education?

This web appendix reports the following statistical analysis results as robustness checks for the results reported in the main paper:

- (1) Means of weight loss at Week 13 and Week 25 using the ITT and per-protocol approaches, with and without regression adjustment (Table A1)
- (2) Full regression results using the ITT and per-protocol approaches, with and without regression adjustment (Table A2)
- (3) Full regression results using heteroscedasticity-consistent standard errors (Table A3)

TABLE A1
 PERCENTAGE WEIGHT LOSS AT WEEK 13 AND WEEK 25 BY GENDER AND TREATMENT CONDITION

	Males			Females		
	Control	IE	IW	Control	IE	IW
Weight loss at Week 13						
Full Sample						
Mean, Unadjusted	0.87	0.87	2.40*** c γ	1.44	1.27	1.03
95% CI	0.31 – 1.43	0.32 – 1.43	1.84 – 2.96	1.02 – 1.87	0.85 – 1.69	0.61 – 1.45
Per-Protocol Subsample						
Mean, Unadjusted	1.12	1.00	2.57*** c γ	1.83	1.53	1.26
95% CI	0.44 – 1.81	0.33 – 1.68	1.93 – 3.21	1.32 – 2.35	1.02 – 2.04	0.76 – 1.76
Full Sample						
Mean, Regression-adjusted [#]	0.82	0.83	2.39*** c γ	1.46	1.29	1.06
95% CI	0.26 – 1.39	0.28 – 1.39	1.83 – 2.95	1.03 – 1.88	0.87 – 1.71	0.64 – 1.48
Per-Protocol Subsample						
Mean, Regression-adjusted [#]	1.08	0.94	2.56*** c γ	1.84	1.57	1.28
95% CI	0.38 – 1.77	0.27 – 1.62	1.92 – 3.20	1.33 – 2.36	1.06 – 2.08	0.78 – 1.78
Weight loss at Week 25						
Full Sample						
Mean, Unadjusted	0.88	1.36	2.53*** a	1.55	1.46	1.66
95% CI	0.16 – 1.61	0.64 – 2.08	1.80 – 3.25	1.01 – 2.10	0.91 – 2.00	1.11 – 2.20
Per-Protocol Subsample						
Mean, Unadjusted	1.14	1.69	2.82**	2.14	1.84	2.00
95% CI	0.27 – 2.02	0.83 – 2.55	2.01 – 3.64	1.49 – 2.79	1.19 – 2.49	1.36 – 2.64
Full Sample						
Mean, Regression-adjusted [#]	0.83	1.34	2.55*** a	1.54	1.47	1.69
95% CI	0.11 – 1.56	0.63 – 2.06	1.83 – 3.27	1.00 – 2.09	0.92 – 2.01	1.14 – 2.23
Per-Protocol Subsample						
Mean, Regression-adjusted [#]	1.10	1.67	2.87*** a	2.12	1.85	2.01
95% CI	0.22 – 1.99	0.81 – 2.53	2.04 – 3.69	1.47 – 2.78	1.20 – 2.50	1.37 – 2.65

Note. The full sample consists of all 710 participants. The per-protocol subsample consists of the 574 participants who attended all the weigh-ins.

[#] These are regression-adjusted estimation of means, with baseline BMI, age, and education indicators added to the regression model as covariates.

*, **, *** indicate $p \leq .05$, $p \leq .01$, and $p \leq .005$, respectively, for comparison with the same-sex Control condition.

^{a, b, c} indicate $p \leq .05$, $p \leq .01$, and $p \leq .005$, respectively, for comparison with the same-sex IE condition.

^{α, β, γ} indicate $p \leq .05$, $p \leq .01$, and $p \leq .005$, respectively, for comparison with the same treatment condition in the opposite gender.

TABLE A2.
FULL REGRESSION RESULTS OF TREATMENT AND GENDER EFFECTS ON
PERCENTAGE WEIGHT LOSS AT WEEK 13 AND WEEK 25

WEEK 13				
	Model			
Predictor variables	(1) Full Sample (<i>n</i> = 710)	(2) Full Sample (<i>n</i> = 710)	(3) Per-Protocol Sample (<i>n</i> = 574)	(4) Per-Protocol Sample (<i>n</i> = 574)
Constant	-0.326 (1.668)	-1.443*** (0.215)	-0.817 (1.952)	-1.834*** (0.261)
Control (<i>Reference</i>)	—	—	—	—
IE	0.169 (0.303)	0.171 (0.304)	0.274 (0.367)	0.301 (0.368)
IW	0.396 (0.303)	0.409 (0.304)	0.561 (0.364)	0.571 (0.365)
Female (<i>Reference</i>)	—	—	—	—
Male	0.632 (0.359)	0.571 (0.358)	0.767 (0.440)	0.712 (0.436)
Male × IE	-0.176 (0.502)	-0.174 (0.504)	-0.143 (0.611)	-0.183 (0.612)
Male × IW	-1.961*** (0.503)	-1.938*** (0.505)	-2.047*** (0.601)	-2.016*** (0.601)
Baseline BMI	.048 (0.040)	—	0.046 (0.048)	—
Age	-0.028 (0.019)	—	-0.024 (0.023)	—
Education	-0.323** (0.129)	—	-0.332* (0.159)	—
WEEK 25				
	Model			
Predictor variables	(1) Full Sample (<i>n</i> = 710)	(2) Full Sample (<i>n</i> = 710)	(3) Per-Protocol Sample (<i>n</i> = 574)	(4) Per-Protocol Sample (<i>n</i> = 574)
Constant	1.503 (2.155)	-1.552*** (0.278)	0.666 (2.493)	-2.141*** (0.332)
Control (<i>Reference</i>)	—	—	—	—
IE	0.078 (0.391)	0.093 (0.393)	0.271 (0.469)	0.306 (0.469)
IW	-0.144 (0.391)	-0.107 (0.393)	0.109 (0.465)	0.139 (0.465)
Female (<i>Reference</i>)	—	—	—	—
Male	0.712 (0.464)	0.669 (0.463)	1.021 (0.562)	1.000 (0.556)
Male × IE	-0.59 (0.648)	-0.573 (0.652)	-0.840 (0.780)	-0.853 (0.781)
Male × IW	-1.574* (0.650)	-1.537* (0.653)	-1.873* (0.767)	-1.821* (0.767)
Baseline BMI	.049 (0.051)	—	0.023 (0.062)	—
Age	-0.065** (0.025)	—	-0.052 (0.030)	—
Education	-0.371* (0.167)	—	-0.256 (0.203)	—

Note. This table presents the ordinary least squares regression coefficient estimates, with standard errors in parentheses. Dependent variables are Week 13 weight change [(Week 13 weight – Week 0 weight) / Week 0 weight]*100 and Week 25 weight change [(Week 13 weight – Week 0 weight) / Week 0 weight]*100, respectively. Control and Female are used as the reference categories. Regression models (1) and (2) are constructed using the ITT approach, whereas models (3) and (4) are constructed using the per-protocol approach.

*, **, *** indicate statistical significance at the 5%, 1%, and 0.5% levels, respectively.

TABLE A3
 REPLICATION OF REGRESSION RESULTS USING A HETEROSCEDASTICITY CONSISTENT STANDARD ERROR

WEEK 13				
	Model			
Predictor variables	(1) Full Sample ($n = 710$)	(2) Full Sample ($n = 710$)	(3) Per-Protocol Sample ($n = 574$)	(4) Per-Protocol Sample ($n = 574$)
Constant	-0.326 (1.834)	-1.443*** (0.201)	-0.817 (2.113)	-1.834*** (0.240)
Control (<i>Reference</i>)	—	—	—	—
IE	0.169 (0.291)	0.171 (0.291)	0.274 (0.352)	0.301 (0.350)
IW	0.396 (0.287)	0.409 (0.293)	0.561 (0.346)	0.571 (0.351)
Female (<i>Reference</i>)	—	—	—	—
Male	0.632 (0.355)	0.571 (0.361)	0.767 (0.447)	0.712 (0.450)
Male \times IE	-0.176 (0.493)	-0.174 (0.497)	-0.143 (0.613)	-0.183 (0.615)
Male \times IW	-1.961*** (0.534)	-1.938*** (0.536)	-2.047*** (0.632)	-2.016*** (0.634)
Baseline BMI	.048 (0.040)	—	0.046 (0.050)	—
Age	-0.028 (0.020)	—	-0.024 (0.024)	—
Education	-0.323* (0.132)	—	-0.332* (0.164)	—
WEEK 25				
	Model			
Predictor variables	(1) Full Sample ($n = 710$)	(2) Full Sample ($n = 710$)	(3) Per-Protocol Sample ($n = 574$)	(4) Per-Protocol Sample ($n = 574$)
Constant	1.503 (2.407)	-1.552*** (0.289)	0.666 (2.767)	-2.141*** (0.320)
Control (<i>Reference</i>)	—	—	—	—
IE	0.078 (0.399)	0.093 (0.400)	0.271 (0.467)	0.306 (0.467)
IW	-0.144 (0.387)	-0.107 (0.396)	0.109 (0.447)	0.139 (0.453)
Female (<i>Reference</i>)	—	—	—	—
Male	0.712 (0.438)	0.669 (0.445)	1.021 (0.530)	1.000 (0.535)
Male \times IE	-0.59 (0.632)	-0.573 (0.634)	-0.840 (0.776)	-0.853 (0.773)
Male \times IW	-1.574* (0.660)	-1.537* (0.663)	-1.873* (0.772)	-1.821* (0.772)
Baseline BMI	.049 (0.053)	—	0.023 (0.064)	—
Age	-0.065* (0.027)	—	-0.052 (0.031)	—
Education	-0.371* (0.171)	—	-0.256 (0.215)	—

This table presents the ordinary least squares regression coefficient estimates, with heteroscedasticity consistent standard errors in parentheses. As in Table A2, dependent variables are Week 13 weight change and Week 25 weight change, respectively. Control and Female are used as the reference categories. Regression models (1) and (2) are constructed using the ITT approach based on the full sample, whereas models (3) and (4) are constructed using the per-protocol approach based on the 574 participants who attended all the 3 weigh-ins.

*, **, *** indicate statistical significance at the 5%, 1%, and 0.5% levels, respectively.