Teaching temptation bundling to boost exercise: A field experiment[☆]

Erika L. Kirgios [a,*], Graelin H. Mandel [b], Yeji Park [c], Katherine L. Milkman [a], Dena M. Gromet [c], Joseph S. Kay [c], Angela L. Duckworth [a,d]

[a] Department of Operations, Information, & Decisions, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, United States
[b] The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, United States
[c] Behavior Change for Good Initiative, University of Pennsylvania, Philadelphia, PA 19104, United States
[d] Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, United States

ARTICLE INFO

ABSTRACT

Temptation bundling—pairing a pleasurable indulgence with a behavior that provides delayed rewards—combats present bias by making behaviors with delayed benefits more instantly-gratifying. If people are sophisticated and capable of following self-set rules to overcome present bias, they could benefit from learning about temptation bundling. Participants in a four-week exercise-boosting program ($N = 6792$) received either an audiobook with encouragement to temptation bundle, only an audiobook, or neither an audiobook nor encouragement to temptation bundle. Giving participants audiobooks and encouraging temptation bundling boosted their likelihood of a weekly workout by 10–14% and average weekly workouts by 10–12% during and up to seventeen weeks post-intervention. Relative to giving audiobooks alone, encouraging temptation bundling had a modest positive effect on exercise on the extensive margin. The marginal benefit of encouraging temptation bundling may be small because free audiobooks leak information: Simply providing an audiobook to exercise program participants suggests they should temptation bundle.

1. Introduction

A dieter is offered chocolate cake at a party. Someone trying to save more money stares longingly at a fancy new cellphone. A student trying to improve her grades considers skipping her studies to watch a binge-worthy TV show instead. These are all examples of *want-should* conflicts (Bazerman, Tenbrunsel, & Wade-Benzoni, 1998; Milkman et al., 2008), in which people are forced to choose between doing what they *want* to get instant gratification and what they *should* to gain long-term benefits. *Wants* are defined as actions that provide short-term pleasure but no long-term benefits, while *shoulds* provide long-term benefits but often involve short-term pain. Due to present bias, defined as the tendency to dramatically overweight immediate rewards relative to delayed rewards, people often choose *wants* over *shoulds* in the heat of the moment, only to later regret their decisions (Milkman et al., 2008).

Unfortunately, evidence shows that over time, repeatedly selecting

wants over *shoulds* can take a toll on people's health as well as their financial, academic, and professional outcomes (Boals, Vandellen, & Banks, 2011; Kaur, Kremer, & Mullainathan, 2010; Stavrova & Kokkoris, 2017; Strömbäck, Lind, Skagerlund, Västfjäll, & Tinghög, 2017). For instance, an estimated 40% of premature deaths can be attributed to repeated decisions that favor *wants* over *shoulds* when it comes to eating, drinking, smoking, exercise, sex, and vehicle safety (Schroeder, 2007).

In light of the large social and personal costs of consistently choosing *wants* in the face of *want-should* conflicts, policymakers and behavioral scientists have tested a wide range of interventions designed to encourage people to overcome their present bias and make more *should* choices (see Duckworth, Milkman, & Laibson, 2018, for a review). Many such interventions rely on people's "sophistication" or self-awareness (O'Donoghue & Rabin, 2011), leaning on their ability to anticipate future temptations and their willingness to adopt strategies that help them resist those temptations. Such interventions have encouraged

---

people to pre-commit to *should* choices such as saving money (Thaler & Benartzi, 2004), to self-impose penalties if they fail to engage in *should* activities such as exercise (Royer, Stehr, & Sydnor, 2015), or to avoid situations that might make it particularly tempting to choose *want* activities such as screen time over *should* activities such as studying (Duckworth, White, Matteucci, Shearer, & Gross, 2016).

These kinds of interventions prioritize helping people engage in *should* activities at the expense of *wants*. But choosing *wants* versus *shoulds* need not always be mutually exclusive.

In the present research, we test the effectiveness of harnessing the appeal of *wants* to increase *should* choices. We teach participants to pair *wants* with *shoulds* to increase their follow-through on *should* behaviors, and we examine whether merely providing them with *wants* in the company of *shoulds* is enough to suggest such pairings. First, we examine whether encouraging people to pair indulgent entertainment with exercise can durably increase gym attendance using a technique called "temptation bundling." In doing so, we shed new light on the degree to which people can effectively follow self-set, personal rules to overcome present bias (Ainslie, 2001; Benabou & Tirole, 2004; Grady, Goodenow, & Borkin, 1988; Thaler, 2000; Wertenbroch, 1998). Second, we explore whether giving people free entertainment as part of a program to encourage exercise can implicitly suggest temptation bundling and therefore boost exercise. Consequently, we illuminate people's capacity to make complex inferences from information policymakers leak (Tannenbaum, Valesek, Knowles, & Ditto, 2013; McKenzie, Liersch, & Finkelstein, 2006; McKenzie & Nelson, 2003). In both cases, we test the extent of individuals' sophistication in overcoming their present bias, adding to a growing literature that suggests people are often aware of present bias and capable of taking steps to reduce its harms (O'Donoghue and Rabin, 2011).

As conceived by Milkman, Minson, and Volpp (2014), temptation bundling pairs *wants* (e.g., getting a pedicure, watching lowbrow TV, eating chocolate) with *shoulds* (e.g., paying taxes, exercising, getting an eye exam). This pairing makes *should* activities more enticing and therefore more likely to be readily executed; it also makes *want* activities less wasteful and guilt-inducing. In our experiments, exercise constitutes the *should* activity and consuming lowbrow entertainment (in the form of television, podcasts, and audiobooks) constitutes the *want* activity.[1]

In the single, small ($N = 226$) prior experiment exploring the benefits of temptation bundling, Milkman et al. (2014) found initial evidence for the short-term effectiveness of this behavior change strategy. However, because the literature on temptation bundling includes just this small field study, a great deal remains unknown about the mechanisms responsible for its impact, whether it can have durable benefits, its general robustness, and whether people can self-generate the idea to temptation bundle given the right stimuli, even if they aren't explicitly taught how it works.

This paper presents a field experiment roughly 15 times larger than the prior study, which tests the benefits of merely encouraging temptation bundling rather than externally imposing temptation bundling. With a seventeen-week follow-up period, we provide the first evidence that encouraging temptation bundling and providing a bundle-worthy *want* has durable benefits, significantly boosting both the likelihood and frequency of weekly gym visits for months. Separately, we find a small benefit of merely *encouraging* temptation bundling, but it accrues solely on the extensive margin rather than the intensive margin. Specifically, it leads marginally more people to visit the gym in a given week during our intervention and it boosts the likelihood of a gym visit significantly during our 10 and 17-week post-intervention follow-up

periods. However, encouragement to temptation bundle alone does not robustly cause people to make *more* gym visits per week. Our field experiment and two follow-up experiments also suggest that people are sophisticated enough to infer they should temptation bundle from programming that "leaks" the idea—specifically, an exercise program that provides the gift of a tempting *want* in the form of an audiobook.

1.1. Can people self-impose constraints to overcome their present bias?

Research shows that some people are well aware they're likely to procrastinate on *should* activities that can be painful in the present but beneficial in the future in favor of *want* activities that are pleasurable now but carry few, if any, long-term benefits. Those who recognize their self-control struggles and are interested in finding ways to prevent self-control failures have been dubbed "sophisticates" (O'Donoghue & Rabin, 1999; O'Donoghue and Rabin, 2011). Past research on sophisticates has largely focused on documenting their taste for commitment devices—technologies that help them avoid future temptations, such as smaller plates, bank accounts that are inaccessible until a predetermined future date, or fines that arise when they fail to achieve their goals (Ariely & Wertenbroch, 2002; Ashraf, Karlan, & Yin, 2006; Bryan, Karlan, & Nelson, 2010; Royer, Stehr, & Sydnor, 2015; Laibson, 1997; Wansink & Cheney, 2005). Anticipating that overcoming present bias will be difficult, sophisticates may also be on the lookout for behavioral tricks or techniques they can try on their own.

Indeed, a growing literature suggests some sophisticates can successfully self-impose behavior change strategies. Teaching people new skills or practical tricks can empower them to make better choices on their own (Drexler, Fischer, & Schoar, 2014; Sedlmeier & Gigerenzer, 2001; for a review of "boosts," see Hertwig & Grüne-Yanoff, 2017). Evidence suggests people can also craft and successfully follow personal guidelines that help them exert self-control (Ainslie, 1992; Ainslie, 2001; Benabou & Tirole, 2004; Cheema & Soman, 2006; Grady et al., 1988; Thaler, 2000; Wertenbroch, 1998). For instance, people can self-impose constraints on purchases of vice items such as cigarettes (Wertenbroch, 1998), assign activities to categories and restrict spending to implicit category-level budgets on their own initiative (Thaler, 2000), and reward themselves for the successful completion of *should* activities to encourage their own goal-oriented behaviors (Grady et al., 1988).

These findings indicate that individuals can be sophisticated and agentic in their pursuit of behavior change and thus might be able and eager to self-impose useful new rules that could help them overcome present bias.

1.2. Temptation bundling

Temptation bundling is one form of personal rule that past research suggests may appeal to sophisticates hoping to reign in *want* behaviors and pursue more *should* behaviors. Sophisticates may like the idea of only allowing themselves to enjoy *wants*, like lowbrow audio-novels or other forms of entertainment, while simultaneously engaging in a *should* behavior, like exercise. One small prior study showed that when access to a *want* (a tempting audio novel) was denied to people unless they engaged in a *should* behavior (visiting the gym), they exercised more often (Milkman et al., 2014). In this same study, a survey determined that the majority of participants would be interested in paying for gym-only access to *wants* because they thought this restriction might help them boost their physical activity.

Although it was not a direct test of temptation bundling (because *wants* could be accessed anytime, not only in the presence of *shoulds*), a separate field experiment found that encouraging high school students to enjoy snacks, music, and colorful markers while completing challenging math problems led to greater persistence than encouraging this *should* behavior without any *wants* incorporated (Woolley & Fishbach, 2016). Gamification interventions that add features such as points and leaderboards to the pursuit of *should* goals in order to make them more

---

[1] Lowbrow entertainment is instantly gratifying but generally lacks long-term benefits (see Milkman, Rogers and Bazerman, 2009). Meanwhile, exercise has long-term benefits but lacks immediate appeal: Exercising more was the most common New Year's resolution in 2020 (Ballard, 2020), but people frequently fail to follow through on this health goal (DellaVigna & Malmendier, 2006).

immediately enjoyable (and more like *wants*) have also successfully increased daily step counts (Patel et al., 2017; Patel et al., 2019; Zuckerman & Gal-Oz, 2014). Taken together, these studies suggest that combining *wants* with *shoulds* can promote *should* behaviors.

These studies all included third parties who bundled *wants* with *shoulds* for participants. We focus on whether it is useful to merely suggest that people *self-impose* such combinations. Given that past research shows that 1) combining *wants* with *shoulds* can make *shoulds* more appealing and that 2) sophisticates can self-impose constraints to overcome present bias, there is reason to believe suggesting temptation bundling might be valuable. However, the only evidence testing this claim comes from an under-powered treatment in the original temptation bundling experiment conducted by Milkman et al. in 2014. In this study, some gym-goers ($N = 75$) received four tempting audiobooks of their choice (e.g., *The Hunger Games, The Da Vinci Code*) and were encouraged to listen to the audiobooks only while exercising at the gym. Compared to participants in a control condition ($N = 76$) who received an equally valued gift card to Barnes & Noble and no instructions about temptation bundling, those encouraged to temptation bundle initially visited the gym marginally more often. However, these marginal benefits quickly fell below detectable levels. Further, the study's design makes it impossible to disentangle the benefits of giving people audiobooks from the benefits of teaching them about temptation bundling.

In this paper, we test the hypothesis that merely teaching people about temptation bundling may have durable benefits. We also run the first well-powered test of the hypothesis that both teaching people about temptation bundling and providing them with free access to a *want* that can be used to temptation bundle has benefits. Formally, we hypothesize the following:

Hypothesis 1. People introduced to the idea of temptation bundling will engage in more of the *should* behavior they are seeking to boost (in this case, exercise).

Hypothesis 2. People introduced to the idea of temptation bundling and provided with a *want* (in this case, a free audiobook) to bundle with the *should* behavior they are seeking to boost (in this case, exercise) will engage in more of that *should* behavior.

1.3. Information leakage

We also investigate whether merely offering people access to bundle-ready *wants* when they are aiming to engage in more of a *should* behavior can facilitate behavior change by indirectly *suggesting* the idea of temptation bundling.

Past research on "information leakage" has shown that people are often able to correctly infer intentions of policymakers based on context clues, such that the context "leaks" information that influences people's judgments and subsequent decisions and behaviors. These context clues may include physical features of the environment such as objects or images that prime people to think competitively (Kay, Wheeler, Bargh, & Ross, 2004; Prinsen, de Ridder, & de Vet, 2013); policy structures such as punitive or rewarding health incentives (Tannenbaum et al., 2013); or choice or information frames such as which options are listed as the default (Hilton, 1995; McKenzie, Liersch, & Finkelstein, 2006; McKenzie & Nelson, 2003; Sher & McKenzie, 2006; Strack, Martin, & Schwarz, 1988). For example, when a default option is present in a choice set, individuals often infer that policymakers are implicitly recommending the default option (otherwise, why would it be the "standard" choice?) and as a result, they are more likely to stick to the default even when given the opportunity to opt out (McKenzie et al., 2006). Individuals also (correctly) infer the presence of negative attitudes towards those who are overweight in an organization based on whether its workplace wellness programs punish employees for being overweight or, instead, reward them for being at a weight that the programs deem healthy (Tannenbaum et al., 2013).

We examine whether information leakage can inspire people to

temptation bundle, such that individuals infer and apply a behavior change tactic from context clues even with no explicit direction or instruction (Ainslie, 1992; Ainslie, 2001; Thaler, 2000; Wertenbroch, 1998). Specifically, we propose that simply giving participants a free audiobook in the context of a program designed to build exercise habits may lead them to infer and apply a temptation bundling strategy to exercise more often. Formally, we hypothesize:

Hypothesis 3. Providing people with a *want* (in this case, a free audiobook) that can be easily bundled with a *should* behavior they are seeking to boost (in this case, exercise) leaks information that the *want* ought to be paired with the *should* to make the *should* more attractive.

Hypothesis 4. People who receive a *want* (in this case, a free audiobook) that can easily be bundled with a *should* behavior they are seeking to boost (in this case exercise) will engage in more of that *should* behavior.

1.4. Current investigation

We present the results of four experiments that test our hypotheses about whether people can self-impose or even infer and apply temptation bundling to boost exercise. Study 1A is a field experiment in which we test the effectiveness of giving people a free audiobook and encouraging a bundling strategy (rather than imposing one) against simply giving them a free audiobook (H1). We find that encouraging temptation bundling only leads to small benefits on the extensive rather than the intensive margin, boosting the likelihood of a weekly gym visit by a marginal 4% during the intervention and significantly lifting it by 6% during the ten weeks post-intervention. In Study 1B, we compare both conditions from Study 1A (teaching participants about temptation bundling and providing them with a free audiobook as well as simply giving them a free audiobook) with a control condition in which participants are not given an audiobook, allowing us to test H2 and H4. We find that teaching participants about temptation bundling and providing them with a *want* (i.e., a free audiobook) that can be linked with exercise increases the likelihood they'll visit the gym by 10–14% both during and up to seventeen weeks post-intervention. We further find that simply providing participants with a free audiobook in the context of an exercise program has no significant effect on the likelihood of a weekly gym visit during the intervention but boosts the likelihood of gym visits by 15% during the four weeks post-intervention. And in two follow-up laboratory experiments, Studies 2A and 2B, we find evidence suggesting that participants given a free audiobook by a gym program will infer they should use a temptation bundling technique to build exercise habits, consistent with H3. Overall, our findings contribute important new insights to scholarly debates regarding (a) people's degree of sophistication and capacity to self-impose rules to counter present bias and (b) their ability to make useful inferences about leaked information.

Our study focuses on boosting gym attendance for two key reasons. First, increasing physical activity is an important goal for many individuals as well as policymakers (Guthold, Stevens, Riley, & Bull, 2018). Fewer than half of American adults exercise sufficiently and 30% of American adults report engaging in no exercise whatsoever (Physical Activity Guidelines, 2018) . Recent estimates suggest that physical inactivity accounts for 9% of premature mortality and a 25% increase in worldwide physical activity levels would be enough to avert 1.3 million deaths each year (Lee et al., 2012). Thus, our study focuses on an important policy problem. Further, many past studies of behavior change have focused on increasing exercise (Acland & Levy, 2015; Babcock, Bedard, Charness, Hartman, & Royer, 2015; Charness & Gneezy, 2009; Condliffe, Isgin, & Fitzgerald, 2017; Royer et al., 2015; DellaVigna & Malmendier, 2006), including the original study of temptation bundling (Milkman et al., 2014). Testing the effectiveness of our interventions in a gym setting makes it easier to directly compare their impact with that of other behavioral interventions.

2. Study 1: A large-scale, preregistered field experiment

In Study 1, we conducted a large-scale field experiment to test Hypotheses 1, 2 and 4. We first report the results of Study 1A, a preregistered experiment in which we evaluate whether encouraging participants to temptation bundle and giving them something appealing to bundle with exercise (a code to download a free audiobook from Audible.com) boosts gym visits relative to simply giving participants the same temptation (a free audiobook) with no further explanation (H1). In Study 1B, we pool a quasi-experimental control condition with our Study 1A data. In this control condition, participants did not receive a free audiobook or encouragement to temptation bundle. Examining this control condition allows us to test whether encouraging temptation bundling and providing participants with a free audiobook increases exercise (H2) and whether simply providing participants with a free audiobook when they sign up for an exercise-boosting program at their gym can increase exercise (H4).

2.1. Study 1A: Does encouraging people to temptation bundle and giving them a free audiobook boost gym visits relative to simply giving them a free audiobook?

In Study 1A, we conducted a preregistered field experiment to test the effectiveness of encouraging people to only enjoy a temptation while working out at the gym.

Method. We partnered with 24 Hour Fitness, one of the largest gym chains in the United States, to conduct a field experiment. All active adult members of 24 Hour Fitness (over four million people) were eligible to participate. To access their gym, 24 Hour Fitness members had to check in by presenting an ID to staff at the front desk, swiping or scanning a member card, or using a fingerprint reader and unique check-in code. We used check-in data provided by 24 Hour Fitness to track participants' gym visits before, during, and after a four-week intervention period.

Participant recruitment. From March 21, 2018, through January 31, 2019, members of 24 Hour Fitness gyms were invited to join the "StepUp Program" through a 24 Hour Fitness marketing campaign that included emails, gym app notifications, social media advertisements, phone calls, postcards, and posters in gyms, as well as on-site, in-person recruitment. All recruitment materials advertised the StepUp Program as "a habit-building, science-based workout program" and informed gym members that they could sign up for free and earn Amazon cash rewards for participating. Members were also informed that simply registering for the StepUp Program would make them eligible for the chance to win a $50 Amazon gift card. Our study was one of 20 preregistered studies embedded in the StepUp Program (see Milkman et al., 2020, for a meta-analysis of all 54 study conditions embedded in these 20 studies).

Participant enrollment. Gym members signed up online for StepUp via the program's website. After consenting to participate in the research, participants provided their gym check-in code and date of birth to verify their gym membership and allow us to sync their online enrollment data with their gym visit data. Participants also provided their name, email address, and phone number upon enrollment. They received an SMS code to verify they could receive text messages from StepUp before being randomized to a study condition (see page 7 of the Online Supplement for the full participant enrollment process).

Experimental design. StepUp registrants were randomly assigned with equal probability to one of two experimental conditions that were designed and preregistered to evaluate the impact of teaching gym members about temptation bundling (http://bit.ly/2NUkEBE; this is referred to as the "Free Audiobook Experiment" in Fig. 2). As detailed below, the *free audiobook with encouragement to temptation bundle* condition introduced and explained the benefits of temptation bundling and provided participants with a code that would allow them to download a free audiobook from Audible.com. In contrast, the *simple free audiobook* condition only provided participants with this Audible.com download

code.

Experimental conditions. The initial registration process was the same for all participants regardless of experimental condition: After enrolling in the 28-day StepUp Program, participants were informed that they had earned 100 points for registering and would receive 300 StepUp points each time they visited the gym, redeemable for Amazon cash after the program's conclusion. Participants learned that points were redeemable at a rate of $5 for 7000 points ($0.22 per gym visit). They were then prompted to plan the dates and times when they would work out at the gym each week during the StepUp Program and learned that they would receive text message reminders to exercise 30 minutes prior to each scheduled workout for the program's 28-day duration.

Participants then answered a multiple-choice question about what they normally do at the gym ("attend fitness classes," "swim," "use cardio equipment," "lift weights," and/or "other"). Next, they were informed that as a special promotion, the StepUp Program had partnered with Audible.com to share one free audiobook with them, and at the end of registration, they would be provided with a code that would allow them to download their free audiobook.

Up to this point, participants in each of our two experimental conditions experienced identical registration processes. Their treatments diverged only after they learned they would receive a free audiobook code. Participants randomly assigned to the *simple free audiobook* condition ($N = 1604$) received their code for a free audiobook and were told their registration process was complete. Meanwhile, participants who were randomly assigned to the *free audiobook with encouragement to temptation bundle* condition ($N = 1685$) received additional content during registration. After learning that they would receive a free audiobook from Audible.com, they were told that "A secret to making working out at the gym into an enjoyable, fun habit is to make a simple rule: I only let myself enjoy my favorite TV shows, audiobooks, or podcasts on my smartphone when exercising. Scientists call this 'temptation bundling.'" Above this text, participants saw a cartoon that visually depicted the concept of temptation bundling (see Fig. 1). On the following screen, participants saw a short video about temptation bundling (see http://bit.ly/2GhdbrM).

Participants were then asked to report what type of entertainment would make them crave their workouts most if it were paired with

Crave Your Workout
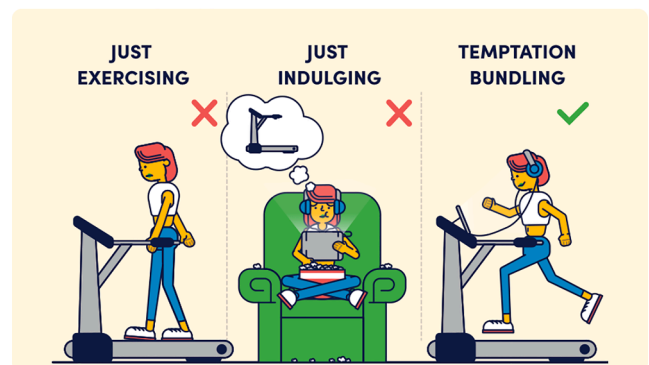
A secret to making working out at the gym into an enjoyable, fun habit is to make a simple rule: I only let myself enjoy my favorite tv shows, audiobooks, or podcasts on my smartphone when exercising. Scientists call this "temptation bundling."

Fig. 1. Participants in the *free audiobook with encouragement to temptation bundle* condition saw this cartoon explaining the concept of temptation bundling during the StepUp Program registration process.

exercise (audiobooks, TV shows, or podcasts) and their preferred genre of that form of entertainment (thrillers, science fiction, fantasy, romance, literary fiction, or dramas). Based on their responses, participants were offered a customized recommendation for media they could use to temptation bundle (e.g., someone who reported preferring sci-fi audiobooks was told that they might enjoy listening to *The Handmaid's Tale, Dune, The Hunger Games,* or *The Water Knife* while exercising). These recommendations were based on the results of a survey of Amazon Mechanical Turk workers ($N = 100$) in which we asked them to share their favorite audiobooks, TV shows, and podcasts by genre. Finally, participants assigned to the *free audiobook with encouragement to temptation bundle* condition received their code for a free audiobook, and were encouraged to sign a pledge promising to only use their favorite form of media at the gym.

Starting the day after they registered for StepUp, participants in both experimental conditions received text message reminders to exercise 30 minutes prior to each scheduled workout and weekly emails with their workout schedule during the four-week program. Some of these text messages and emails also included reminders to download their free audiobook from Audible.com. Additionally, in the *free audiobook with encouragement to temptation bundle* condition, all emails and text messages included reminders to temptation bundle and reminders of participants' pledge to temptation bundle if they had signed the pledge. At the end of the four weeks, all participants received a goodbye text message confirming their completion of the program. Descriptions of text messages and emails sent across conditions can be found in the Online Supplement on pages 15–17 and all study materials are in the Online Supplement on pages 8–14.

Participants. Following our preregistered analysis plan, we only analyzed data from participants who indicated (prior to randomization) that they typically use cardio machines and/or lift weights at the gym. This is because participants who prefer to swim or take fitness classes at the gym could not easily take advantage of temptation bundling. Fig. 2 presents a diagram of the flow of study participants in Study 1A.

Participants were also dropped from our analyses if they requested to withdraw from the study or signed up for the program twice. Specifically, nine participants in the *simple free audiobook* condition and eight participants in the *free audiobook with encouragement to temptation bundle* condition were excluded from our analyses for these reasons.

After excluding the 955 participants who did not indicate that they typically use cardio machines and/or lift weights at the gym and the 17 participants who signed up twice or requested to withdraw their data, we were left with a total of 2,334 24 Hour Fitness members in this study (63% female, average age = 38.84, $SD_{age} = 13.16$; *free audiobook with encouragement to temptation bundle* condition: $N = 1211$; *simple free audiobook* condition: $N = 1123$). These participants were residents of 18 states (57% Californian, 16% Texan, 6% Coloradan, 4% Washingtonian, 4% Oregonian, and 14% from other states). Using Census data, we inferred participants' race based on their first and last names (following Morton, Zettelmeyer, & Silva-Risso, 2003 and Berger & Milkman, 2012). This technique suggests that our study participants were 50% White, 22% Hispanic, 12% Asian, 2% Black, and 14% Other or Unknown. Balance checks reported in Table 1 suggest that random assignment was successful, as we see no difference in observable participant characteristics across experimental conditions.

Statistical analyses. All of our analyses in Study 1A followed our preregistered analysis plan: http://bit.ly/2NUkEBE.

Dependent variables. Our primary outcome of interest was gym attendance, which we measured in two ways. First, we examined the number of gym visits made by each study participant per week (following Acland & Levy, 2015; Charness & Gneezy, 2009; Milkman, Minson, & Volpp, 2014; Royer et al., 2015; Beshears, Lee, Milkman, Mislavsky, & Wisdom, 2020). Second, we examined an indicator for whether each participant visited the gym at least once in a given week ([1 if yes, 0 if no]; following Royer et al., 2015; Beshears et al., 2020).

We analyzed these outcome variables during the four-week StepUp

Program, which we refer to as our "intervention period," and for two different, preregistered post-intervention periods: four weeks and 10 weeks post-intervention.[2] 24 Hour Fitness provided us with up to 52 weeks of pre-intervention gym visit data for each participant, which allowed us to control for participants' pre-intervention gym attendance (following our preregistered analysis plan) and substantially increased our statistical power.

To evaluate the effect of explaining temptation bundling and encouraging participants to commit to bundling, we compared the gym attendance of participants in the *free audiobook with encouragement to temptation bundle* condition to the *simple free audiobook* condition. We relied on a difference-in-difference regression framework, including data on participants' gym visits pre-, during-, and post-intervention and including participant fixed effects in our analyses.

Regression specifications. To measure the effect of the *free audiobook with encouragement to temptation bundle* condition over and above the *simple free audiobook* condition, we ran the following ordinary least squares regression model, with clustered standard errors by participant:

$$y_{it} = \beta_0 + \beta_1 treatment_i * during\_intervention_{it}$$
$$+ \beta_2 treatment_i * post\_intervention_{it} + CX_i + DZ_t + \varepsilon_{it}$$

where i indexes participants and t indexes weeks. The left-hand side y_{it} is either the number of gym visits made by participant i in week t, or a binary indicator for whether participant i visited the gym in week t. On the right-hand side of the equation is an indicator for a participant's experimental condition interacted with an indicator for whether week t is an intervention week (1 if yes, 0 if no) and an indicator for a participant's experimental condition interacted with an indicator for whether week t is a post-intervention week (1 if yes, 0 if no). We also controlled for week fixed effects (denoted by DZ_t terms) and participant fixed effects (denoted by CX_i terms).[3]

Results

StepUp Program registration survey results. We found that 65.4% of participants in our *free audiobook with encouragement to temptation bundle* condition completed all components of the StepUp enrollment process and received their free audiobook code at the end of the enrollment survey, a percentage that is statistically indistinguishable ($p = .816$) from the 65.9% of participants in our *simple free audiobook* condition who completed all components of the StepUp enrollment process and received their free audiobook code. Furthermore, the majority of participants in the *free audiobook with encouragement to temptation bundle* condition (78.6%) signed a pledge to temptation bundle during the four-week StepUp Program. Over a quarter of the participants in the *free audiobook with encouragement to temptation bundle* condition (29.9%) indicated that audiobooks were their preferred type of entertainment to bundle with exercise. Participants downloaded their free Audible audiobook at similar (albeit low) rates across the two experimental conditions: 14% of participants in the *free audiobook with encouragement to temptation bundle* condition and 13% of participants in the *simple free audiobook* condition downloaded their free audiobooks, $z = 0.815$, $p = .415$.

---

[2] We also report supplemental analyses that include 17 weeks of post-intervention data because we were able to obtain complete data on all study participants' gym attendance for up to 17 weeks post-intervention.

[3] Data from Study 1A and Study 1B cannot be posted publicly according to the researchers' agreement with 24 Hour Fitness. Researchers interested in the data used in Study 1A and Study 1B should contact the corresponding author, Erika Kirgios (ekirgios@wharton.upenn.edu), who upon request, will ask 24 Hour Fitness for approval to share de-identified study data. If approval is obtained, the researcher(s) will be granted access to data files and analysis code posted at https://osf.io/ubzh6/. De-identified data and analyses from Studies 2A and 2B are posted at https://osf.io/8jby5/?view_only=2b3e560987af4831b5043a50f01678fb.
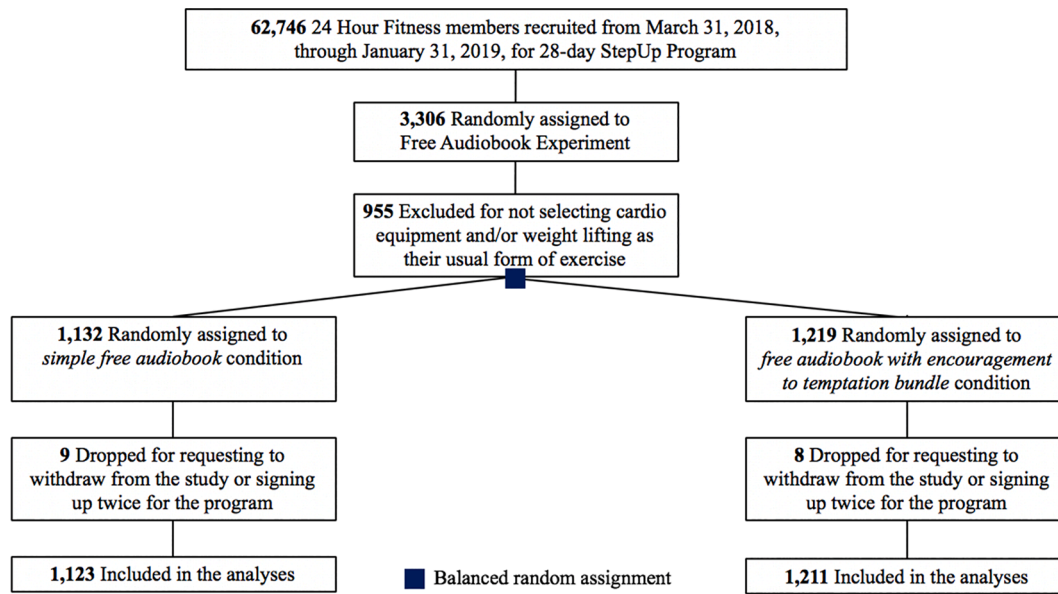
Fig. 2. Flow of study participants in Study 1A.

Table 1
Balance table for Study 1A comparing participant characteristics across treatments.

	Study 1A		
	Simple free audiobook	Free audiobook with encouragement to temptation bundle	*p*-value from *F*-test
Age	38.7 (13.1)	39.0 (13.2)	0.615
Average number of weekly gym visits in the four weeks before joining StepUp	1.2 (1.5)	1.2 (1.4)	0.671
Weeks of gym membership prior to joining StepUp	35.8 (20.4)	36.8 (19.8)	0.262
Female	62.1%	63.0%	0.630
White	50.3%	49.6%	0.688
Black	1.9%	2.0%	0.652
Asian	12.3%	12.3%	0.873
Hispanic	21.7%	22.8%	0.527
Other race	13.7%	13.3%	0.887
Sample size	1,123	1,211	

Note. Participants' age, gender, and weeks of gym membership prior to joining StepUp were provided by 24 Hour Fitness. Each participant's race was inferred using their first name, last name, and Census data (following Morton, Zettelmeyer, and Silva-Risso, 2003 and Berger & Milkman, 2012). Each participant's average number of weekly gym visits in the four weeks before joining StepUp was calculated from 24 Hour Fitness data, and for weeks when participants were not 24 Hour Fitness members, they were presumed to make no gym visits. Standard deviations for means are reported in parentheses. *F*-tests were conducted as a joint test of equality across all treatments.

The impact of encouraging temptation bundling. Participants in Study 1A visited the gym an average of 1.67 times per week during the StepUp Program, and, on average, 59.6% of participants visited the gym at least once a week. Fig. 3 depicts participants' extra weekly gym visits during and after the StepUp Program by experimental condition. As this figure illustrates, the two experimental conditions produced a similar number of workouts during the StepUp Program, but directionally, participants in the *free audiobook with encouragement to temptation bundle* condition were more likely to visit the gym after the program ended. To evaluate the significance of these patterns, we turn to our preregistered regression models described in the section above entitled *Regression*

specifications.

As reported in Table 2, Model 1, participants in the *free audiobook with encouragement to temptation bundle* condition were a marginal, regression-estimated 2.2 percentage points more likely to visit the gym at least once in a given week during the intervention period than participants in the *simple free audiobook* condition ($SE = 0.013$, $p = .095$; a 4.3% increase). However, as Table 2, Model 1 reports, we found no difference in participants' average estimated weekly gym visits across experimental conditions during the intervention period ($b = -0.001$, $SE = 0.051$, $p = .989$).

We next examined differences between conditions after the conclusion of the StepUp Program. During this four-week follow-up period, as reported in Table 2, Model 3, participants in the *free audiobook with encouragement to temptation bundle* condition were also a marginal, regression-estimated 2.5 percentage points more likely to visit the gym in a given week than participants in the *simple free audiobook* condition ($SE = 0.014$, $p = .080$; a 4.9% increase). Furthermore, as reported in Table 2, Model 4, participants went to the gym a marginal, regression-estimated 0.09 more times per week (equivalent to a 6.5% lift in average weekly gym visits) in the *free audiobook with encouragement to temptation bundle* condition than in the *simple free audiobook* condition during the four weeks post-intervention ($SE = 0.049$, $p = .081$).

We saw similar results when we extended our analyses to a 10-week follow-up period. As reported in Table 2, Model 5, participants who were encouraged to temptation bundle visited the gym at least once a week at a regression-estimated 3.2 percentage point higher rate than others ($SE = 0.01$, $p = .012$; a 6.2% increase). However, as reported in Table 2, Model 6, the marginal increase in participants' total number of weekly gym visits produced by encouragement to temptation bundle in our four-week follow-up period dropped below marginal significance levels when our analyses included gym visits up to 10 weeks post-intervention ($b = 0.071$, $SE = 0.045$, $p = .113$).

Overall, these results offer some evidence that participants who were educated about temptation bundling increased their exercise compared to those who did not learn about temptation bundling, supporting Hypothesis 1. Although small, the effects of educating participants about temptation bundling relative to simply giving them a free audiobook were durable, and surprisingly, seemed stronger post-intervention than during the program itself. Teaching people about temptation bundling seemed to be most effective at dissuading them from entirely abandoning the gym, as the effects of learning about temptation bundling
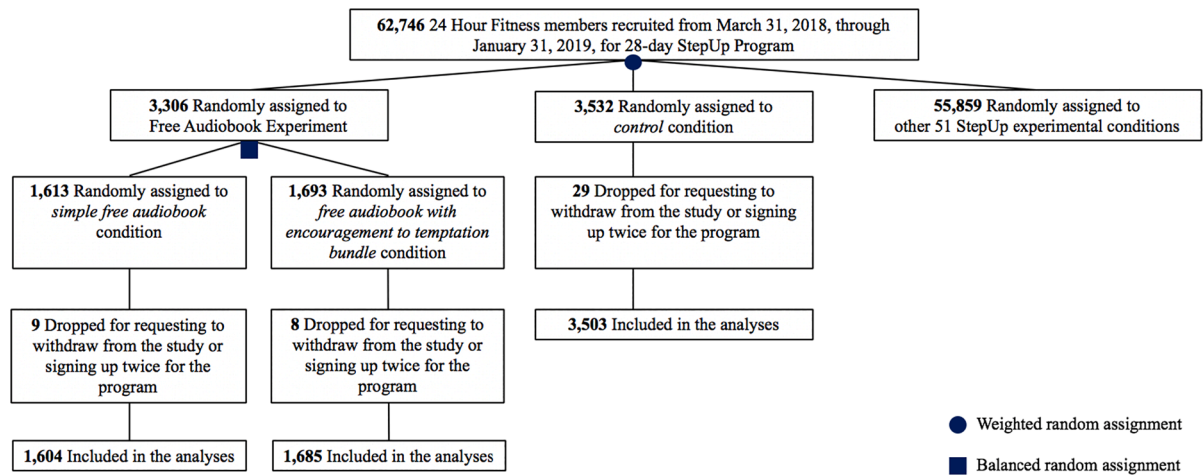
Fig. 3. Flow of study participants in Study 1B.

Table 2
Regression-estimated effects of providing encouragement to temptation bundle in Study 1A during and post-intervention.

Time period of interest	Four-week intervention period		Four weeks post-intervention		Ten weeks post-intervention	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dependent variable	Any weekly gym visit? (Y/N)	Total weekly gym visits	Any weekly gym visit? (Y/N)	Total weekly gym visits	Any weekly gym visit? (Y/N)	Total weekly gym visits
Free audiobook with encouragement to temptation bundle × Indicator for time period of interest	0.022† (0.013)	−0.001 (0.051)	0.025† (0.014)	0.086† (0.049)	0.032* (0.013)	0.071 (0.045)
Fixed effects for participants?	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects for weeks since start of intervention?	Yes	Yes	Yes	Yes	Yes	Yes
N of participants	2334	2334	2334	2334	2334	2334
N of observations	94,146	94,146	103,482	103,482	117,486	117,486
Adjusted R^2	0.437	0.542	0.428	0.535	0.422	. 528

Note. The table reports the results of preregistered ordinary least squares (OLS) regressions predicting study participants' likelihood of visiting the gym at least once in a given week during the four-week StepUp Program intervention period (Model 1), the four-week post-intervention period (Model 3), and the 10-week post-intervention period (Model 5). The table also reports the results of OLS regressions predicting study participants' total weekly workouts during the four-week intervention period (Model 2), the four-week post-intervention period (Model 4), and the 10-week post-intervention period (Model 6). In all models, the key predictor variable is an indicator for being in the *free audiobook with encouragement to temptation bundle* experimental condition interacted with an indicator for the time period of interest (either the four-week intervention period, the four-week post-intervention period, or the 10-week post-intervention period).
Standard errors clustered by participant are reported in parentheses. † $p < 0.10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

were strongest when we focused on whether participants visited the gym at all in a given week.

Discussion. The results of Study 1A suggest that the benefits of education about temptation bundling are small and accrue primarily in the long term. Moreover, encouragement to temptation bundle seems to operate primarily on the extensive margin, increasing the likelihood that people will go to the gym at least once a week, but exerting less influence on their weekly visit total. Although our treatment effect is smaller than in the Milkman et al. (2014) experiment, this is likely due to the differences between the control conditions used in the two experiments. In Milkman et al. (2014), participants in the control condition received a gift certificate to Barnes & Noble that was of equal value to the audiobooks on offer in the treatment condition. In our experiment, participants received a code that would allow them to download a free audiobook from Audible.com in *both* conditions. Merely giving participants a free audiobook could have helped them exercise more often by leaking information about the intentions of the exercise program's creators (per Hypothesis 3): When participants signed up for a program offered by 24 Hour Fitness to help them exercise more regularly and received a free audiobook as part of that program, they assumed the audiobook was intended to help them exercise.

There are two results providing initial support for this contention. First, Fig. 4 indicates that participants in both conditions increased their gym visits during the StepUp intervention period relative to baseline.

Second, findings from a follow-up, preregistered survey of a separate sample of gym-goers[4] ($N = 75$) point to the possibility that the free audiobooks leaked information. In this survey, participants saw screenshots of all study stimuli included in the *simple free audiobook* condition of the StepUp Program. Participants were asked to predict how StepUp Program designers intended for the free audiobook to be used, first in an open-ended response—intended to get at their unbiased opinion—and then in a multiple-choice question in which they indicated whether the audiobook was intended for "entertainment at the gym," "for entertainment outside the gym," or "whenever they wanted." Three independent coders blind to our hypotheses indicated whether each open-ended response mentioned using the audiobooks for entertainment at the gym or not. Sixty out of 75, or 80% of participants, mentioned that the audiobook was intended for use at the gym in their free responses.[5] Furthermore, in a subsequent multiple choice question, 80% of

---

[4] https://aspredicted.org/g39ff.pdf.
[5] Each of the three RAs was given the same Word document containing instructions about how to code the free responses. The text of this Word document is available upon request.
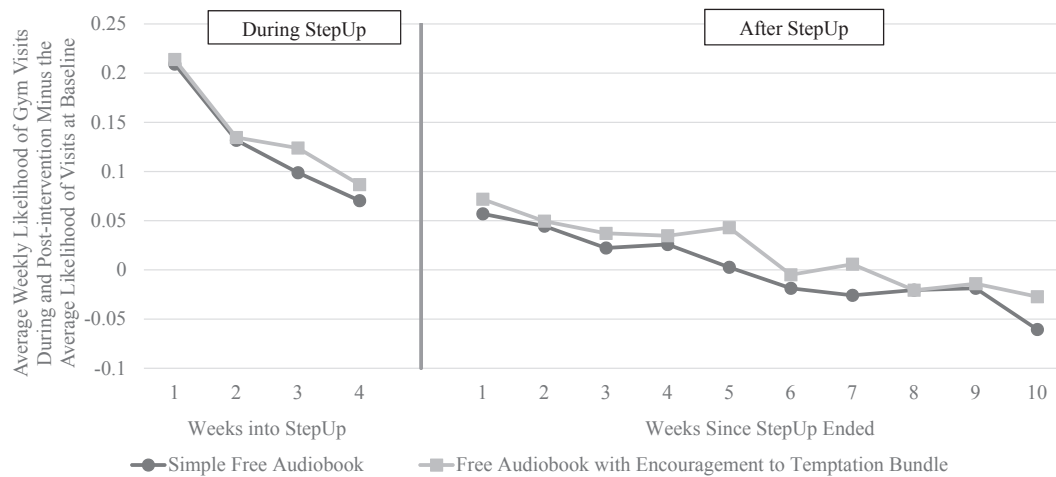
Fig. 4. This graph depicts Study 1A participants' average likelihood of visiting the gym in a given week during and post-intervention by experimental condition after subtracting off their baseline gym attendance (averaged over four weeks pre-intervention). Encouraging participants to temptation bundle leads to a durable boost in average weekly likelihood of gym visits (as shown in the graph), but has negligible effects on average total weekly gym visits. *Note:* When participants were not members of 24 Hour Fitness during the entire, four-week pre-intervention baseline period, zero gym visits were assumed for any missing weeks to ensure balanced panel.

participants indicated that they thought the free audiobook was meant to be used for entertainment at the gym, even though they were given no information about temptation bundling.[6] This is a significantly larger proportion than would be expected by chance ($z = 5.60$, $p < .001$) and suggests that simply providing a free audiobook in this context led participants to infer that they should temptation bundle.

To test whether merely providing participants with a free audiobook in a workout context is sufficient to promote behavior change via information leakage (Hypothesis 4), we make use of our second, quasi-experimental control condition from the StepUp mega-study. Specifically, the StepUp mega-study included a control condition in which participants were exposed to incentives, planning, and pre-workout reminders, exactly like participants in Study 1A, but were not given a free audiobook. In Study 1B, we compare each of the two *free audiobook* conditions from Study 1A to this control condition to assess the effect of giving participants a free audiobook on gym visits.

2.2. Study 1B: Does simply giving participants a free audiobook boost gym visits?

In Study 1B, we explored whether giving free audiobooks to gym members in the context of an exercise-promoting program increased their gym attendance to test Hypothesis 4, and we also examined the benefits of providing both a free audiobook and encouragement to temptation bundle to test Hypothesis 2. We conducted parallel analyses to those presented in Study 1A (closely following our Study 1A preregistration, though Study 1B was not preregistered), this time comparing gym attendance across three conditions: *control*, *simple free audiobook,* and *free audiobook with encouragement to temptation bundle*.

Methods. As in Study 1A, all participants were members of 24 Hour Fitness who enrolled in the StepUp Program. All recruitment and enrollment details were identical to Study 1A.

Experimental design. The *simple free audiobook* and *free audiobook with encouragement to temptation bundle* experimental conditions were described in Study 1A. The control condition in the StepUp mega-study had all of the same features as our two temptation bundling experimental conditions but did not describe a partnership with Audible.com,

ask participants about what they typically do at the gym, or give participants a code that they could use to download a free audiobook from Audible.com.

Randomization. To ensure as many as possible of the 20 separately preregistered experiments in the StepUp mega-study would be adequately powered, StepUp Program enrollees were randomly assigned to experimental conditions with time-varying probabilities. At any given moment, the plurality of participants (40–60%) were randomly assigned with equal probability to one of the conditions in a given, specific experiment (one of the 20 experiments that were separately preregistered) within the overall mega-study, and the remaining participants were assigned across the other experimental conditions with equal probability. After a prespecified number of participants enrolled in StepUp, the probabilities shifted and a new experiment received the plurality of participants. The probabilities shifted a total of 27 times; each time defines a randomization cohort (see Fig. 4 for a flowchart depicting the randomization process in greater detail).

Since the two experimental conditions in which participants were given free audiobooks were designed as part of a single experiment, participants always had an equal probability of being assigned to either of these two conditions in Study 1A. However, the *control* condition we analyze in Study 1B was not designed as part of this single experiment, so at any given time, the probability of being assigned to this condition differed from the probability of being assigned to either of the *free audiobook* conditions. As a result, comparisons between our *free audiobook* experimental conditions and the *control* condition require sample weighting to ensure balance, making these analyses quasi-experimental. Below, in a section of our paper entitled *Sample weighting,* we provide detailed information about the weights we assigned to participants in our analyses to account for the different, time-varying probabilities of being assigned to the *control* versus *free audiobook* conditions in our experiment.

Experimental conditions. As noted above, participants were randomized with unequal probabilities into one of three different conditions: the *control* condition, the *simple free audiobook* condition, or the *free audiobook with encouragement to temptation bundle* condition. Participants assigned to the *control* condition experienced the same registration process as participants in the two *free audiobook* conditions (described in Study 1A), but they did not receive a code that would allow them to download a free audiobook or information about temptation bundling. Every other detail of the participants' experience was identical across the *control* and *free audiobook* conditions.

---

[6] Meanwhile, 14.7% and 5.3% of participants indicated that they thought the free audiobook was intended for use "whenever they wanted" or "outside of the gym," respectively.

Table 3
Balance table for Study 1B comparing participant characteristics across treatments.

	Study 1B			
	Control	Simple free audiobook	Free audiobook with encouragement to temptation bundle	*p*-value from *F*-test
Age	39.2 (13.3)	39.6 (13.4)	39.6 (13.3)	0.294
Average number of weekly gym visits in the four weeks before joining StepUp	1.3 (1.5)	1.2 (1.5)	1.2 (1.4)	0.084
Weeks of gym membership prior to joining StepUp	35.4 (20.3)	35.9 (20.3)	36.9 (19.9)	0.418
Female	57.6%	63.5%	63.6%	0.189
White	43.4%	50.7%	49.8%	0.845
Black	2.8%	2.2%	2.4%	0.888
Asian	9.5%	12.9%	11.8%	0.532
Hispanic	19.5%	20.6%	21.8%	0.769
Other race	24.6%	13.5%	14.2%	0.190
Sample size	3503	1604	1685	

Note. Participants' age, gender, and weeks of gym membership prior to joining StepUp were provided by 24 Hour Fitness. Each participant's race was inferred using their first name, last name, and Census data (following Morton, Zettelmeyer, and Silva-Risso, 2003 and Berger & Milkman, 2012). Each participant's average number of weekly gym visits in the four weeks before joining StepUp was calculated from 24 Hour Fitness data, and for weeks when participants were not 24 Hour Fitness members, they were presumed to make no gym visits. Standard deviations for means are reported in parentheses. *F*-tests were conducted as a joint test of equality across all treatments. *N*s for *simple free audiobook* and *free audiobook with encouragement to temptation bundle* conditions are larger than *N*s reported in Table 1 (Study 1A) as we no longer exclude participants who did not select cardio equipment and/or weight lifting as their usual form of exercise.

Descriptions of text messages and emails sent across conditions can be found in the Online Supplement on pages 15–17, and all study materials (including screenshots of the enrollment and registration process in each experimental condition) are presented in the Online Supplement on pages 8–14.

Participants. As in Study 1A, participants were dropped from the analyses if they requested to withdraw from the study or signed up for the program twice. In total, this led us to drop 30 participants from the *control* condition, nine participants from the *simple free audiobook* condition and eight participants from the *free audiobook with encouragement to temptation bundle* condition.

After these exclusions, a total of 6,792 24 Hour Fitness members were assigned to Study 1B's three different experimental conditions (64% female, average age = 39.26, *SD* = 13.26; *simple free audiobook*: *N* = 1604; *free audiobook with encouragement to temptation bundle*: *N* = 1685; *control*: *N* = 3503). These participants were residents of 26 states (55% Californian, 15% Texan, 6% Coloradan, 4% Washingtonian, 4% Oregonian, and 15% from other states). As in Study 1A, we used Census data to infer participants' race based on their first and last names (following Morton et al., 2003 and Berger & Milkman, 2012). This technique suggests that our study participants were 50% White, 22% Hispanic, 11% Asian, 3% Black, and 14% Other or Unknown. Balance checks reported in Table 3 suggest that random assignment was successful, as we see no difference in observable participant characteristics across experimental conditions after including sample weights (described below) to account for imbalanced random assignment.

Note that in comparing the two *free audiobook* conditions to the *control* condition, we no longer exclude participants from our study whose self-reported typical activities at the gym were incompatible with temptation bundling because these screening questions were not asked

of participants in the *control* condition.[7] This is why Study 1B includes 955 more participants in the two *free audiobook* conditions (*N* = 3289) than the total participant count in Study 1A (*N* = 2334). Naturally, the inclusion of additional participants who could not plausibly temptation bundle vastly reduces our power to detect the effects of encouraging temptation bundling and test Hypothesis 1, which our preregistered Study 1A was designed to measure. We therefore rely on Study 1B solely for the exploration of Hypotheses 2 and 4.

Statistical analyses. We relied on the same general analysis framework described in Study 1A, but this time we compared the two *free audiobook* conditions to the *control* condition. We accounted for the quasi-experimental nature of this comparison by including sample weights to adjust for the different probabilities of assignment to conditions across randomization cohorts in our regression analyses and by including cohort-by-week fixed effects, as detailed below.

Regression specifications: free audiobook conditions versus *control.* To measure the effect of offering participants either a free audiobook or a *free audiobook with encouragement to temptation bundle*, we relied on the following ordinary least squares regression model, with standard errors clustered by participant and sample weights as described in the *Sample weighting* section below:

$$y_{it} = \beta_0 + \sum_{g=1}^{G} \beta_1^g d_i^g \times during\_intervention_{it} + \sum_{g=1}^{G} \beta_2^g d_i^g \times post\_intervention_{it} + \sum_{c=1}^{C} \sum_{t'=2}^{T} \gamma^{c,t'} d_{it}^{c,t'} + CX_i + \varepsilon_{it},$$

where *i* indexes participants, *t* indexes weeks relative to the start of the intervention, *g* indexes experimental condition (our *control* condition was set as the omitted reference group, and *G* = 2 to account for the two *free audiobook* conditions), and *c* indexes randomization cohorts. The left-hand side, y_{it}, represents one of the two dependent variables discussed above: either the number of gym visits made by participant *i* in week *t*, or a binary variable indicating whether participant *i* visited the gym in week *t* (1 = yes, 0 = no). On the right-hand side of the equation are two interaction terms: first, an interaction between an indicator, d_i^g, for whether participant *i* was assigned to condition *g* and an indicator for whether week *t* was during the intervention period for participant *i*, and second, an interaction between d_i^g and an indicator for whether week *t* was during the post-intervention period of interest for participant *i*. We controlled for cohort-by-week interactions (as denoted by $d_{it}^{c,t'}$ terms) and participants fixed effects (as denoted by CX_i terms).

The coefficients β_1^g and β_2^g refer to the treatment effects of experimental condition *g* compared to the *control* condition during the intervention period and the post-intervention period of interest, respectively.

Sample weighting. To account for compositional differences across randomization cohorts when comparing the two *free audiobook* conditions to the *control* condition, we sample-weighted each observation in our regressions such that each experimental group was equally weighted within a cohort, and each cohort was weighted proportionally to its duration. The weight specification used for each observation was as follows:

$$w_{it}^{g,c} = \frac{1}{T_i} * \frac{1}{N^{g,c}} * \frac{1}{G} * \frac{L^c}{L}$$

where *i* indexes participants; *t* indexes weeks relative to the start of the intervention; *g* indexes experimental condition; *c* indexes randomization cohort; T_i indicates the length, in weeks, of the panel data for participant *i*; $N^{g,c}$ indicates the number of participants assigned to experimental

condition *g* during cohort *c*[8]; *G* indicates the number of experimental conditions in the analysis; L^c indicates the length of cohort *c*, in weeks; and *L* is the overall length of the recruitment period, in weeks.

The inclusion of these sample weights allows us to control for both compositional differences between cohorts and for each cohort's timing. By including the inverse of the number of participants assigned to treatment *g* in cohort *c* ($1/N^{g,c}$) in our weighting specification, we weight each treatment equally within a cohort (rather than overweighting treatments that were assigned more participants). For example, if one-fourth of our participants in January were assigned to our control condition and three-fourths were assigned to our two *free audiobook* conditions, we would weight each participant in the control condition three times as heavily as each participant in our *free audiobook* conditions in January. Similarly, by including the relative duration of each cohort (L^c/L) in our weighting formula, we weight each cohort in proportion to its duration. This approach ensures any imbalance in random assignment will not bias our regression-estimated effects of the *free audiobook* conditions.

Results. Participants in Study 1B visited the gym an average of 1.57 times per week during the intervention, and on average, 57.7% of them visited the gym in a given week. Fig. 5 depicts participants' extra gym attendance during and post-intervention by experimental condition after accounting for their baseline gym attendance. As this figure illustrates, participants in both the *simple free audiobook* condition and the *free audiobook with encouragement to temptation bundle* condition exercised directionally more than those in the *control* condition during and after the intervention. To evaluate the significance of these differences, we compared the change in frequency of gym visits for participants in each of the two *free audiobook* treatments to participants in the *control* condition using the regression models described in *Regression specifications: free audiobook conditions* versus *control*.

As Table 4, Models 1–2 report, we found that during the intervention period, participants in the *simple free audiobook* condition visited the gym an estimated 0.23 times more per week (*SE* = 0.114, *p* = .040; a 17.6% increase) than participants in the *control* condition, supporting Hypothesis 4, though they were not significantly more likely to visit the gym in a given week (*b* = 0.021, *SE* = 0.033, *p* = .528). We find that during the intervention period, participants in the *free audiobook with encouragement to temptation bundle* condition visited the gym a marginal, regression-estimated 0.14 times more per week (*SE* = 0.080, *p* = .076; a 10.0% increase) than participants in the *control* condition, and were a regression-estimated 6.7 percentage points more likely to visit the gym in a given week (*SE* = 0.024, *p* = .005; a 13.6% increase), consistent with Hypothesis 2.

As Table 4, Models 3–4 report, we observe slightly larger benefits of the two free audiobook conditions in the four weeks post-intervention. During this period, participants in the *simple free audiobook* condition made an estimated 0.24 more weekly gym visits, on average (*SE* = 0.097, *p* = .015; an 18.4% increase), than participants in the *control* condition and were a regression-estimated 7.6 percentage points more likely to visit the gym in a given week (*SE* = 0.030, *p* = .013; a 15.4% increase), supporting Hypothesis 4. Similarly, participants in the *free audiobook with encouragement to temptation bundling* condition made an estimated 0.16 more average weekly gym visits (*SE* = 0.079, *p* = .041; a 12.2% increase) than participants in the *control* condition and were a

regression-estimated 6.4 percentage points more likely to visit the gym in a given week (*SE* = 0.026, *p* = .012; a 13.0% increase), supporting Hypothesis 2.

When we run the same analyses but instead look at a 10-week post-intervention follow-up period, the results are remarkably consistent, as reported in Table 4, Models 5–6. Participants in the *simple free audiobook* condition made a directional but insignificant estimated 0.14 more weekly gym visits (*SE* = 0.094, *p* = .127; a 10.7% increase) than participants in the *control* condition, and were directionally but insignificantly more likely to visit the gym in a given week (*b* = 0.035; *SE* = 0.025, *p* = .155; a 7.1% increase). Participants in the *free audiobook with encouragement to temptation bundle* condition made an estimated 0.13 more weekly gym visits (*SE* = 0.066, *p* = .046; a 9.9% increase) than participants in the *control* condition and were a regression-estimated 5.1 percentage points more likely to visit the gym each week than participants in the *control* condition during the 10-week post-intervention period (*SE* = 0.019, *p* = .006; a 10.3% increase), supporting Hypothesis 2.

Although we focus on our preregistered analyses in Study 1A to compare the two *free audiobook* conditions, we also ran Wald tests to compare the regression-estimated impact of the *simple free audiobook* condition with that of the *free audiobook with encouragement to temptation bundle* condition in Study 1B. These comparisons differ substantially from the comparisons presented in Study 1A because for comparability with our *control* condition, we needed to include in our analyses 955 additional participants who reported that they did not typically use cardio equipment or lift weights at the gym (but rather swam or took classes) and thus would not be able to temptation bundle while exercising. Unsurprisingly, when diluted by an influx of 41% more study participants who could not benefit from temptation bundling (and whom our preregistered Study 1A analyses excluded), our *simple free audiobook* and *free audiobook with encouragement to temptation bundle* conditions no longer differ significantly in any comparisons (in all Wald tests, *p*'s > 0.10).

Alternative post-intervention follow-up periods. To ensure the robustness of our Study 1 regressions presented in Tables 2 and 4 to alternative follow-up periods after the conclusion of our intervention, we ran supplementary analyses that are detailed in the Online Supplement. Specifically, we extended our post-intervention analyses to include all 17 weeks of post-intervention data made available to us by 24 Hour Fitness, and we found that our results are remarkably robust in this lengthier follow-up period, as reported in Online Supplement Tables S1 and S2.

Heterogeneity Analyses. We tested whether our estimated treatment effects varied by gender, age, race, exercise frequency during the four weeks pre-intervention, or length of gym membership. We found that none of these participant characteristics were robustly associated with any change in our estimated treatment effect.[9]

Discussion. In Study 1B, we find that offering participants access to free audiobooks, regardless of whether temptation bundling was explicitly encouraged, significantly increased their gym visits during the StepUp Program and for at least 10 weeks after it concluded, supporting

---

[8] During high-volume registration periods, bugs in our system led to mistakes in random assignment for 6.6% of the participants in the overall StepUp megastudy. These individuals were accidentally assigned to treatment conditions that differed from their intended treatment condition according to the pre-defined randomization matrix. Our weighting accounts for this bug because we define $N^{g,c}$ as the number of participants who were *actually* assigned to condition *g* during cohort *c*, instead of the number of participants who were *intended* to be assigned to condition *g* during cohort *c* according to the randomization matrix.

[9] We did find that the estimated effects of our *simple free audiobook* condition on whether a participant visited the gym in a given week were larger for men than for women, but only during the four (*p* = .029) and ten weeks (*p* = .014) post-intervention.
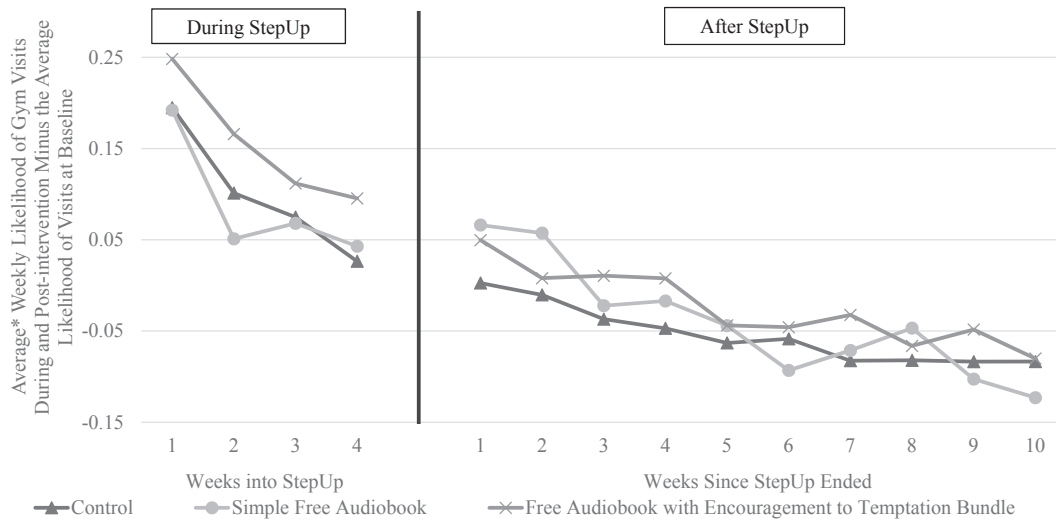
Fig. 5. This graph depicts Study 1B participants' average weekly gym visits during and post-intervention by experimental condition after subtracting off their baseline gym attendance (averaged over four weeks pre-intervention). *Note:* When participants were not members of 24 Hour Fitness during the entire, four-week pre-intervention baseline period, zero gym visits were assumed for any missing weeks to ensure a balanced panel. Average weekly gym visits were weighted to account for the different probabilities of being assigned to different conditions as described in *Sample weighting*.

Table 4
Regression-estimated effects of providing a free audiobook and of providing encouragement to temptation bundle in Study 1B during and post-intervention.

Time period of interest	Four-week intervention period		Four weeks post-intervention		Ten weeks post-intervention	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dependent variable	Any weekly gym visit? (Y/N)	Total weekly gym visits	Any weekly gym visit? (Y/N)	Total weekly gym visits	Any weekly gym visit? (Y/N)	Total weekly gym visits
Simple free audiobook × Indicator for time period of interest	0.021	0.235*	0.076*	0.236*	0.035	0.143
	(0.033)	(0.114)	(0.030)	(0.097)	(0.025)	(0.094)
Free audiobook with encouragement to temptation bundle × Indicator for time period of interest	0.067**	0.142†	0.064*	0.161*	0.051**	0.131*
	(0.024)	(0.080)	(0.026)	(0.079)	(0.019)	(0.066)
Fixed effects for participants?	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects for weeks since start of intervention × Cohort interaction?	Yes	Yes	Yes	Yes	Yes	Yes
N of participants	6792	6792	6792	6792	6792	6792
N of observations	270,892	270,892	298,060	298,060	338,812	338,812
Adjusted R^2	0.451	0.570	0.431	0.551	0.424	0.539

Note. The table reports the results of ordinary least squares (OLS) regressions predicting study participants' likelihood of visiting the gym at least once in a given week during the four-week StepUp Program intervention period (Model 1), the four-week post-intervention period (Model 3), and the 10-week post-intervention period (Model 5). The table also reports the results of OLS regressions predicting study participants' total weekly workouts during the four-week intervention period (Model 2), the four-week post-intervention period (Model 4), and the 10-week post-intervention period (Model 6). In all models, the key predictor variables are an indicator for being in the *simple free audiobook* condition interacted with an indicator for the time period of interest and an indicator for being in the *free audiobook with encouragement to temptation bundle* condition interacted with an indicator for the time period of interest.
Standard errors clustered by participant are reported in parentheses. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

both Hypotheses 2 and 4.[10]

The results of Study 1B suggest that when people were offered a free audiobook by their gym as part of a program designed to boost their

---

[10] Interestingly, the effects of our temptation bundling and free audiobook conditions increased in magnitude and significance during the four-week post-intervention period (see Table 4). This pattern of results may stem from a ceiling effect at the outset of the intervention period. Across all conditions, participants tended to concentrate their visits during the first few weeks of the intervention period, and then spread visits more evenly across weeks during the post-intervention period. This may have made it difficult for differences between conditions to emerge during the intervention period. A ceiling effect could also explain why, in Study 1A, the benefits of explaining temptation bundling relative to simply offering participants a free audiobook grew in the post-intervention period (see Table 2).

exercise, the audiobook's intended use was clear: It was meant to be enjoyed during workouts at the gym. It may be that there was little added benefit of explaining temptation bundling to participants because providing the free audiobook in this context was sufficiently instructive, consistent with Hypothesis 3.

Of course, there are alternative explanations for the results of Study 1B. Receiving the free audiobook could have induced positive feelings, "warm glow," or gratitude toward the gym for participants, which made gym visits more appealing (Andreoni, 1995; Armenta, Fritz, & Lyubomirsky, 2017; Falk, 2007; Isen, Shalker, Clark, & Karp, 1978; List & Lucking-Reiley, 2002). Listening to or recalling the free audiobook outside the gym may also have kept the gym top-of-mind for participants, reminding them to visit the gym more often (Bordalo, Gennaioli, & Shleifer, 2013; Chetty, Looney, & Kroft, 2009; Dohmen, Falk, Huffman, & Sunde, 2006; Finkelstein, 2009).

To assess the plausibility of these alternative explanations for the

benefits of simply providing free audiobooks and to test Hypothesis 3, we conducted two follow-up scenario experiments.

3. Study 2: Follow-up lab studies

3.1. Study 2A

In this study, we tested Hypothesis 3: that providing people with a *want* (in this case, a free audiobook) that can be easily bundled with a *should* behavior they are seeking to boost (in this case, exercise) leaks information that the *want* ought to be paired with the *should* to make the *should* more attractive. We also explored alternative explanations for why providing a free audiobook to StepUp participants might have increased gym visits.

Methods: At an East Coast university that includes gym membership as part of its tuition and fees, 304 students with university gym memberships completed an online survey in exchange for $5 and a potential $1 bonus, both paid via Amazon gift card.

Procedure. This experiment was a three-condition scenario study. In all conditions, participants first read a brief description of the StepUp Program. They were then told that after learning more about StepUp, they would be asked to make predictions about how participants reacted to the program.

Survey participants were randomly assigned to view screenshots of the actual StepUp registration content viewed by participants in one of Study 1's three experimental conditions: the *simple free audiobook* condition, the *free audiobook with encouragement to temptation bundle* condition, or the *control* condition. Participants received a $1 bonus if they correctly answered comprehension check questions about the content they viewed.

After viewing StepUp screenshots and completing comprehension check questions, participants predicted how many days a week they thought StepUp participants visited the gym during the StepUp Program. Then, participants rated the likelihood that StepUp participants visited the gym because they (1) "wanted to listen to an audiobook while working out at the gym," (2) "were reminded to visit the gym whenever they listened to an audiobook," (3) "felt grateful to the gym," (4) "felt positively about the gym," or (5) "wanted to earn money for going to the gym." Each rating was collected on a 5-point scale (from "Very Unlikely" to "Very Likely").

Finally, participants indicated how frequently they exercised and, for both *free audiobook* conditions, if and where they would listen to their free audiobook if they were in the StepUp Program. Participants also answered optional questions about their age, gender, and ethnicity. Complete survey materials and screenshots are available in the Online Supplement on pages 18–35.

Results and discussion

Ninety-eight percent of survey-takers passed our comprehension check.[11] To assess the possibility that gratitude or positive feelings toward the gym generated by simply receiving a free audiobook would increase gym members' exercise, we compared survey participants' ratings of these potential motivators across our *control* and *simple free audiobook* conditions. As depicted in Fig. 6, we did not find any evidence that gym members thought the receipt of a *simple free audiobook* in the StepUp program would create gratitude or positive feelings that would boost exercise over and above the *control* condition (all *p*'s from pairwise comparisons >0.35).

However, as depicted in Fig. 6, and consistent with Hypothesis 3,[12]

we find participants did anticipate that receiving a *simple free audiobook* in the StepUp program would cause gym members to temptation bundle more ($\Delta = 0.398$; $t(204) = 2.55$, $p = .011$) and remind them to attend the gym more ($\Delta = 0.523$; $t(204) = 3.19$, $p = .002$), which would both, in turn, boost their exercise. The *free audiobook with encouragement to temptation bundle* condition generally moved in parallel with the *simple free audiobook* condition, though participants expected gym members to temptation bundle more ($\Delta = 0.427$; $t(197) = 2.82$, $p = .005$) and think about the gym more ($\Delta = 0.459$; $t(197) = 2.79$, $p = .006$) in the *free audiobook with encouragement to temptation bundle* condition than in the *simple free audiobook* condition. These results are consistent with the larger long-term boost in weekly gym visits in the *free audiobook with encouragement to temptation bundle* condition relative to the *simple free audiobook* condition detected in Study 1A and consistent with Hypothesis 1. Pairwise comparisons between all experimental conditions are reported in Online Supplement Table S3 and depicted in Fig. 6.

3.2. Study 2B

Study 2A demonstrated that gym members inferred StepUp participants were likely motivated to visit the gym in order to temptation bundle their audiobook with their workouts, consistent with Hypothesis 3. In Study 2B, we build on this finding to assess whether this inference was driven by the fact that the entity providing the *want* item for bundling with a *should* was also encouraging engagement in the *should* behavior. An information leakage account would suggest that participants only infer they should temptation bundle if they receive their free audiobook from an exercise-related context. Thus, testing whether different inferences are made about the gift of an audiobook when its provider is encouraging exercise versus some unrelated behavior can offer further insight into the validity of Hypothesis 3.

Methods. Participants included 602 Prolific users who completed a preregistered online experiment in exchange for $0.50.[13]

Procedure. This experiment was a two-condition scenario study. Participants were randomly assigned to one of two conditions: the *exercise program* condition and the *reading program* condition. In the *exercise program* condition, participants were instructed to imagine they joined StepUp, a free 28-day workout rewards program designed by scientists to help people build exercise habits. Participants in the *reading program* condition were asked to imagine they joined ReadUp, a free 28-day virtual book club designed by scientists to help people build reading habits. In both conditions, participants were told they received a free audiobook when they signed up for their 28-day program.

After reading these instructions and completing comprehension check questions, participants were asked to indicate the likelihood that they would listen to their free audiobook while engaging in a series of different activities on a scale from -3 ("very unlikely") to 3 ("very likely"). The activities were presented in random order and included exercising, taking a break from work, commuting, running errands, doing chores, other and "whenever, no preferences." Participants then indicated when they thought program designers most hoped they would listen to their free audiobook, selecting one option from the aforementioned list of choices.

Participants also answered optional questions about their age, gender, and ethnicity at the end of the survey. Complete survey materials and screenshots are available in the Online Supplement on pages 36–40.

Results and discussion

Following our preregistration, we analyzed whether participants' reported likelihood of listening to the free audiobook while exercising differed across the two conditions. Participants in the *exercise program*

---

[11] We include the data from all survey takers, including those who failed the comprehension check, in our main analyses. Our results do not change when we exclude the six participants who failed the comprehension check.

[12] Note that this hypothesis to explain our Study 1B results was generated ex post rather than ex ante.

[13] This study was preregistered on AsPredicted.org: https://aspredicted.org/pg8g3.pdf.
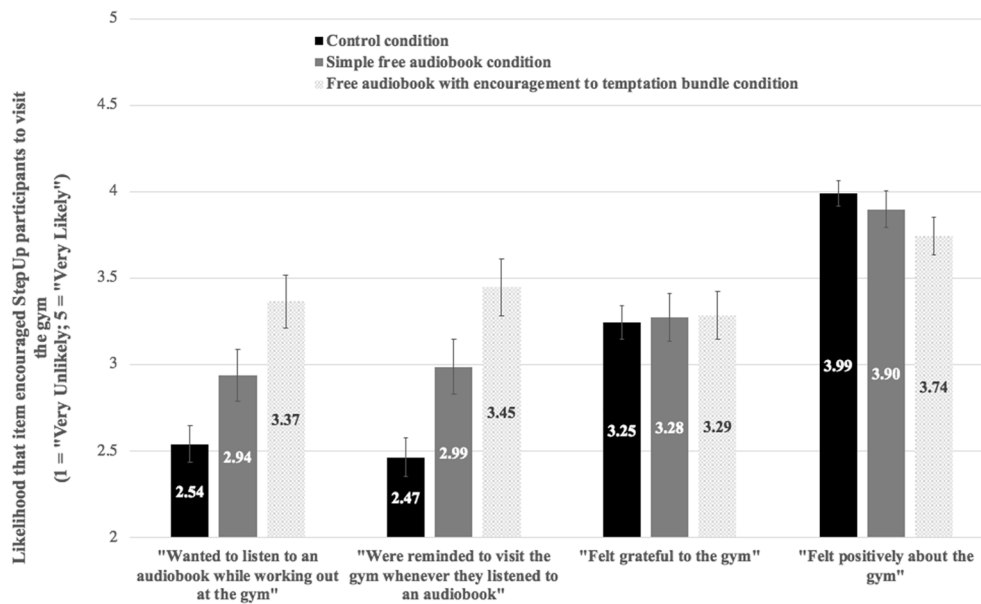
Fig. 6. Study 2A participants' ratings of the likelihood that each of four possible motivators encouraged participant gym visits during the StepUp Program by experimental condition. The error bars depict 95% confidence intervals.

condition were more likely to indicate that they would listen to their audiobook while exercising than those in the *reading program* condition ($\Delta = 0.629$; $t(600) = 3.50$, $p < .001$). Furthermore, 66.2% of participants in the *exercise program* condition indicated that program designers most hoped they would listen to their free audiobook while exercising, while only 2.64% made this choice in the *reading program* condition ($z(600) = 16.35$, $p < .001$).

We also tested whether likelihood of listening to the audiobook while exercising was mediated by inferences about program designers' intentions (following our preregistration). First, we documented a significant main effect of assignment to the *exercise program* condition on participants' perception that program designers intended for the audiobook to be used while exercising ($b = 0.640$, $SE = 0.028$, $p < .001$). The relationship between indicating that the audiobook was intended for use while exercising and participants' reported likelihood of listening to the audiobook while exercising was also significant ($b = 0.926$, $SE = 0.254$, $p < .001$). Consistent with our mediation hypothesis, the effect of assignment to the *exercise program* condition on study participants' likelihood of using the audiobook while exercising ($b = 0.630$, $SE = 0.180$, $p < .001$) was eliminated when controlling for participants' inferences about whether the audiobook was meant to be used while exercising ($b = 0.037$, $SE = 0.241$, $p = .877$). A Sobel test confirmed that this reduction in effect size was significant ($b = 0.593$, $SE = 0.165$, $p < .001$), and a 5,000-sample bootstrap analysis (MacKinnon, Fairchild, & Fritz, 2007; Shrout & Bolger, 2002) also produced a 95% bias-corrected confidence interval for the size of the indirect effect that excluded zero (95% CI [0.287, 0.919]).

Consistent with an information leakage account and Hypothesis 3, these results demonstrate that participants infer they are meant to use their free audiobook while exercising when the free audiobook is provided in the context of an exercise program, but not when it is provided in a context unrelated to exercise. The supplier of the audiobook leaked information about its intended purpose, which may help explain why participants in our field experiment who received a free audiobook (but no encouragement to temptation bundle) from a program designed to boost their exercise increased their gym attendance.

4. General discussion

We provide evidence that people can successfully self-impose

temptation bundling as a behavior change strategy, bundling *wants* with *shoulds* to increase their adherence to *should* behaviors. In one of the largest-ever field experiments studying exercise frequency, we find that teaching people about temptation bundling and providing them with unrestricted access to tempting content can durably increase their gym visits. Specifically, pairing a *free audiobook with encouragement to temptation bundle* increased the likelihood of a weekly gym visit by 10–14% and average weekly gym visits by 10–12% during and for up to seventeen weeks after a four-week intervention period.

We also provide suggestive evidence that people can infer and apply behavior change strategies on their own when given a gift that leaks information about its intended use. This may explain why encouraging temptation bundling and providing a free audiobook only had modest benefits over and above giving our study participants a free audiobook with no encouragement to bundle. If it was obvious that the gift of a free audiobook from their gym meant temptation bundling was encouraged, then explicit encouragement to bundle would be of little value. Indeed, simply giving gym members free audiobooks boosted gym visits by 11% during the four-week intervention and by 18% for four weeks post-intervention. Follow-up lab studies provide support for our information leakage account, demonstrating that participants likely inferred they should temptation bundle and helping to rule out alternative explanations for the benefits of giving away free audiobook (e.g., reciprocity or warm glow). Our findings suggest both that people can be highly sophisticated in pursuing behavior change and that extremely low-touch interventions can successfully promote change. In fact, we find evidence that a well-timed gift can be enough to help people improve their habits.

Overall, our findings suggest that practitioners who wish to encourage the adoption of *should* behaviors may be able to effect change by providing bundle-ready *wants* in the right context and encouraging temptation bundling. Importantly, the effectiveness of this strategy seems to be remarkably robust across various subpopulations, suggesting that many people can benefit from learning how to temptation bundle and receiving a bundle-ready *want*. Finally, it's likely that the more accessible practitioners can make bundle-ready *wants* to participants, the better.

It's worth noting that our findings with respect to Hypothesis 2 are consistent with those of Milkman and colleagues' 2014 temptation bundling study. In one experimental condition, Milkman et al. (2014)

provided participants with iPods equipped with free audiobooks (so no initiative was required on the part of participants to obtain the books), monitored an initial workout while participants temptation bundled, explained temptation bundling in a live laboratory session, and produced an initial (marginal) 0.27 extra gym visits per week relative to a control group that was given a gift certificate to Barnes & Noble (not an audiobook) and no explanation of temptation bundling. Meanwhile, our intervention was fully digital and provided participants with a code to download a free audiobook from Audible.com on their own. Our temptation bundling intervention induced an extra 0.13–0.16 gym visits per week compared with a control group given no instructions about temptation bundling and no free audiobook. The fact that our digital intervention was roughly half as effective as this face-to-face intervention where 100% of participants received audiobooks seems reasonable. Interestingly, the benefits of our digital intervention were far longer-lasting.[14]

The benefits of encouraging people to temptation bundle that we document are consistent with the literature suggesting some people have the capacity to set and stick to personal rules even without employing commitment devices (Ainslie, 1992; Ainslie, 2001; Thaler, 2000; Wertenbroch, 1998). We demonstrate that teaching people about a strategy intended to help them avoid the pitfalls of temptation and fulfill their long-term goals can produce lasting behavior change. These findings add support for the theory that people who are sophisticated about their present bias can self-impose rules or strategies to help them exert self-control.

Among our most intriguing findings is the suggestive evidence that explicit instructions about how temptation bundling works may not be necessary when the technique is suggested implicitly. Our results indicate that providing a tempting *want* in the context of a program designed to increase a *should* behavior (i.e., a gym handing out free audiobooks) may be sufficient to inspire temptation bundling. This finding aligns with and builds on prior work on information leakage (Tannenbaum et al., 2013; McKenzie et al., 2006; McKenzie & Nelson, 2003; Sher & McKenzie, 2006), suggesting that goal-seekers can be sophisticated about gleaning useful recommendations from their environments.

Future work confirming this hypothesized mechanism for Study 1B's findings would be valuable. The design of our field experiment made it impossible to verify whether participants who received a free audiobook actually used it to temptation bundle, although comparable download rates across experimental conditions are suggestive of similar usage patterns whether or not bundling was encouraged. We also could not fully rule out gratitude or warm glow effects as drivers of gym visits in the free audiobook conditions without a *control* condition where participants received a cash gift of equal monetary value to a free audiobook. While the results of Studies 2A and 2B suggest our field experiment participants likely inferred they should use the free audiobook at their gym, further confirmation would be useful.

Future work might also explore why teaching participants about temptation bundling appears particularly beneficial in the long run. We speculate that participants who were taught about temptation bundling could apply this concept more generally than those who were only provided with a free audiobook. Teaching participants how to temptation bundle may have more firmly cemented the idea of pairing exercise with temptations than merely receiving a free audiobook from an exercise program (Hertwig & Grüne-Yanoff, 2017). Furthermore, participants who were taught how to temptation bundle may have enacted a broader and more flexible personal rule, which produced stickier

behavior change (Beshears et al., 2020; Mollick & Rothbard, 2014; Woolley & Fishbach, 2016). Because participants taught to temptation bundle were encouraged to bundle whatever media they most preferred (e.g., TV shows, podcasts) with workouts, they may have bundled exercise with their favorite *want*, not only with audiobooks. In contrast, participants who simply received a free audiobook might have formed a narrower personal rule which only pertained to pairing the gifted audiobook with exercise. These participants might not have considered how to make temptation bundling work best for them personally (e.g., trying out an enjoyable podcast instead of another audiobook) or repeatedly (e.g., they may not have thought to download a new audiobook after the first concluded).[15]

Although a strength of our research is the focus on an objective outcome variable measured unobtrusively over many months' time, this measure is nonetheless imperfect. We could not, for instance, measure the impact of our interventions on exercise quality or intensity. It is possible that temptation bundling could be a distraction from exercise, decrease the intensity of exercise, or crowd out certain forms of exercise in favor of those that are more easily bundled with a temptation. On the other hand, prior work suggests that people persist more on activities they find fun (Woolley & Fishbach, 2016), so temptation bundling may have increased the quality and/or intensity of exercise by making it more enjoyable. Future research should explore how temptation bundling influences the quality and intensity of engagement in the target *should* activity.

CRediT authorship contribution statement

---

[14] Our effects were likely long-lasting in part because the original temptation bundling study was interrupted by a gym closure for Thanksgiving break. But 24 Hour Fitness gyms rarely close. Furthermore, our participants were recruited over a roughly one-year period with staggered start dates. Thus, any disruptions to participants' schedules or to their consumption of tempting audiobooks induced by holidays would have been smoothed across participants.

[15] This could also help explain why encouraging temptation bundling and providing a free audiobook produced far more durable benefits in Study 2B than in the 2014 study by Milkman et al. Participants in the Milkman et al. (2014) study of temptation bundling were specifically instructed to listen to their free audiobook while visiting the gym, so they might not have formed a flexible and adaptable temptation bundling habit.

References

2018 Physical Activity Guidelines Advisory Committee Scientific Report A-4. (2018). Retrieved from https://health.gov/paguidelines/second-edition/report/pdf/PAG_Advisory_Committee_Report.pdf.

Acland, D., & Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science, 61*(1), 146–160. https://doi.org/10.1287/mnsc.2014.2091.

Ainslie, G. (1992). *Studies in rationality and social change. Picoeconomics: The strategic interaction of successive motivational states within the person*. Cambridge, England: Cambridge University Press.

Ainslie, G. (2001). *Breakdown of will*. Cambridge, England: Cambridge University Press.

Andreoni, J. (1995). Cooperation in public-goods experiments: Kindness or confusion? *American Economic Review, 85*(4), 891–904.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science, 13*(3), 219–224.

Armenta, C. N., Fritz, M. M., & Lyubomirsky, S. (2017). Functions of positive emotions: Gratitude as a motivator of self-improvement and positive change. *Emotion Review, 9*(3), 183–190. https://doi.org/10.1177/1754073916669596.

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

Babcock, P., Bedard, K., Charness, G., Hartman, J., & Royer, H. (2015). Letting down the team? Social effects of team incentives. *Journal of the European Economic Association, 13*(5), 841–870. https://doi.org/10.1111/jeea.12131.

Ballard, Jamie. Exercising more and saving money are the most popular 2020 new year's resolutions. YouGov, YouGov, 2 Jan. 2020, today.yougov.com/topics/lifestyle/articles-reports/2020/01/02/new-years-resolutions-2020-health-finance.

Bazerman, M. H., Tenbrunsel, A. E., & Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review, 23*(2), 225–241.

Bénabou, R., & Tirole, J. (2004). Willpower and personal rules. *The American Economic Review, 112*(4), 848–886.

Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research, 49*(2), 192–205.

Beshears, J., Lee, H. N., Milkman, K. L., Mislavsky, R. & Wisdom, J. (2020). Creating exercise habits using incentives: The tradeoff between flexibility and routinization. Manuscript in preparation.

Boals, A., Vandellen, M. R., & Banks, J. B. (2011). The relationship between self-control and health: The mediating effect of avoidant coping. *Psychology & Health, 26*(8), 1049–1062. https://doi.org/10.1080/08870446.2010.529139.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy, 121*(5), 803–843. https://doi.org/10.1086/673885.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics, 2*(1), 671–698.

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica, 77*(3), 909–931.

Cheema, A., & Soman, D. (2006). Malleable mental accounting: The effect of flexibility on the justification of attractive spending and consumption decisions. *Journal of Consumer Psychology, 16*(1), 33–44. https://doi.org/10.1207/s15327663jcp1601_6.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *The American Economic Review, 99*(4), 1145–1177.

Condliffe, S., Işgın, E., & Fitzgerald, B. (2017). Get thee to the gym! A field experiment on improving exercise habits. *Journal of Behavioral and Experimental Economics, 70*, 23–32. https://doi.org/10.1016/j.socec.2017.07.007.

DellaVigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *American Economic Review, 96*(3), 694–719.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2006). Seemingly irrelevant events affect economic perceptions and expectations: The FIFA World Cup 2006 as a natural experiment. https://ssrn.com/abstract=928830.

Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics, 6*(2), 1–31. https://doi.org/10.1257/app.6.2.1.

Duckworth, A. L., White, R. E., Matteucci, A. J., Shearer, A., & Gross, J. J. (2016). A stitch in time: Strategic self-control in high school and college students. *Journal of Educational Psychology, 108*(3), 329–341. https://doi.org/10.1037/edu0000062.

Duckworth, A. L., Milkman, K. L., & Laibson, D. (2018). Beyond willpower: Strategies for reducing failures of self-control. *Psychological Science in the Public Interest, 19*(3), 102–129. https://doi.org/10.1177/1529100618821893.

Falk, A. (2007). Gift exchange in the field. *Econometrica, 75*(5), 1501–1511. https://doi.org/10.1111/j.1468-0262.2007.00800.x.

Finkelstein, A. (2009). E-ZTAX: Tax salience and tax rates. *Quarterly Journal of Economics, 124*(3), 969–1010. https://doi.org/10.1162/qjec.2009.124.3.969.

Grady, K. E., Goodenow, C., & Borkin, J. R. (1988). The effect of reward on compliance with breast self-examination. *Journal of Behavioral Medicine, 11*(1), 43–57. https://doi.org/10.1007/bf00846168.

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: A pooled analysis of 358 population-based surveys with 1·9 million participants. *The Lancet Global Health, 6*(10). https://doi.org/10.1016/s2214-109x(18)30357-7.

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science, 12*(6), 973–986. https://doi.org/10.1177/1745691617702496.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin, 118*(2), 248–271. https://doi.org/10.1037/0033-2909.118.2.248.

Isen, A. M., Shalker, T. E., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology, 36*(1), 1–12. https://doi.org/10.1037/0022-3514.36.1.1.

Kaur, S., Kremer, M., & Mullainathan, S. (2010). Self-control and the development of work arrangements. *American Economic Review, 100*(2), 624–628. https://doi.org/10.1257/aer.100.2.624.

Kay, A. C., Wheeler, S. C., Bargh, J. A., & Ross, L. (2004). Material priming: The influence of mundane physical objects on situational construal and competitive behavioral choice. *Organizational behavior and human decision processes, 95*(1), 83–96.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics, 112*(2), 443–478.

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. *The Lancet, 380*(9838), 219–229. https://doi.org/10.1016/s0140-6736(12)61031-9.

List, J. A., & Lucking-Reiley, D. (2002). The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. *Journal of Political Economy, 110*(1), 215–233. https://doi.org/10.1086/324392.

McKenzie, C. R., Liersch, M. J., & Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science, 17*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2006.01721.x.

McKenzie, C. R., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin & Review, 10*(3), 596–602.

Milkman, K. L., Gromet, D., Ho, H., Kay, J. S., Lee, T., Pandilosky … Duckworth, A. L. (2020). A mega-study approach to evaluating interventions. Working paper.

Milkman, K. L., Minson, J. A., & Volpp, K. G. M. (2014). Holding the Hunger Games hostage at the gym: An evaluation of temptation bundling. *Management Science, 60*(2), 283–299. https://doi.org/10.1287/mnsc.2013.1784.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2008). Harnessing our inner angels and demons: What we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making. *Perspectives on Psychological Science, 3*(4), 324–338.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2009). Highbrow films gather dust: Time-inconsistent preferences and online DVD rentals. *Management Science, 55*(6), 1047–1059.

Mollick, E. R., & Rothbard, N. (2014). Mandatory fun: Gamification and the impact of games at work. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.2277103.

Morton, F. S., Zettelmeyer, F., & Silva-Risso, J. (2003). Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities? *Quantitative Marketing and Economics, 1*(1), 65–92.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89*(1), 103–124.

O'Donoghue, T., & Rabin, M. (2011). Doing it now or later. In: C. F. Camerer, G. Loewenstein, & M. Rabin (Eds.), Advances in behavioral economics (pp. 223–251). https://doi.org/10.2307/j.ctvcm4j8j.12.

Patel, M. S., Benjamin, E. J., Volpp, K. G., Fox, C. S., Small, D. S., Massaro, J. M., … Murabito, J. M. (2017). Effect of a game-based intervention designed to enhance social incentives to increase physical activity among families. *JAMA Internal Medicine, 177*(11), 1586–1593. https://doi.org/10.1001/jamainternmed.2017.3458.

Patel, M. S., Small, D. S., Harrison, J. D., Fortunato, M. P., Oon, A. L., Rareshide, C. A. L., … Hilbert, V. (2019). Effectiveness of behaviorally designed gamification interventions with social incentives for increasing physical activity among overweight and obese adults across the United States. *JAMA Internal Medicine, 179*(12), 1624–1632. https://doi.org/10.1001/jamainternmed.2019.3505.

Prinsen, S., de Ridder, D. T., & de Vet, E. (2013). Eating by example. Effects of environmental cues on dietary decisions. *Appetite, 70*, 1–5.

Royer, H., Stehr, M., & Sydnor, J. (2015). Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a Fortune-500 company. *American Economic Journal: Applied Economics, 7*(3), 51–84.

Schroeder, S. A. (2007). We can do better—Improving the health of the American people. *New England Journal of Medicine, 357*(12), 1221–1228. https://doi.org/10.1056/nejmsa073350.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130*(3), 380–400. https://doi.org/10.1037/0096-3445.130.3.380.

Sher, S., & McKenzie, C. R. (2006). Information leakage from logically equivalent frames. *Cognition, 101*(3), 467–494.

Stavrova, O., & Kokkoris, M. D. (2017). Struggling to be liked: The prospective effect of trait self-control on social desirability and the moderating role of agreeableness. *International Journal of Psychology, 54*(2), 232–236. https://doi.org/10.1002/ijop.12444.

Strack, F., Martin, L. L., & Schwarz, N. (1988). Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology, 18*(5), 429–442.

Strömbäck, C., Lind, T., Skagerlund, K., Västfjäll, D., & Tinghög, G. (2017). Does self-control predict financial behavior and financial well-being? *Journal of Behavioral and Experimental Finance, 14*, 30–38. https://doi.org/10.1016/j.jbef.2017.04.002.

Tannenbaum, D., Valasek, C. J., Knowles, E. D., & Ditto, P. H. (2013). Incentivizing wellness in the workplace. *Psychological Science, 24*(8), 1512–1522. https://doi.org/10.1177/0956797612474471.

Thaler, R. H. (2000). Mental accounting matters. *Journal of Behavioral Decision Making, 12*(3), 183–206.

Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112*(S1). https://doi.org/10.1086/380085.

Wansink, B., & Cheney, M. M. (2005). Super bowls: serving bowl size and food consumption. *Jama, 293*(14), 1727–1728.

Wertenbroch, K. (1998). Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing Science, 17*(4), 317–337. https://doi.org/10.1287/mksc.17.4.317.

Woolley, K., & Fishbach, A. (2016). For the fun of it: Harnessing immediate rewards to increase persistence in long-term goals. *Journal of Consumer Research, 42*(6), 952–966. https://doi.org/10.1093/jcr/ucv098.

Zuckerman, O., & Gal-Oz, A. (2014). Deconstructing gamification: Evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. *Personal and Ubiquitous Computing, 18*(7), 1705–1719. https://doi.org/10.1007/s00779-014-0783-2.