On or Off Track: How (Broken) Streaks Affect Consumer Decisions

JACKIE SILVERMAN
ALIXANDRA BARASCH

New technologies increasingly enable consumers to track their behaviors over time, making them more aware of their “streaks”—behaviors performed consecutively three or more times—than ever before. Our research explores how these logged streaks affect consumers’ decisions to engage in the same behavior subsequently. In seven studies, we find that intact streaks highlighted via behavioral logs increase consumers’ subsequent engagement in that behavior, relative to when broken streaks are highlighted. Importantly, this effect is independent of *actual* past behavior and depends solely on how that behavior is represented within the log. This is because consumers consider maintaining a logged streak to be a meaningful goal *in and of itself*. In line with this theory, the effect of intact (vs. broken) logged streaks is amplified when consumers attribute a break in the streak to themselves rather than to external factors, and attenuated when consumers can “repair” a broken streak. Our research provides actionable insights for companies seeking to benefit from highlighting consumers’ streaks in various consequential domains (e.g., fitness, learning) without incurring a cost (e.g., reduced engagement or abandonment) when those streaks are broken.

Keywords: streaks, behavioral tracking and logging, technology, goals and motivation, engagement

Jackie Silverman (jasilv@udel.edu) is an assistant professor of marketing at Lerner College of Business and Economics, University of Delaware, 209 Lerner Hall, Newark, DE 19716, USA. Alixandra Barasch (alix.barasch@colorado.edu) is an associate professor of marketing at Leeds School of Business, University of Colorado at Boulder, 995 Regent Drive, Boulder, CO 80309, USA, and a visiting associate professor of marketing at INSEAD, Boulevard de Constance, 77300 Fontainebleau, France. All correspondence can be addressed to Jackie Silverman. This article is based on the first author’s doctoral dissertation. The authors thank Amit Bhattacharjee, Ike Silver, Sydney Scott, Uri Barnea, Rom Schrift, and Deborah Small for helpful comments on earlier versions of this article. They also thank Hazel Hamelin at INSEAD for editing assistance, Carrie Motschwiller at New York University’s Global Walking Challenge for assistance with data collection for study 1, and Shawn Zamechek at the University of Pennsylvania for coding assistance. Partial financial support for this research comes from the Wharton School’s Baker Retail Center and the Wharton Behavioral Lab, both at the University of Pennsylvania. Supplementary materials are included in the web appendix accompanying the online version of this article.

Editors: J. Jeffrey Inman and Stacy Wood

Associate Editor: Leonard Lee

Advance Access publication 30 June 2022

INTRODUCTION

Sustained consumer engagement—that is, repeated interaction with a product or service—is important to a company’s revenues and valuation, particularly in the digital marketplace (Venkatesan and Kumar 2004). In the battle to keep consumers engaged, companies frequently rely on new technological developments and individual consumer data to captivate and retain customers. One increasingly common strategy is to use behavioral logs to highlight consumers’ recent behavior within the company’s platform, capturing when a consumer has repeated the same desired behavior for multiple time periods. Consumers, too, seek technologies that allow them to track and log their behavior across various domains, from health (Fox 2013) to finance (Malcolm 2015). While recent research has examined the general impact of behavioral tracking on different dimensions of consumer wellbeing (Etkin 2016; Karapanos et al. 2016), none has explored how highlighting specific patterns of past behavior through such technology actually affects subsequent behavior.

 • Vol. 00 • 2022
https://doi.org/10.1093/jcr/ucac029

One behavioral pattern made increasingly salient through tracking technology is a *consumer streak*—an unbroken series of three or more consecutive behaviors. For example, language-learning apps like Duolingo highlight the number of consecutive days users have completed a lesson by sending daily reminders and displaying graphics with the length of their streaks. Fitness apps like Peloton feature icons on users' profiles when they exercise multiple days or weeks in a row. Gaming platforms like Pokémon Go and Wordle award badges and send notifications when users have completed multiple actions consecutively (e.g., catching a Pokémon; solving a word game).

Companies presumably highlight streaks via behavioral logs to motivate consumers to continue interacting with the product and ensure the desired behavior persists. Anecdotally, consumers report that these behavioral logs increase their awareness of streaks, which draws them in and keeps them engaged (Lorenz 2017). Yet the effectiveness of this strategy (and when it might backfire) has not been investigated. While displaying streaks may keep consumers engaged when they are on track, consumers will inevitably experience interruptions to their streaks, despite their best intentions. For example, a busy travel schedule may make it impossible to exercise or a bad internet connection may prevent an activity from being logged. Indeed, community forums of many apps document complaints about broken streaks that occurred for reasons outside the user's control (e.g., power outages caused by extreme weather), and some users contact companies directly for help in restoring their streaks (Leskin 2019).

In this article, we investigate the effects of intact versus broken streaks, highlighted through logging, on consumer decisions. Across seven studies, we find that consumers are more likely to engage in a behavior when it contributes to an intact streak in their log, relative to when it follows a broken streak in their log. These effects hold even when consumers' past behavior is held constant, and the only difference is whether that behavior and its contribution to a (broken) streak are displayed within a behavioral log.

Our article makes theoretical contributions to a number of literatures. First, we contribute to the growing area of research on how new technologies are reshaping consumer behavior (Deighton, Goldenberg, and Stephen 2017; Schmitt 2019). In particular, recent research has explored various experiential outcomes for consumers who track their activities, such as decreased enjoyment (Etkin 2016) or enhanced feelings of autonomy (Karapanos et al. 2016), as well as individual traits that predict the likelihood of logging (e.g., increased conscientiousness: Maltseva and Lutz 2018). Building on this work, our research is the first to examine how *highlighting* a specific series of past behaviors—streaks—through tracking technologies can affect consumers' subsequent choices. It thus has practical implications both for consumers seeking to motivate themselves to pursue desirable behaviors and for companies seeking to optimize consumer engagement.

Our research also adds to the literature on consumer goal setting and pursuit (Bagozzi and Dholakia 1999; Locke and Latham 1990). Although prior research has explored how progress toward an end goal (or lack thereof) affects subsequent goal-consistent behavior (Cochran and Tesser 1996; Kivetz, Urminsky, and Zheng 2006; Soman and Cheema 2004), it does not consider how highlighting specific patterns of goal-consistent behavior can become motivating in itself and affect consumers' subsequent decisions. Importantly, while a streak of goal-consistent behaviors may enable progress toward a desirable end goal, we show that even beyond these effects, maintaining a logged streak can become a goal *in and of itself*, and thus independently influence subsequent behavior. Indeed, we find that logged streaks not only affect consumer decisions in prototypical goal-oriented domains (e.g., exercising, learning a language) but also for activities that do not necessarily contribute directly to a quantitative, specific end goal (e.g., playing games).

Theoretical Background

Repeated Behavior. Research on the effects of past behavior on present behavior, notably the tendency for consumers to repeat specific behaviors over time, has yielded mixed findings. On the one hand, consumers seek variety, choosing to vary their experiences as a source of stimulation (Menon and Kahn 1995) or to alleviate boredom (Fishbach, Ratner, and Zhang 2011). This is consistent with theories of decreasing marginal utility: satiated by multiple experiences in a row, consumers are likely to switch behaviors in subsequent time periods (Galak and Redden 2018). They also tend to "balance" (or alternate) attributes when making trade-offs between options (e.g., choosing an expensive but tasty dessert after a cheap but bland entrée: Dhar and Simonson 1999).

On the other hand, people tend to behave consistently with previous decisions in a wide variety of contexts (see Ouellette and Wood 1998 for a review). The likelihood of engaging in the same behavior in the next period increases if the consumer (a) chose that behavior in the *most recent* (immediately preceding) time period and (b) has a high *overall rate*, or frequency, of the past behavior. For example, consumers tend to select products they have just chosen (Dhar, Huber, and Khan 2007), and to abstain from purchase when they have just done so (Tykocinski, Pittman, and Tuttle 1995). Research on brand loyalty also reveals that consumers who bought a product more frequently in the past are more likely to continue buying it in the future (Hoyer 1984; Nunes and Drèze 2006).

Streaks. A streak is defined as a series of at least three repeated, consecutive events or behaviors (Carlson and

Shu 2007).[1] By consecutive we mean that the behavior is repeated across contiguous opportunities or temporal intervals (e.g., once every day) without a "break" (i.e., a missed opportunity) in between.

To date, the majority of related research has focused on how observers use streaks of past events outside their control to forecast future events, particularly in the sports and gambling domains (Gilovich, Vallone, and Tversky 1985). This past work suggests that people over-interpret noise within small samples, seeing random streaks of past events (e.g., getting heads for three consecutive coin flips) as meaningful patterns that shape their subsequent predictions. However, the question of how consumers' *own* streaks of freely chosen behaviors affect their subsequent decisions about whether to keep engaging in those behaviors remains unexplored, as are the psychological mechanisms involved. This is surprising given the ubiquity of streaks in consumers' lives—almost everyone exhibits streaks of behavior at some point, whether purposefully or inadvertently.

More importantly, as companies increasingly deploy technology to track and highlight consumers' behavioral streaks, their salience and potential impact on consumer decision-making grow ever more prevalent. The market is flooded with apps to help users quantify their behaviors and highlight their streaks (Nield 2019), and the proliferation of tracking technology has been widely discussed in the popular press (Austin 2019; Fox 2013).

In the current research, we seek to answer the question of how consumers' awareness of their own streaks affects their subsequent choices. That is, we examine how the presence of an intact versus broken streak, as highlighted by logging, impacts the consumer's decision to engage in the same behavior moving forward. We also investigate the psychological process underlying the influence of logged streaks on consumer behavior.

The Current Research

We propose that consumers consider maintaining a logged streak to be a meaningful goal in and of itself. Findings from several literatures support the notion that people value streaks and seek to preserve them. For instance, streaks may be valued as indicators of consistency, which can engender feelings of cognitive balance, promote psychological well-being, and serve as a source of motivation and positive reinforcement (Singer 1966; Spence 1956). The literature on self-concept suggests that even in the absence of social feedback, people regard behavioral consistency as a goal, with implications for the way they perceive themselves and their likelihood of future success (Markus and Wurf 1987; Reed et al. 2012). Streaks may also satisfy an inherent preference for order, uniformity, and completeness (Barasz et al. 2017; Evers, Inbar, and Zeelenberg 2014), which can drive consumers to collect complete sets of products and experiences (Belk 1988; Gao, Huang, and Simonson 2014; Keinan and Kivetz 2011). Being aware of their own "sets" of repeated behaviors may thereby allow consumers to derive greater meaning and psychological utility from them, with motivational implications for their future behavior.

We further suggest that the *symbolic representation* of streaks (i.e., behavioral logs) plays a crucial role in determining consumers' awareness of their streaks. Cues that shift attention to certain aspects of the environment—by making them more visible or concrete—may enhance how they are processed, encoded, and remembered (Liberman, Trope, and Stephan 2007; Schneider and Shiffrin 1977; Taylor and Fiske 1978). For example, logging meals on a weight-loss app makes consumers more conscious of their recent food intake (Zepeda and Deal 2008), and wearing a fitness-tracking device directs greater attention to recent physical activity (Sjöklint, Constantiou, and Trier 2015). Some work also suggests that the symbols used to represent streaks within behavioral logs (e.g., checkmarks, stars, badges) can come to be valued in and of themselves. For instance, behavioral feedback can be inherently reinforcing, even independent of the information it provides (Hsee, Yang, and Ruan 2015; Locke, Cartledge, and Koeppel 1968; Verplanck 1956). Consumers may thereby seek to accumulate symbolic media (e.g., airline points), engage in gamified behaviors even when they offer no practical utility (Hamari, Koivisto, and Sarsa 2014; Hsee et al. 2003), and work harder to reach end goals when reward programs endow them with even an illusory sense of progress (Kivetz et al. 2006). Accordingly, by highlighting intact (vs. broken) streaks, behavioral logs may elevate their importance and impact on subsequent behavior.

Logged Streaks as Goals. Given the benefits of behavioral consistency and the reinforcing nature of behavioral logs, we posit that consumers will value their logged streaks and adopt a goal of maintaining them, thus affecting their subsequent decisions. Goals exert a profound influence on consumers' self-concept, feelings of self-efficacy, and motivation (Hollenbeck, Williams, and Klein 1989; Latham and Locke 2006; Schunk 1989). Behaviors that contribute to goal progress generate feelings of pride and accomplishment (Fredrickson 2001; Scott and Nowlis 2013) and instill confidence in one's abilities (Atkinson 1957; Bandura 1977; McClelland 1961). Thus, if

1 Other research supports the idea that three behaviors (or items) are often a "tipping point" when consumers start to gain meaning from a set (e.g., consumer collections: Gao et al. 2014). Furthermore, in a separate study (study S1; see web appendix), we examined how streaks are perceived and defined in consumer behavior contexts by asking participants to report the "streakiness" of 25 different patterns of behaviors logged on an app. Consistent with prior findings in other contexts, participants viewed consumer behaviors as streaky (i.e., significantly higher than the scale midpoint) whenever there are at least three behaviors in a row without a recent "miss" ($ps < .001$).

consumers view maintaining a streak of logged behavior as a goal in and of itself, an intact streak should increase their sense of accomplishment and motivate them to continue engaging in that behavior. By the same token, a broken streak should have the opposite effect on their sense of accomplishment and subsequent behavior. Encountering failure during goal pursuit triggers feelings of discouragement (Heath, Larrick, and Wu 1999) and undermines belief in one's ability to ultimately reach that goal (Bandura and Locke 2003; Latham and Locke 2006). Consequently, failure is demotivating and reduces the likelihood of future goal-consistent behaviors (Soman and Cheema 2004). In sum, behavioral logs (and the streaks highlighted therein) may be a double-edged sword, increasing subsequent behavioral engagement when a streak is intact but decreasing engagement when a streak is broken.

Notably, our theory predicts that subsequent decisions will be influenced by logged streaks *themselves*, even when holding *actual* behavior constant. In other words, simply representing the exact same behavior as contributing to an intact streak (vs. following a broken streak) within a behavioral log will affect consumers' actual future behavior. In this research, we test the following hypotheses:

> **H1:** Consumers will be more likely to engage in a target behavior when it contributes to an intact streak in their behavioral log than when the same behavior follows a broken streak in their behavioral log.
>
> **H2:** This effect will be driven by consumers' goal of maintaining logged streaks.

Our theory allows us to make two notable ancillary predictions about moderating factors that will increase or decrease the effects of logged streaks on subsequent behavior. First, the effect will be *magnified* when consumers attribute the break to themselves compared to when they attribute the break to an external factor. Prior research has shown that when individuals feel responsible for goal failure, they experience lower self-efficacy and are less likely to continue pursuing the goal (Bandura and Locke 2003; Tolli and Schmidt 2008). Relatedly, individuals often try to attribute responsibility for goal failure to others, as self-attribution intensifies the negative consequences on their self-image (Steele 1988; Zuckerman 1979). Thus, to the extent that consumers view maintaining a logged streak as a goal, we predict that when they feel responsible for a broken streak they will be even *less* likely to continue the streak-related behavior.

> **H3:** The effect of intact (vs. broken) logged streaks will be amplified when consumers attribute responsibility for the break to themselves as opposed to external factors.

Another ancillary prediction that follows from our theory is that the effect of logged streaks will be *diminished* when consumers have the opportunity to "repair" streaks within their behavioral log. Prior work suggests that when consumers engage in goal-inconsistent behaviors, they seek to preserve their self-concept by justifying the behavior in some way, for example by reframing inconsistent behavior as an exception (Mazar, Amir, and Ariely 2008). Similarly, allowing a mental "reset" of progress after goal failure can promote subsequent goal pursuit (e.g., with the start of a new month: Dai, Milkman, and Riis 2014). Moreover, replacing the symbols of broken streaks with symbols that represent streak-contributing behaviors should reinforce continued engagement. Accordingly, we predict that providing consumers with the opportunity to repair broken streaks in their behavioral logs (e.g., by allowing certain actions to "fill in" the break) should reactivate the streak goal, thus increasing subsequent engagement in the streak-related behavior.

> **H4:** The effect of intact (vs. broken) logged streaks will be attenuated when consumers have the opportunity to repair broken streaks within the behavioral log.

To summarize, we propose that maintaining an uninterrupted series of past behaviors (i.e., a streak) in a consumer's behavioral log becomes a goal in and of itself, and thus affects subsequent decisions. Unlike prior research that considered the effects of goal progress and failure with respect to a salient end state (Kivetz et al. 2006; Soman and Cheema 2004), we suggest that a logged streak of behavior is not simply a means to an end, but an end in itself that independently influences consumer behavior.

Study Overview. We test our hypotheses in seven studies (as well as a pilot study and six additional studies in our web appendix) using a multimethod approach. First, through descriptive pilot data, we document the pervasiveness of apps that log or track consumer streaks across a variety of domains. Then, we examine how logged streaks (i.e., whether they are intact or broken) affect the decision to engage in a behavior in the future (hypothesis 1), both in the field with fitness app users' actual step count data (study 1), as well as in controlled lab settings where participants use apps that track either common goal-oriented behaviors (exercising in study 2; language learning in studies 3 and 4) or behaviors that are not necessarily associated with the pursuit of specified end goals (playing games in studies 5, 6, and 7). All studies test the effects of logged streaks on real (i.e., not hypothetical) consumer decisions with consequences for participants (e.g., how they spend their time). Table 1 summarizes the key attributes of all studies reported in the main text.

Across these studies, we test our proposed mechanism (hypothesis 2) in multiple ways. First, we directly measure the extent to which participants explicitly report that maintaining their logged streak is a goal they are pursuing, and find that it mediates the effect of an intact versus broken logged streak on subsequent behavior (study 4). Second, we demonstrate that an intact (vs. broken) logged streak

TABLE 1

KEY ATTRIBUTES OF STUDIES 1–7

Study	Domain	Design	Broken streak operationalization	DV
1	Fitness (step counts)	Correlational field data	App users did not meet the fitness challenge of 7,000+ steps on a given day	Meet the fitness challenge on a given day or not
2	Fitness (strength exercises)	Two conditions (*logged streak: intact* or *broken*) between subjects	Completed exercise cannot be added to log	Continue with strength exercises or switch to a stretching exercise
3	Language learning	2 (*logged streak: intact* or *broken*) × 2 (*behavioral log: present* or *absent*) between-subjects design	Quota message prevents completion of a Portuguese question	Continue with a Portuguese question or switch to a Hawaiian question
4	Language learning	Three conditions (*intact logged streak*, *broken logged streak*, and *no log*) between subjects	Completed language question cannot be added to log	Continue with a language question or switch to viewing a funny video
5	Word and number games	Three conditions (*intact logged streak*, *broken logged streak*, and *no log*) between subjects	Only one category of game counts toward logged streak	Continue or stop playing games
6	Word and number games	Three conditions (*intact streak*, *externally-attributed broken streak*, and *self-attributed broken streak*) between subjects	*Externally-attributed:* quota message prevents completion of a number game *Self-attributed:* unable to complete a difficult number game	Continue with a number game or switch to a word game
7	Word and number games	Three conditions (*intact streak*, *broken streak*, and *repairable broken streak*) between subjects	Quota message prevents completion of a game	Continue with initial game type or switch to the other game type

affects several key indicators of goal progress, including participants' sense of accomplishment and emotions (studies 4, 5, and 7). Third, we support our proposed mechanism through two theoretically derived moderation tests, which show that the focal effect is amplified when participants feel responsible for a broken streak in their behavioral log (compared to when they attribute that break to an external factor; hypothesis 3; study 6), and is attenuated when participants have the opportunity to repair a broken streak in their behavioral log (hypothesis 4; study 7).

All experimental studies involved sample sizes between 400 and 600 participants. These were determined in advance to provide at least 80% power to detect the focal effect in each study, based on preliminary effect size estimates from pilot studies. For our field study (study 1), the sample size was determined by the number of users who enrolled in the fitness challenge. We report all measures assessed, and no conditions or participants were dropped from any of the analyses. Studies 4, 5, 6, and 7 were preregistered. All data, materials, and preregistrations, as well as our web appendix, can be found in our Open Science Framework (OSF) repository: https://osf.io/kpjh9/ (Last Accessed July 7, 2022).

CONSUMER LOGGING PILOT STUDY

To investigate just how pervasive behavioral logging has become for consumers, we collected descriptive data from 100 participants on MTurk ($M_{\text{age}} = 31.86$ years, 41.00% female). We asked them several questions about their real-life usage of and attitudes toward apps that track and highlight their patterns of behavior.

First, we asked participants to list any apps that they currently used that highlight streaks. Together, they named 101 unique apps (total listed: $N = 202$; per person: $M = 2.02$, SD = 1.52), which fell into several distinct categories. The largest portion of participants (48%) named gaming apps, followed by social media/messaging apps (40%), health and fitness apps (39%), and language-learning apps (9%); 13% named other types of apps, including apps focused on reading (e.g., Bible verses: 6%) and financial planning (4%).

Next, we asked participants how these apps influenced their personal experiences and behaviors. Fifty-nine percent of participants reported that they had "gone out of their way" to maintain (or avoid breaking) their streaks on an app, and 27% said they had specifically engaged in behavior "outside of the app" (i.e., offline) to maintain their logged streak. For example, three participants described how they set an alarm on their phone to make sure that they did not break their streak, and one participant wrote that he "didn't want to work out . . . but did a quick 7-minute workout in the back of a bar to make sure [his] streak remained intact" on his fitness app. We then asked participants to respond to several statements about how they felt when apps notified them of having a streak (from 1 "Strongly disagree" to 7 "Strongly agree"). They reported that they liked the fact that apps highlighted their streaks

($M = 5.10$, SD $= 1.86$; vs. scale midpoint: $t(99) = 5.91$, $p < .001$, $d = 0.59$), and that seeing information about their streaks increased their awareness of their streaks, enhanced their overall experience, and motivated them to keep using the app (Ms > 4.90, ts > 5.00, ps $< .001$, ds > 0.50).[2]

In sum, this pilot study reveals the prevalence of apps that highlight streaks to consumers, which underscores the timely nature of our investigation. Moreover, consumers reported that these logged streaks affected their behaviors both within apps and offline, and that highlighting streaks improved their experiences with apps and were a source of motivation. To further explore the influence of this technology, we conducted a field study to examine how streaks displayed within an app can affect consumers' actual step counts.

STUDY 1: THE EFFECT OF STREAKS ON STEP COUNTS IN THE FIELD

To establish the effects of tracking technology and streaks in a real-world setting, we conducted a study in collaboration with a step-counting fitness app that notifies users of streaks in reaching a daily step goal. We chose to examine consumer fitness because it is a consequential consumer behavior often tracked via apps and other devices; indeed, 39% of participants in our pilot study reported using health and fitness tracking apps, and 21% of American adults report regularly using a wearable fitness tracker (Vogels 2020). In this study, we observed participants' recent behavior logged in the app—in particular, whether they had an intact or broken streak of daily steps—and used it to predict their subsequent stepping behavior.

Methods

For this study, we partnered with a university's wellness program that runs an annual 30-day fitness challenge encouraging university employees to walk at least 7,000 steps a day. To participate, employees tracked their daily steps (e.g., with a FitBit) and synced their step tracker with a third-party app. The app's interface showed users a variety of information about their physical activity, including their step count in real time. Like many other apps commonly used by consumers, it also highlighted users' current streak of "meeting the challenge" via an in-app behavioral log only visible to the user which informed them of how many days in a row they had walked at least 7,000 steps (see web appendix for a sample screenshot).

Our study took place in the fall of 2018. In the month prior to the start date, all university employees across four campuses could sign up online and download the corresponding app to register for the challenge. During registration, they reported some basic information of interest to the university (see web appendix for descriptive statistics). Ultimately, 980 employees registered and participated in at least 1 day of the challenge. We did not exclude any individuals who may have decided to stop participating after the challenge began. Through our partnership with the wellness program, we obtained a daily step count for each individual throughout the challenge. We used meeting or exceeding the 7,000-step threshold (vs. not meeting it) as the target behavior in this study.

Results

We ran several binary logistic models to examine the effects of individuals' past logged behavior on their subsequent behavior—specifically, whether the individual completed the target behavior on a given day (coded as "1" when they walked 7,000 or more steps and "0" when they walked fewer than 7,000 steps). Due to the repeated measures nature of our data, the models clustered standard errors by individual, thus accounting for nonindependence of observations. The models also included a variable for each individual's overall rate of meeting the challenge (i.e., a value from 0 to 1 representing the percentage of days the individual met the challenge), allowing us to control for the possibility that users who had a higher rate of meeting the challenge may also have been more likely to exhibit streaks.

In our primary model, we tested the effect of an intact versus broken streak on subsequent behavior via a contrast variable, coded as "1" when the individual had an intact streak (i.e., they had met the challenge for three or more days in a row, inclusive of the previous day), "−1" when they had a broken streak (i.e., they had not met the challenge on the previous day, after doing so for three or more days prior) and "0" when they had neither pattern of behavior. The model revealed that individuals were more likely to engage in the target behavior (i.e., meet the challenge) on a given day when they had an intact streak compared to when they had a broken streak ($b = 0.38$, standard error [SE] $= .03$; $Z = 11.27$, $p < .001$). Additional models using dummy variables to investigate the independent effects of intact streaks and broken streaks also found significant effects (*intact*: $b = 0.25$, SE $= .05$; $Z = 5.06$, $p < .001$; *broken*: $b = -1.01$, SE $= .07$, $Z = -14.94$, $p < .001$). Our web appendix includes several robustness check models with additional controls (e.g., day effects, state dependence) and using various streak lengths, all of which replicate these effects.

2 Participants did not report feeling overwhelmed, bothered, or distracted by such information, or that it took away from their experience in using the app (Ms < 2.60, ts > 7.75, ps $< .001$, ds > 0.75).

Discussion

This field study documents the effect of streaks in a real-life setting with consequential outcomes for consumers. Fitness app users were more likely to meet a step-count challenge on a given day when doing so contributed to their logged streak, compared to when it followed a broken logged streak. These results provide initial support for our primary prediction (hypothesis 1): users of an app that highlighted their own freely chosen sequences of behavior exhibited stark differences in behavioral engagement following intact versus broken streaks. However, while allowing for self-selection enhances this study's ecological validity and realism, it limits our ability to isolate the causal effects of users' logged streaks *themselves* on subsequent decisions. For instance, the correlational nature of this study prevents us from separating the effects of streaks from factors that may have made certain people more likely to both have an intact streak and to log their behavior on a given day (e.g., their enjoyment of the app). Accordingly, our subsequent studies experimentally manipulated participants' streaks within a behavioral log to establish causal effects on subsequent behavior, and thus serve as a crucial complement to this field evidence.

STUDY 2: THE EFFECT OF LOGGED STREAKS ON EXERCISE

To isolate the effects of logged streaks in a controlled setting, in this study, we simply manipulated whether the same behavior was successfully or unsuccessfully logged in an app (and thus *appeared* as an intact or broken streak in a behavioral log). While our goal in this and subsequent studies was to isolate the causal effects of logged streaks on subsequent behavior, we also wanted to ensure that our paradigm maintained high external validity. Therefore, we developed an app that mirrored consumers' real-world experiences, which included a "tracker" displaying participants' logged behaviors (e.g., completed exercises). This app could be integrated into our survey, with a customizable background, color scheme, and content that made it distinct from the rest of the study (see materials for screenshots).

In this study, we again examined the effects of logged streaks within the consequential domain of consumer fitness. Here, participants engaged in the same sequence of real exercises developed by a fitness instructor (Rogers 2020). We chose these exercises because they are meant to be performed at a desk or on a couch, making them accessible to many adults who often face difficulties in finding time or energy to exercise (Schmall 2019).

Methods

We recruited 601 participants from MTurk ($M_{age} = 38.63$, 44.93% female). Participants learned that they would be testing a fitness app under development, which provided instructions and diagrams from a fitness instructor to guide them through several strength exercises (see OSF for materials). Participants were informed that they would log their exercises as they progressed through them by typing in the name of each exercise as soon as they completed it (as with common fitness apps like MyFitnessPal). A checkmark then appeared for each completed exercise on the behavioral log featured at the top of the app. Participants were told that only strength exercises could be recorded on their behavioral log.

Next, all participants did the same three strength exercises (chest squeezes, core twists, and leg scissors) and successfully logged them on the app—hence everyone had a streak of three checkmarks on their behavioral log. They then completed a fourth strength exercise (sit stands). Participants were randomly assigned to one of two between-subjects conditions. Those in the *intact logged streak* condition successfully logged their fourth exercise. Those in the *broken logged streak* condition, however, were informed that the behavioral log was unable to load correctly after their fourth exercise (but that this error would not occur again) and thus saw an "X" on their behavioral log.[3] This design thus controlled for participants' *actual* behavior (i.e., everyone completed the same four exercises), and manipulated only what participants viewed in their behavioral log (i.e., whether the *logged* streak representing that series of behaviors was intact or broken; figure 1).

Participants then chose their next activity within the app: they could either continue with another strength exercise (i.e., the target behavior) or switch to a stretching exercise instead. This decision served as our key dependent variable, allowing us to test whether logged *representations* of behavior could affect subsequent *actual* behavior.

Before continuing with the exercise of their choice, participants answered two questions about their personal fitness level (see materials for all questions). Participants also reported how many exercises they actually did during the study; the majority of participants (76.71%) reported completing all four exercises, while only 1.83% said that they did none of them. Finally, participants completed their chosen exercise and answered demographic questions.

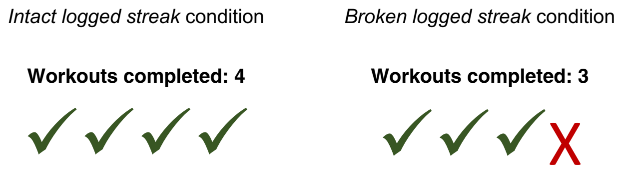
Results

A chi-square test revealed that more participants engaged in the target behavior (another strength exercise)

3 In an additional preregistered study (study S4 in the web appendix), the broken logged streak was conveyed without any graphic; participants were simply informed if they had a broken streak in their log.

FIGURE 1

SCREENSHOTS OF HOW INTACT VERSUS BROKEN STREAKS WERE PORTRAYED WITHIN THE BEHAVIORAL LOG IN STUDY 2



Note: A color version of this figure appears in the online version of this article.

when they had an intact streak in their behavioral log (66.23%) than when they had a broken streak in their behavioral log (57.86%; X^2 (1) = 4.47, p = .035; odds ratio [OR] = 1.43).

Discussion

Study 2 found that participants were more likely to continue engaging in real exercise behavior when their behavioral log showed an intact streak versus a broken streak. Importantly, this effect occurred even when their *actual* behavior was identical across conditions. That is, even though all participants had completed four consecutive strength exercises in reality, simply having that same series of behaviors framed as an intact versus broken streak within the app's behavioral log influenced their subsequent *real* behavior outside the app. Notably, by holding actual behavior constant, this study controlled for any potential unobserved effects that might result from maintaining versus breaking a streak by engaging in *different* behaviors (e.g., feelings of boredom or expertise; perceptions of one's own liking or motivation; habit formation or automaticity). These findings thus provide an initial demonstration of the importance and causal impact of *logged* streaks themselves.

STUDY 3: EXAMINING THE EFFECTS OF LOGGED STREAKS IN BOTH DIRECTIONS

Building on the previous study, study 3 used a different approach to further investigate the effects of logged streaks on consumers' subsequent choices. Moreover, we also tested whether the positive effects of intact logged streaks could be distinguished from the negative effects of broken logged streaks. To achieve these objectives, we directly manipulated participants' series of behaviors and examined how the effect of intact versus broken streaks on subsequent behavior was affected by the mere presence (vs. absence) of a behavioral log. In other words, we tested whether logging *itself* could amplify the effects of the same series of actual behaviors, and do so in *both* directions. We expected that participants would be more likely to subsequently engage in the target behavior when their intact streak was highlighted by a behavioral log within an app (vs. not), but *less* likely to engage in that target behavior when their *broken* streak was highlighted (vs. not).

Additionally, this study examined another consequential consumer behavior: language learning. Like exercise apps, language-learning apps have grown increasingly popular (the most widely used app, Duolingo, has over 500 million registered users; de León 2020) and commonly log consumers' patterns of behavior (see Consumer Logging Pilot Study section).

Methods

We recruited 602 MTurk participants (M_{age} = 36.03, 51.16% female). Participants learned that they would be testing Portuguese and/or Hawaiian language-learning questions for an app under development, and read basic information about the app, which was modeled after Duolingo. All participants were also told that at some point, they might see a "quota" message when enough participants had already tested out a specific language-learning question, and that this would mean that they were not needed to test that particular question. The instructions emphasized that the quota message pertained to a specific question, not the availability of all questions in general. Participants were also informed that this message was not an indication of their abilities and that they would see this message at most once during the study.[4] Participants were informed they would start with the Portuguese module, then proceeded to completing the language-learning questions.

Participants were then randomly assigned to condition in a 2 (*streak: intact* or *broken*) by 2 (*behavioral log: absent* or *present*) between-subjects design. In the *intact streak* condition, participants completed four Portuguese questions in a row. In the *broken streak* condition, participants completed three questions in a row and then saw the quota message in place of the fourth question. We also manipulated whether this series of behaviors was highlighted to participants via a behavioral log within the app. In the *behavioral log absent* condition, participants did not see a behavioral log at all; rather, they completed these Portuguese questions in the typical Qualtrics survey format. In the *behavioral log present* condition, participants completed the language-learning questions within an app that highlighted their recent series of behaviors, similar to study 2. These

4 A post-test (N = 374) confirmed that the quota message did not have any effect on participants' beliefs about how functional or likeable the app was (ts < 0.50, ps > 0.65, ds < 0.12; see web appendix).

participants were informed that this app would show them their progress in completing questions, as many apps typically do to help motivate users in practice. For every Portuguese question completed, a checkmark would be added to their behavioral log featured at the top of the screen. Participants logged a question by clicking an additional button on the app after completing it. In addition, within the *behavioral log present* condition, participants in the *intact logged streak* condition saw a notification that they had a streak in their behavioral log, while participants in the *broken logged streak* condition were informed that, due to the quota, they had a broken streak in their behavioral log. Importantly, these visual symbols and notifications within the behavioral log were modeled on real-life apps like Duolingo which highlight consumers' past behavior in similar ways.

All participants then chose which language they would like to learn for their next question. They could either continue with the target behavior (i.e., learning Portuguese) or switch to learning Hawaiian. This choice served as our primary-dependent variable. Then participants completed this language question and demographics.

Results

We ran a binary logit with *streak* condition, *behavioral log* condition, and their interaction as factors. This model revealed a main effect of *streak* condition ($F(1, 598) = 49.90$, $p < .001$, OR = 3.17): Participants were more likely to engage in the target behavior (i.e., choose Portuguese) when they had an intact streak (78.33%) than when they had a broken streak (53.31%). There was no main effect of *behavioral log* condition ($F(1, 598) = 2.66$, $p = .104$). More importantly, we found the expected significant interaction between the *streak* and *behavioral log* conditions ($F(1, 598) = 35.02$, $p < .001$; figure 2).

Separate chi-square tests revealed that highlighting participants' behavior through the behavioral log had an effect both when the participant broke their streak and when they maintained it, but in opposite directions. Specifically, participants with an intact streak were *more* likely to engage in the target behavior when their recent streak was highlighted via the log (92.47%) compared to when it was not (64.94%; X^2 (1) = 33.47, $p < .001$, OR = 6.63). However, participants with a broken streak were *less* likely to engage in the target behavior when their broken streak was highlighted via the log (45.21%) compared to when it was not (60.90%; X^2 (1) = 7.46, $p = .006$, OR = 0.53).

Discussion

This study provides convergent evidence that simply *logging* a series of behaviors can itself affect consumers' subsequent decision to engage in that behavior. We documented this effect in the context of a language-learning app, where consumers are accustomed to seeing their behavioral streaks visualized in behavioral logs, providing initial support for the generalizability and persistence of these effects.

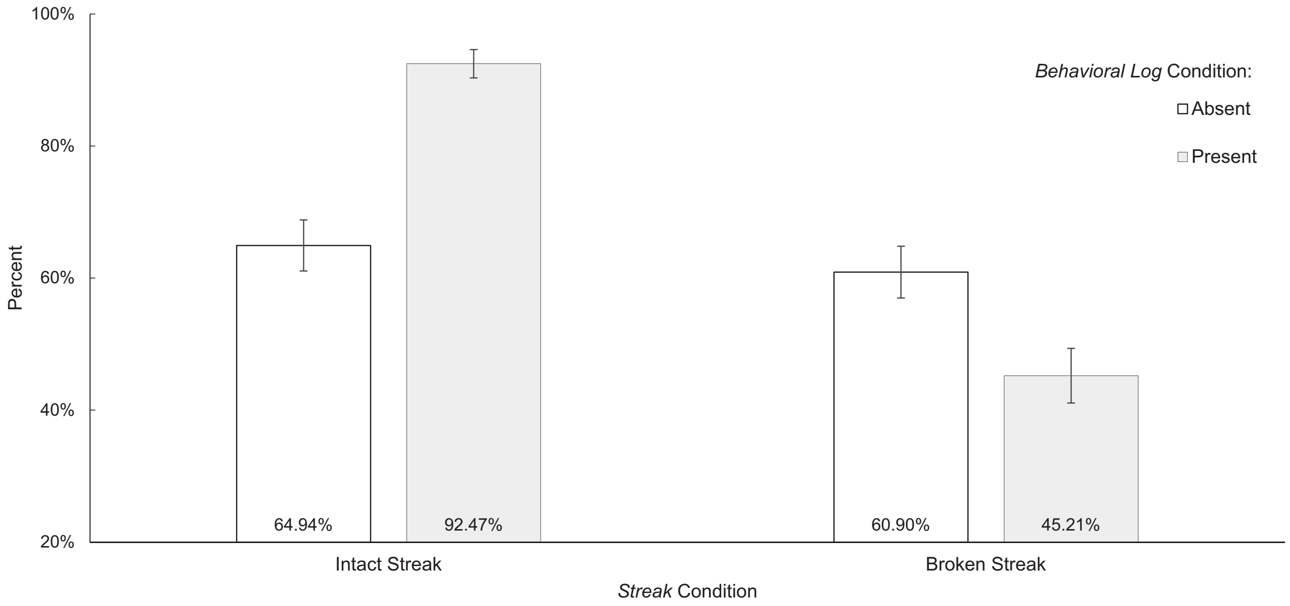
Interestingly, these findings also suggest that intact and broken logged streaks may exert *separate* effects, independently shaping downstream decisions in opposite directions (compared to the same behaviors in the absence of a behavioral log). Specifically, holding actual behavior constant, participants were significantly more likely to engage in the target behavior when they had an intact streak highlighted by their behavioral log (vs. when it was not highlighted), but were significantly less likely to engage in the target behavior when they had a broken streak highlighted by their behavioral log (vs. when it was not highlighted). We continue to examine these contrasts in the next two studies, which also have the added benefits of controlling for both participants' actual series of behaviors and their use of the app interface across all conditions.

Notably, this pattern of results is inconsistent with a potential alternative explanation: that breaking a streak due to a quota message or a logging failure (in study 2) triggered negative inferences about the app, thus affecting subsequent behavior. Because participants in both *broken streak* conditions received the same quota message in this study, any such inferences are constant across conditions and cannot explain the negative effect of simply highlighting that broken streak within a behavioral log.

To further isolate the effect of a broken streak from such potential inferences and to examine the effects of intact and broken logged streaks separately, we ran two preregistered follow-up studies (S2a and S2b; see web appendix). In both, we controlled for participants' most recent logged behavior within the app, such that all participants either successfully engaged in and logged the target behavior (study S2a, $N = 156$) or experienced a "miss" in their behavioral log due to a quota (study S2b, $N = 218$). We simply manipulated whether this most recent behavior was *preceded* by a streak. This allowed us to test the effect of an intact logged streak above and beyond any positive effects of successfully logging the most recent behavior (study S2a in the web appendix), as well as the effect of a broken logged streak above and beyond any negative effect of being unable to log the most recent behavior (study S2b in the web appendix). Consistent with our theory, in study S2a (web appendix), more participants engaged in the target behavior when it contributed to an intact streak (69.95%) than when it did not (48.72%; X^2 (1) = 5.93, $p = .015$; OR = 2.23), while in study S2b (web appendix), fewer participants engaged in the target behavior when it followed a broken logged streak (37.61%) than when it followed a single miss without a preceding streak (64.22%; X^2 (1) = 14.39, $p < .001$, OR = 0.34).

FIGURE 2

PERCENTAGE OF PARTICIPANTS IN STUDY 3 WHO ENGAGED IN THE TARGET BEHAVIOR AS A FUNCTION OF HAVING AN INTACT VERSUS BROKEN STREAK, AND WHETHER THAT PATTERN OF BEHAVIOR WAS HIGHLIGHTED IN A BEHAVIORAL LOG (VS. NOT). ERROR BARS REPRESENT 95% CONFIDENCE INTERVALS



STUDY 4: TESTING THE MECHANISM FOR THE EFFECT OF LOGGED STREAKS

The primary purpose of study 4 was to directly test the proposed mechanism underlying the effects documented in our earlier studies. Specifically, we measured the extent to which participants explicitly reported maintaining a logged streak as an active goal, as well as their reported sense of accomplishment (i.e., an indicator of goal progress: Bandura and Locke 2003; Scott and Nowlis 2013). We tested whether these factors mediated the effect of logged streaks on subsequent behavior. In addition, we changed our key behavioral measure to assess whether participants were willing to sacrifice the opportunity to engage in a more enjoyable activity (watching a fun video clip) to keep their logged streaks intact. If, as we posit, consumers view streak maintenance as a meaningful goal in itself, they should be more willing to continue engaging in the target behavior even when doing so entails forgoing a more hedonically appealing outside option. Accordingly, these salient opportunity costs make this study a strong test of our predictions (Spiller 2011).

This study again implemented the approach of study 2, whereby all participants engaged in the exact same series of actual behaviors, and the only difference between the *intact* and *broken logged streak* conditions was what was displayed in their behavioral logs. Building on study 3, we also included a control condition where participants engaged in these behaviors within the same app interface but were not provided with a behavioral log at all. We expected this *no-log* condition to fall between the two *logged* conditions, such that participants with an intact logged streak would engage in the target behavior at a higher rate, while participants with a broken logged streak would do so at a lower rate.

Methods

We recruited 601 participants from MTurk ($M_{\text{age}} = 37.08$, 37.44% female). Like in study 3, all participants were told that they would be testing out vocabulary questions for a language-learning app. After reading the instructions, all participants proceeded to the app and completed four Portuguese vocabulary questions.

Participants were randomly assigned to one of three between-subjects conditions. Similar to studies 2 and 3, participants in the *intact logged streak* and *broken logged streak* conditions learned that for every Portuguese question they completed, a checkmark would be added to their behavioral log featured at the top of the screen. Participants logged questions by clicking an additional button on the app. Participants in the *intact logged streak* condition successfully logged all four questions and saw a notification that they had a streak on their behavioral log.

Participants in the *broken logged streak* condition, however, were informed that the log was unable to load correctly after their fourth question (but that this error would not occur again). Because they could not log their fourth question, they saw an "X" on their log and were informed that they had a broken streak in their behavioral log. Participants in the *no-log* condition completed the same four Portuguese questions on a version of the app that lacked the behavioral log feature but was otherwise identical.

We were interested in whether consumers would be willing to sacrifice the opportunity to engage in a more enjoyable activity to keep their streaks intact. To test this, our key dependent variable was participants' choice between continuing with another Portuguese question on the app (the target behavior) or switching to watching a short, fun video of a kitten and dog playing.[5] Before continuing with their chosen activity, participants answered several questions about the thoughts and feelings they experienced while making this choice, which were intended to measure our proposed process (all from 1 "Not at all" to 11 "Extremely" or "A great deal"). First, they completed an item directly measuring our mechanism: "To what extent did you have a goal of maintaining a streak of completing language-learning questions?" Second, they completed several measures capturing additional indicators of goal progress drawn from prior literature showing that successful goal pursuit typically leads to a greater sense of achievement and increased motivation to engage in goal-consistent behaviors (Bandura 1977; Hollenbeck et al. 1989; Scott and Nowlis 2013), whereas goal failure precipitates discouragement and demotivation (Bandura and Locke 2003; Heath et al. 1999). Accordingly, participants answered five items regarding their *feelings of achievement* (e.g., "How much did you feel like you achieved something?"; Scott and Nowlis 2013) and five items regarding their *motivation* to continue the target behavior (e.g., "How determined were you to continue completing language-learning questions?"; αs > .95; see web appendix and materials for all items). As preregistered, we combined all 10 items into one measure of *sense of accomplishment*, but the results were similar when these subscales were analyzed separately (see web appendix for analyses with these subscales for all relevant studies, including serial mediation).

Participants then answered three exploratory questions regarding additional behavioral intentions that might be influenced by logged streaks. First, they reported how likely they would be to continue using the app and to recommend it to a friend in the future (both from 1 "Extremely unlikely" to 11 "Extremely likely"). Additionally, participants in the two *logged streak* conditions reported the extent to which they would like to see the behavioral log if they were to use the app in the future (from 1 "Definitely no" to 11 "Definitely yes"). Finally, participants completed their chosen activity and answered demographics.

Results

Target Behavior. A logit model revealed a significant effect of condition on subsequent behavior ($F(2, 598) = 13.08$, $p = .001$; figure 3). As expected, more participants in the *intact logged streak* condition engaged in the target behavior (65.84%) than in the *broken logged streak* condition (47.98%; X^2 (1) = 13.02, $p < .001$; OR = 2.09). Additionally, the *no-log* condition fell between the other two conditions; significantly more participants in the *intact logged streak* condition engaged in the target behavior than the *no-log* condition (54.73%; X^2 (1) = 5.20, $p = .023$; OR = 1.59), while directionally fewer participants in the *broken logged streak* condition engaged in the target behavior compared to the *no-log* condition (X^2 (1) = 1.82, $p = .178$; OR = 0.76).

Streak Maintenance Goal. A one-way ANOVA revealed a significant effect of condition on the extent to which participants explicitly reported that maintaining their streak of engaging in the logged target behavior was an active goal ($F(2, 598) = 3.35$, $p = .036$). An independent *t*-test revealed that participants with an intact logged streak reported adopting this goal to a greater extent ($M = 7.24$, SD = 3.38) than participants with a broken logged streak ($M = 6.41$, SD = 3.22, $t(398) = 2.52$, $p = .012$, $d = 0.25$).[6] The *no-log* condition again fell between these two conditions; participants reported directionally lower goal adoption than those in the *intact logged streak* condition ($M = 6.72$, SD = 3.17; $t(401) = 1.60$, $p = .111$; $d = 0.16$), and directionally greater goal adoption than those in the *broken logged streak* condition ($t(397) = 0.98$, $p = .33$; $d = 0.10$).

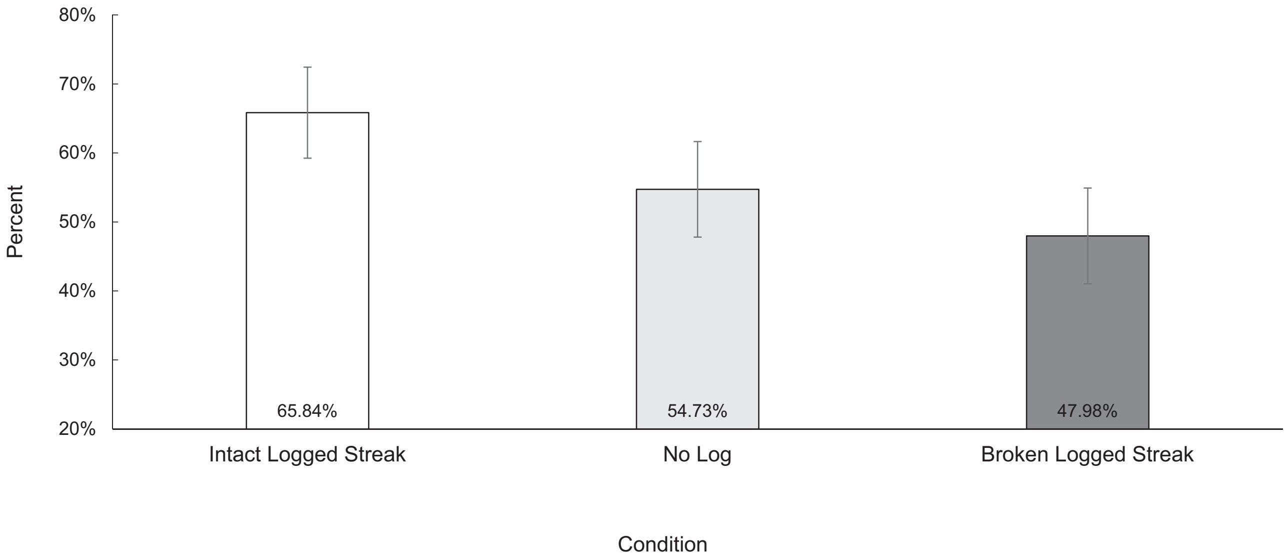
Indicator of Goal Pursuit. Results were consistent for participants' reported sense of accomplishment ($F(2, 598)$

5 In a post-test at the end of the survey, we confirmed that participants indeed preferred watching fun videos over completing language-learning questions (when asked which they liked more, from 1 "Definitely language-learning" to 11 "Definitely fun videos"; M = 6.40, SD = 3.32, vs. scale midpoint: $t(600) = 2.95$, $p = .003$, d = 0.12).

6 This effect was replicated in study S3a in the web appendix; participants with an intact logged streak reported adopting a streak maintenance goal to a greater extent than participants with a broken logged streak ($F(1, 592) = 48.70$, $p < .001$, $d = 0.66$). Additionally, study S3b (web appendix) conceptually replicated this effect using two secondary measures of our proposed process ("How much did you think about consistency in your logging behavior?" and "How much did you think about streaks you may have in your log?"; $F(1, 503) = 43.37$, $p < .001$, $d = 0.59$). Moreover, in both studies these measures mediated the effect of an intact versus broken logged streak on subsequent behavior (*S3a:* Indirect effect = 0.62, SE = 0.11, 95% CI = [0.42, 0.85]; *S3b:* Indirect effect = 0.33, SE = 0.08, 95% CI = [0.19, 0.53]; see web appendix for full methods and results).

FIGURE 3

PERCENTAGE OF PARTICIPANTS IN STUDY 4 WHO ENGAGED IN THE TARGET BEHAVIOR AS A FUNCTION OF HAVING AN INTACT LOGGED STREAK, A BROKEN LOGGED STREAK, OR NO BEHAVIORAL LOG. ERROR BARS REPRESENT 95% CONFIDENCE INTERVALS



$= 7.85$, $p < .001$). Participants in the *intact logged streak* condition felt a greater sense of accomplishment ($M = 7.49$, $SD = 2.71$) than participants in the *broken logged streak* condition ($M = 6.51$, $SD = 2.55$; $t(398) = 3.71$, $p < .001$; $d = 0.37$). In addition, the *no-log* condition fell between these two conditions: participants in the *intact logged streak* condition felt a directionally greater sense of accomplishment relative to this control ($M = 7.24$, $SD = 2.39$; $t(401) = 0.99$, $p = .32$; $d = 0.10$), while those in the *broken logged streak* condition felt a significantly lower sense of accomplishment ($t(397) = 2.93$, $p = .004$; $d = 0.30$).

Mediation Analyses. We conducted mediation analyses using a bootstrap procedure with 10,000 samples (SAS Process Macro, Model 4; Hayes 2017) to test if the effect of logged streaks on subsequent behavior was driven by the extent to which participants had a goal of streak maintenance. Our primary model compared the *intact logged streak* condition (1) to the *broken logged streak* condition (0) as the independent variable, the *streak maintenance goal* measure as the mediator, and choice as the dependent variable. As predicted, the extent to which participants reported streak maintenance as a goal mediated the effect of an intact versus broken logged streak on subsequent engagement in the target behavior (Indirect effect $= 0.20$, $SE = 0.08$, 95% confidence interval [CI] $= [0.05, 0.37]$). Results were similar when we instead used our *sense of accomplishment* measure as the mediating variable (Indirect effect $= 0.31$, $SE = 0.09$, 95% CI $= [0.15, 0.52]$; see web appendix for mediation analyses relative to the *no-log* condition).

Behavioral Intentions. Participants in the *intact logged streak* condition were more likely to report that they would continue using the app, recommend it to a friend, and want to see the app's behavioral log in the future, compared to participants in the *broken logged streak* condition ($ts > 2.50$, $ps < .02$, $ds > 0.25$). Participants in the *no-log* condition were more likely to continue using the app and recommend it to a friend than those in the *broken logged streak* condition ($ts > 3.00$, $ps < .01$, $ds > 0.30$), but were not significantly different from participants in the *intact logged streak* condition for these measures ($ts > 0.40$, $ps < .70$, $ds > 0.05$).

Discussion

Study 4 replicated the finding that participants were more likely to continue engaging in real, consequential behavior when they had an intact (vs. broken) streak within their behavioral log, controlling for their actual past behavior. This study also provides evidence for our mechanism (hypothesis 2): the effect of intact versus broken logged streaks on subsequent behavior was mediated by the extent to which participants reported streak maintenance as an explicit goal, as well as by an indicator of goal pursuit—their sense of accomplishment. Additionally, participants were willing to forgo a more enjoyable outside option to keep their logged streaks intact.[7] Besides demonstrating the value of such streak maintenance goals, this finding is of practical relevance as consumers continually face trade-offs between engaging in less immediately appealing, goal-consistent behaviors that they feel they should do (e.g., studying) versus more enjoyable behaviors that they want to do (e.g., watching TV; Bitterly et al. 2015).

STUDY 5: HOW CATEGORIZING BEHAVIORS AS CONTRIBUTING TO A LOGGED STREAK AFFECTS BEHAVIOR

Study 5 extends our previous findings by isolating the effects of logged streaks while holding participants' actual behavior constant in a different way: by manipulating the *types* of behaviors that counted toward their logged streaks. To do so, we leveraged previous work on categorization (Redden 2008) and defined participants' streaks by logging either broader or narrower categories of behavior. Accordingly, the same behavior was either logged as a contribution to an intact streak (when the category of logged behaviors was broader) or as a break in a streak (when the category was narrower). As in studies 2 and 4, this design held participants' actual behavior constant across conditions and varied only how it was logged, thereby enabling a clean test of how logged streaks affect subsequent behavior. Moreover, as in the previous study, we included a *no-log* condition where participants were not provided with a behavioral log.

This study also examined yet another consumer behavior: playing games on an app (specifically, word and number games). We chose to examine this behavior because gaming apps are quite popular and often incorporate behavioral logs (see Consumer Logging Pilot Study section), with over 60% of Americans playing games on their smartphones (Lynkova 2020). For robustness, we also used a different focal dependent variable in this study: participants' choice between continuing the target behavior versus stopping (rather than specifying an outside option).

Methods

We recruited 805 MTurk participants ($M_{age} = 35.65$, 46.68% female), who learned that they would be testing up to two different types of games for an app under development: (1) "Number Sums," where they had to find two numbers in a matrix that summed to 200 and (2) "Word Jumbles," where they had to unscramble letters to form a word. All participants chose whether to start with Word Jumbles or Number Sums.

Participants were assigned to one of three between-subjects conditions. As in study 4, in the *intact logged streak* and *broken logged streak* conditions, we informed participants that the app featured a behavioral log, which was again modeled on real-life apps. Those in the *intact logged streak* condition read that completing both game types (i.e., Word Jumbles *and* Number Sums) would give them a checkmark on the behavioral log, while those in the *broken logged streak* condition read that only completing the game type that they had chosen would give them a checkmark on the behavioral log (i.e., either Word Jumbles *or* Number Sums). Thus, the types of games that "counted" in the behavioral log differed by condition: broader categorization of behaviors allowed participants to maintain their streaks by completing either type of game, while narrower categorization meant that participants could only maintain their streaks by completing the game type they chose at the beginning (and would thus break their streak if they completed the other game type). Similar to the previous study, participants in the *no-log* condition played the same four games on a version of the app that lacked the behavioral log but was otherwise identical.

After completing several instruction comprehension checks, all participants could complete three games of the type they chose. Participants in this study were allowed to complete as many or as few games as they wanted; that is, they were given the option to continue or stop after each game. When participants reached the fourth game, we informed them that they would instead complete the other type of game mentioned in the instructions (e.g., if they chose to start with Word Jumbles, they would now complete a Number Sum). Thus, all participants completed the exact same series of games (e.g., three Word Jumbles and then one Number Sum), but the fourth game either maintained or broke their logged streak, depending on condition. Our key dependent measure was participants' decision of whether to continue playing games (i.e., the target behavior) or stop.

Regardless of their decision, all participants then answered the same *sense of accomplishment* measures as in study 4. Furthermore, to explore the role of emotions in these effects, participants also answered four items about how they felt right after completing their most recent game (how annoyed, angry, upset, and sad they felt; all measured from 1 "Not at all" to 11 "Extremely"; $\alpha = .94$). We expected that if participants viewed maintaining their logged streaks as a goal, then broken streaks may induce more negative emotion (Bagozzi and Pieters 1998). However, we did not have strong a priori predictions about whether these emotions would *drive* the effects of logged streaks on behavior.

Participants also answered a free-response question about their decision to continue or stop playing games, as

7 We replicated this finding in two additional, preregistered studies (studies S3a and S4; see web appendix). Specifically, participants were more likely to continue language learning versus watching a fun video (S3a) and to continue language learning versus viewing funny online content (S4) when they had an intact logged streak, compared to when they had a broken logged streak ($Fs > 8.00$, $ps < .005$, ORs > 1.60).

well as a manipulation check about how aware they were that they had a streak of attempting games (from 1 "Not aware at all" to 11 "Extremely aware"). Finally, participants answered five exploratory measures about their overall attitude toward the games, one question about how often they used similar apps, and demographics. After these questions, they were told that they were finished with the study and did not need to play any more games.

Results

Three hundred fifty-three participants chose to stop playing games before our manipulation and thus could not be included in our final sample or analyses. Importantly, attrition rates did not differ between conditions (X^2 (2) = 1.97, p = .37; see web appendix for details). This left a final sample of 452 participants, as targeted (see preregistration).

Manipulation Check. A one-way ANOVA confirmed an effect of condition on awareness of having a streak of playing games ($F(2, 449) = 14.13$, $p < .001$). Participants in the *intact logged streak* condition ($M = 8.80$, SD = 2.46) and the *broken logged streak* condition ($M = 8.61$, SD = 2.19) were more aware of having a streak than participants in the *no-log* condition ($M = 7.29$, SD = 3.23; ts > 4.15, ps < .001, ds > 0.45). There was no detectable difference between the two *logged streak* conditions ($t(306) = 0.71$, $p = .48$, $d = 0.08$).

Target Behavior. A binary logit revealed an effect of condition on behavior ($F(2, 449) = 4.61$, $p = .010$; figure 4). As predicted, participants in the *intact logged streak* condition were more likely to engage in the target behavior (82.55%) than participants in the *broken logged streak* condition (67.92%; X^2 (1) = 8.78, $p = .003$; OR = 2.23). In addition, participants in the *intact logged streak* condition were directionally more likely to engage in the target behavior than participants in the *no-log* condition (77.78%; X^2 (1) = 1.05, $p = .31$; OR = 1.35), while participants in the *broken logged streak* condition were marginally less likely to engage in the target behavior than participants in the *no-log* condition (X^2 (1) = 3.69, $p = .055$, OR = 0.61).

Indicators of Goal Pursuit. As expected, participants in the *intact logged streak* condition felt a significantly greater sense of accomplishment ($M = 8.49$, SD = 1.99) and less negative emotion ($M = 2.80$, SD = 2.87) than participants in the *broken logged streak* condition ($M_{\text{accomplishment}} = 7.45$, SD = 2.38; $M_{\text{emotion}} = 5.05$, SD = 3.51; ts > 4.10, ps < .001, ds > 0.45). Additionally, the *no-log* condition fell between these two conditions for both measures ($M_{\text{accomplishment}} = 8.18$, SD = 2.10; $M_{\text{emotion}} = 3.25$, SD = 3.22), with participants in the *intact logged streak* condition feeling a directionally greater sense of accomplishment and directionally lower negative emotion (ts > 1.25, ps < .22, ds > 0.14), and participants in the *broken logged streak* condition feeling a significantly lower sense of accomplishment and significantly greater negative emotion (ts > 2.80, ps < .006, ds > 0.32). Separate exploratory mediation models revealed that while the *sense of accomplishment* measure mediated the effect of an intact versus broken logged streak on subsequent behavior (Indirect effect = 0.49, SE = 0.15, 95% CI = [0.25, 0.82]), the *negative emotion* measure did not (Indirect effect = 0.11, SE = 0.10, 95% CI = [−0.07, 0.32]; see web appendix for mediation analyses relative to the *no-log* condition).

Discussion

Study 5 replicated the effects of intact versus broken logged streaks on subsequent behavior using a novel manipulation: the categorization of behaviors in a behavioral log. That is, defining participants' streaks by allowing a broader (vs. narrower) category of behavior to count in the behavioral log led them to be more likely to engage in the target behavior. Thus, by again holding participants' actual behavior constant across conditions and only varying whether it was logged on the app, we show the value they placed on their logged streaks. Moreover, since the broken streak was caused by engaging in a behavior that participants knew *beforehand* would not count on their behavioral log, its effects on subsequent behavior cannot be explained by negative beliefs or feelings toward the app due to a perceived service failure (i.e., the inability to log a behavior).

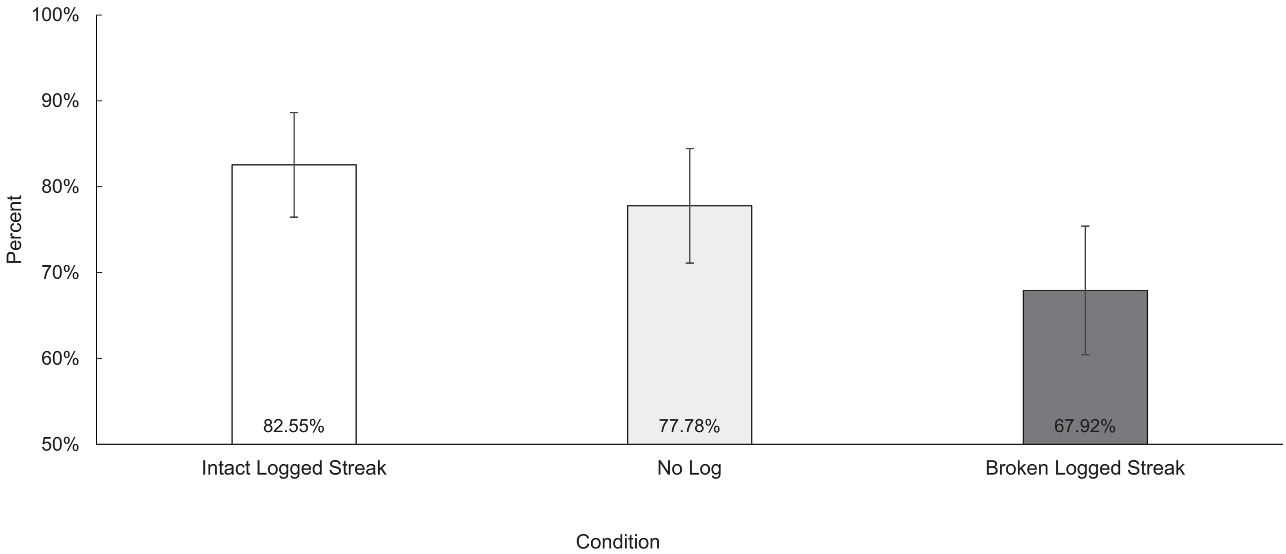
As in study 4, this study featured a *no-log* condition, which allowed us to again examine the effects of highlighting intact and broken streaks relative to not seeing a behavioral log at all (similar to the comparisons with the *behavioral log absent* conditions in study 3). However, because these studies were not powered specifically to estimate these individual contrasts (as per our preregistrations), these comparisons yielded mixed evidence. Specifically, relative to the absence of a behavioral log, highlighting an intact logged streak had a stronger positive effect in study 4, while highlighting a broken logged streak had a stronger negative effect in study 5. That said, although the significance levels of these contrasts differed for these two studies, the *no-log* condition always fell between the two *logged streak* conditions for all measures (as predicted), and there were no differences in the estimated effect sizes of these contrasts (Gelman and Stern 2006). We revisit this issue in the General Discussion section as an avenue for future research.

STUDY 6: THE EFFECT OF LOGGED STREAKS IS AMPLIFIED WHEN CONSUMERS ATTRIBUTE RESPONSIBILITY FOR THE BREAK TO THEMSELVES

In studies 2–5, participants in the *broken streak* conditions experienced breaks that were caused by factors

FIGURE 4

PERCENTAGE OF PARTICIPANTS IN STUDY 5 WHO ENGAGED IN THE TARGET BEHAVIOR AS A FUNCTION OF HAVING AN INTACT LOGGED STREAK, A BROKEN LOGGED STREAK, OR NO BEHAVIORAL LOG. ERROR BARS REPRESENT 95% CONFIDENCE INTERVALS



outside their control (e.g., an issue with the app). But of course, consumers frequently encounter situations where they feel personally responsible for their broken streaks, such as when they forget to log or run out of time before completing the relevant behavior. Past research suggests that attributing goal failure to one's own actions reduces self-efficacy and future goal pursuit (Bandura and Locke 2003). Accordingly, in study 6, we investigated whether the effect of intact versus broken streaks would be moderated by participants' attribution of the break to themselves, versus an external factor. If consumers view maintaining a logged streak as a goal in itself, as theorized, then attributing a broken logged streak to themselves should amplify the effect.

This study also builds on study 4's finding that participants were willing to sacrifice the opportunity to engage in a more enjoyable activity to maintain an intact streak. Here, we examined participants' willingness to watch an advertisement to repair or maintain their streaks. If, as posited, participants value their logged streaks, we would expect them to be willing to engage in an undesirable behavior (watching an advertisement) to keep them intact.

Methods

We recruited 802 MTurk participants ($M_{age} = 39.10$, 54.36% female). As in study 5, all participants were informed that they would be testing out a gaming app that featured a behavioral log. The app included both Number Sums and Word Jumble games, but all participants started by testing Number Sums. As in study 3, all participants were told that they might see a "quota" message when enough participants had already played a specific game, meaning that they were not needed to test that particular game. They completed several comprehension checks about these instructions, then proceeded to play Number Sum games.

Participants were assigned to one of three between-subjects conditions (*intact streak*, *externally-attributed broken streak*, or *self-attributed broken streak*). Participants in the *intact streak* condition were told that any Number Sum game attempt would give them a checkmark on their behavioral log (even if they did not successfully complete it), while participants in both *broken streak* conditions were told that only successfully completing a Number Sum game would give them a checkmark on their log.

All participants first played three Number Sum games that were relatively easy. Our manipulation varied what happened during the fourth Number Sum game. Participants in the *intact streak* condition attempted to play a fourth game that was quite difficult (in fact, only two participants in the entire sample correctly completed it); because all game attempts for participants in this condition counted on the behavioral log, these participants maintained their logged streaks. Similar to study 3, participants in the *externally-attributed broken streak* condition saw the aforementioned quota message and were not able to play the fourth game, which resulted in a broken streak. Finally, participants in the *self-attributed broken streak* condition played the same difficult fourth Number Sum game as the *intact streak* condition; however, because only correctly completed Number Sum games were added to the behavioral log, this resulted in a broken streak. Thus, participants in both *broken streak* conditions saw the *same* broken streak within their behavioral log, with only the purported cause of this break (and therefore responsibility for it) differing by condition: either an external factor (because the fourth game was unavailable) or participants themselves (because they were not able to correctly complete the fourth game). A manipulation check confirmed that participants in the *self-attributed broken streak* condition ($M = 7.70$, $SD = 3.18$) attributed the broken streak to their own actions (vs. something outside of their control) more than participants in the *externally-attributed broken streak* condition ($M = 1.95$, $SD = 2.30$; $t(290) = 17.74$, $p < .001$, $d = 2.07$).

Next, all participants chose between continuing to play a Number Sums game (the target behavior) and switching to the other game available on the app (Word Jumbles). This choice served as our key dependent measure. Regardless of their decision, all participants then answered the same questions as in study 4 regarding their behavioral intentions to continue using the app and recommend it to a friend. Additionally, they reported whether they would be willing to watch an advertisement to maintain their logged streak (*intact streak* condition) or to repair their broken logged streak (*broken streak* conditions; "Yes" or "No"). Finally, participants completed their chosen game and answered demographics.

Results

We recruited 802 participants over 4 days. As preregistered, our final sample included all participants who correctly completed the first three games (which did not differ by condition: X^2 (2) $= 3.52$, $p = .172$; see web appendix for details). This left a final sample of 418 participants. Although lower than our target sample size (as the completion rate was lower than anticipated), we were concerned about the quality of additional participants after having already recruited for 4 days. Given that the final sample contained over 130 per condition, we felt it was adequately powered and therefore stopped data collection (see web appendix for analyses with the full sample, which replicated the effects).

Target Behavior. A binary logit revealed an effect of condition on behavior ($F(2, 415) = 8.02$, $p < .001$; figure 5). Similar to previous studies, participants in the *intact streak* condition were marginally more likely to engage in the target behavior (53.17%) compared to participants in the *externally-attributed broken streak* condition (42.00%; X^2 (1) $= 3.43$, $p = .064$, $OR = 1.57$). Critically, this difference was larger when comparing the *intact streak* condition to the *self-attributed broken streak* condition (28.87%; X^2 (1) $= 16.39$, $p < .001$, $OR = 2.80$). Moreover, participants in the *self-attributed broken streak* condition were less likely to engage in the target behavior than participants in the *externally-attributed broken streak* condition (X^2 (1) $= 5.48$, $p = .019$, $OR = 0.56$). Thus, participants were more likely to engage in the target behavior when they attributed the break to an external factor rather than to themselves.

Behavioral Intentions. Notably, 45% of participants were willing to watch an advertisement to maintain their logged streak, and 43% were willing to watch an advertisement to repair their broken logged streak (48% in the *externally-attributed* condition, and 39% in the *self-attributed* condition). There was no significant effect of *streak* condition on their likelihood to recommend or continue using the app (Fs < 2.00, ps $> .20$).

Discussion

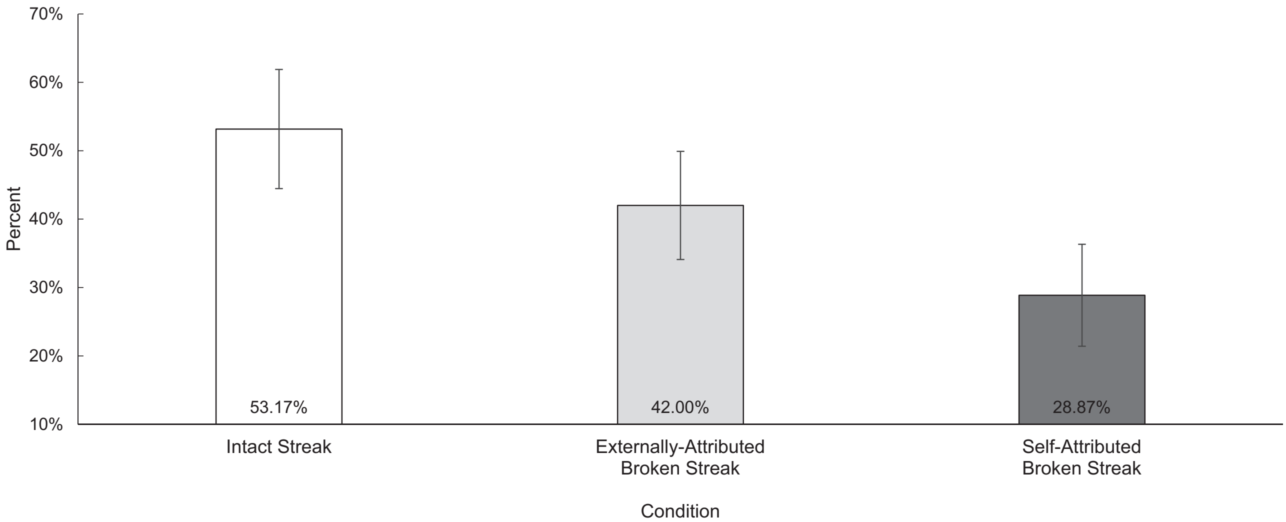
Study 6 replicated the effect of intact versus broken logged streaks on subsequent behavior. Moreover, this effect was amplified when participants attributed the broken streak to themselves rather than an external factor (hypothesis 3). This result supports our theory, as individuals who feel responsible for their goal failure are even less likely to continue pursuing that goal. This amplification is particularly striking given that the *self-attributed broken streak* condition again controlled for participants' actual behavior relative to the *intact streak* condition. Indeed, similar to the categorization manipulation in study 5, whether the participant's streak was intact or broken simply depended on what behaviors "counted" in their behavioral log.

Interestingly, almost half the participants reported being willing to engage in an unrelated, undesirable behavior—watching an advertisement—to maintain or repair their logged streaks. This result is notable given the effort and costs consumers often incur to *avoid* watching advertisements (Anderson and Gans 2011).[8] Building on these results, our next study directly manipulated participants' ability to repair their broken streaks and examined how this affected their subsequent behavior.

8 We found similar results in an additional study (study S3a; see web appendix). Specifically, 51% of participants were willing to watch an advertisement to maintain their logged streaks, and 46% of participants were willing to watch an advertisement to repair their broken logged streaks.

FIGURE 5

PERCENTAGE OF PARTICIPANTS IN STUDY 6 WHO ENGAGED IN THE TARGET BEHAVIOR AS A FUNCTION OF HAVING AN INTACT STREAK, AN EXTERNALLY-ATTRIBUTED BROKEN STREAK, OR A SELF-ATTRIBUTED BROKEN STREAK IN THEIR BEHAVIORAL LOG. ERROR BARS REPRESENT 95% CONFIDENCE INTERVALS



STUDY 7: THE EFFECT OF LOGGED STREAKS IS ATTENUATED BY THE ABILITY TO REPAIR THE BREAK

In our final study, we examined whether providing participants with the ability to repair a broken logged streak (e.g., by allowing re-engagement in the behavior to "fill in" the miss that broke the streak) would moderate the effect of intact versus broken streaks on subsequent behavior. If consumers view maintaining their logged streaks as a goal, as we theorize, then a repair opportunity should enable them to resume goal pursuit and reduce feelings of goal failure (Cochran and Tesser 1996). Thus, allowing streak repair should *attenuate* the effect of intact versus broken logged streaks.

Methods

We recruited 601 MTurk participants ($M_{age} = 35.80$, 45.59% female). All participants read about the same gaming app as in previous studies and were given a choice of two types of games to start with: either Word Jumbles or Number Sums. They were also informed that only the game type they chose would count toward the behavioral log featured on the app. As in studies 3 and 6, all participants read that they might encounter a "quota" message once while playing, indicating that they did not need to play that particular game. They answered instruction comprehension checks before proceeding to the app.

Participants were randomly assigned to one of three conditions (*intact streak*, *broken streak*, and *repairable broken streak*). Participants in the *intact streak* condition completed four games in a row and were informed that they had an intact streak on their behavioral log, while participants in both *broken streak* conditions completed three games and then saw the "quota" notification, leading to a broken streak on their log.

Next, all participants chose what type of game they wanted to play next—either the same game type they started with or the other available game type. This served as our key dependent variable. Participants in the *repairable broken streak* condition were told that choosing the same game type would restore their streak in the behavioral log: specifically, that it would "be repaired to 4 in a row." As in previous studies, participants in the other two conditions were not told anything about potential streak repair. After making this choice, all participants answered similar *sense of accomplishment* ($\alpha = .92$) and *negative emotion* measures ($\alpha = .97$) as in previous studies. Participants also answered a free-response question about their decision and five exploratory measures about their experience playing the games. They then played their chosen game and answered demographics.

Results

Target Behavior. A binary logit found a significant effect of condition on behavior (X^2 (2) = 43.64, $p < .001$;

figure 6). Participants in the *intact streak* condition were more likely to engage in the target behavior (93.14%) than in the *broken streak* condition (68.66%; X^2 (1) = 39.41, $p < .001$, OR = 6.20) and the *repairable broken streak* condition (85.20%; X^2 (1) = 6.56, p = .010, OR = 2.36). As predicted, participants in the *repairable broken streak* condition were more likely to engage in the target behavior than in the *broken streak* condition (X^2 (1) = 15.26, $p <$ .001, OR = 2.63). That is, the opportunity for streak repair attenuated the effect of an intact (vs. broken) streak within the behavioral log.

Indicators of Goal Pursuit. Participants in the *intact streak* condition felt a greater sense of accomplishment (M = 8.56, SD = 1.90) and less negative emotion (M = 2.11, SD = 2.33) than those in the *broken streak* condition ($M_{\text{accomplishment}}$ = 7.37, SD = 2.31; M_{emotion} = 3.16, SD = 2.77) and the *repairable broken streak* condition ($M_{\text{accomplishment}}$ = 8.10, SD = 1.97; M_{emotion} = 3.33, SD = 2.83; ts > 2.30, ps < .020, ds > 0.20). Consistent with our theory, participants in the *repairable broken streak* condition also felt a greater sense of accomplishment than those in the *broken streak* condition (t(395) = 3.35, $p < .001$; d = 0.34). The *broken streak* conditions did not differ in terms of negative emotion (t(398) = 0.59, p = .256; d = 0.06). Replicating study 5, sense of accomplishment mediated the effect of an intact versus broken logged streak on engagement in the target behavior (Indirect effect = 0.28, SE = 0.09, 95% CI = [0.13, 0.48]), while negative emotions did not (Indirect effect = −0.08, SE = 0.06, 95% CI = [−0.21, 0.02]; see web appendix for mediation analyses involving the *repairable broken streak* condition).

Discussion

Study 7 replicated the effect of intact versus broken logged streaks on subsequent behavior and also demonstrated that the opportunity for streak repair attenuated this effect. This provides convergent evidence for our theory: the ability to repair one's streak within the behavioral log (and thus continue the goal of streak maintenance) motivated participants to continue engaging in the actual target behavior (hypothesis 4).[9]

Interestingly, streak repair did not fully eliminate the effect of logged streaks, possibly because participants perceived a repaired streak as less valuable or authentic than one never broken in the first place. Still, the demonstrated attenuation due to streak repair underscores the value consumers place on their streaks within behavioral logs. That is, even though participants in both *broken streak* conditions engaged in the same series of actual behaviors and had the same negative experience of a broken streak, simply having the opportunity to restore the streak within their behavioral log influenced their subsequent behavior.

GENERAL DISCUSSION

Technology has made it increasingly easy for consumers to log and track their repeated behaviors over time, making them more aware of their patterns of behavior than ever before. Specifically, apps frequently highlight consumers' streaks of behaviors and alert consumers when they break those streaks. Our research is the first investigation into how and why these logged streaks influence consumers' motivation and subsequent decisions. Across seven studies, we found that participants were more likely to engage in a target behavior when they had an intact (vs. broken) streak highlighted via a behavioral log, even when their series of actual behaviors was exactly the same. Notably, all of these studies showed effects on real consumer choices, both in the field (study 1) and controlled lab paradigms (studies 2–7), as well as for a variety of consumer-relevant, consequential domains, including exercise (studies 1 and 2), language learning (studies 3 and 4), and games (studies 5, 6, and 7).

We also found evidence that this effect arises because consumers adopt a goal of maintaining a logged streak. We demonstrated this process via mediation using both a direct measure of goal adoption (study 4) and a key indicator of goal progress (sense of accomplishment; studies 4, 5, and 7). Consistent with our theory, we also found that the effect of intact versus broken streaks was magnified when consumers felt responsible for the break in their behavioral log (study 6), and attenuated when they had the opportunity to repair it (study 7).

Theoretical Contributions

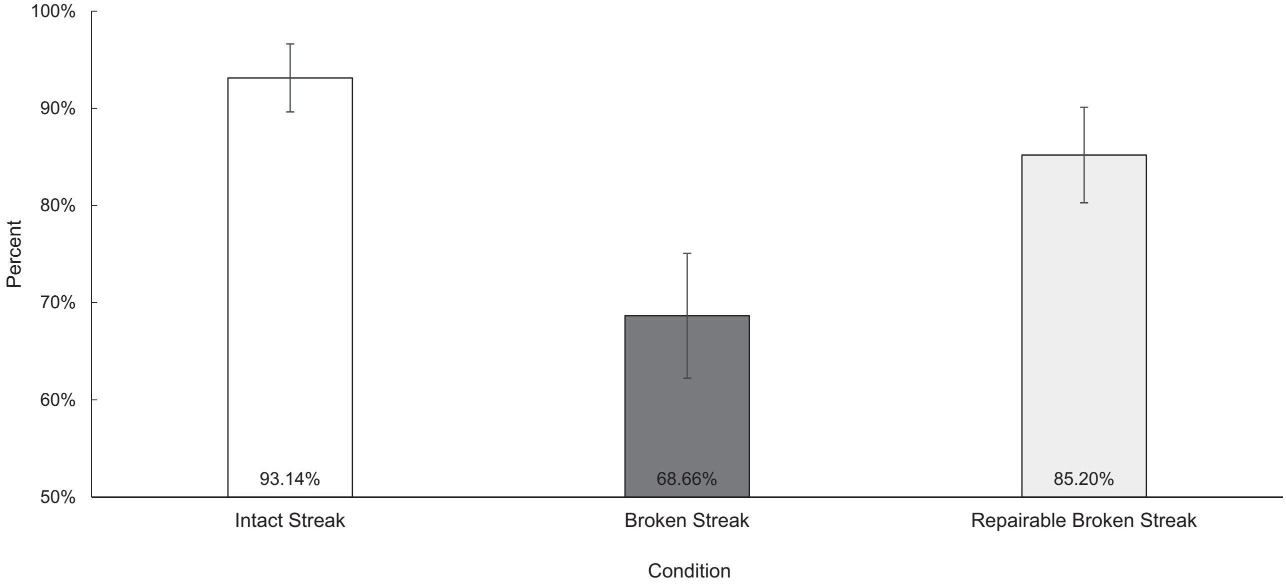
Our work offers several novel insights for consumer research. First, it advances prior work on goal progress by being the first to demonstrate that a highlighted pattern of behavior (i.e., a logged streak) can become a goal in and of itself. While most work on goal progress has examined goals defined by specific end states or outcomes (e.g., the number of coffees needed to earn a freebie: Kivetz et al. [2006]; completing a collection: Gao et al. [2014]), our research identifies a goal that is defined by an ongoing process without a clear end state (since streaks can only be maintained, not "completed"). As such, it extends recent findings on maintenance goals (Yang, Stamatogiannakis, and Chattopadhyay 2015) and "do your best" goals (Wallace and Etkin 2018), which have revealed novel motivational effects of goals distinguished by a process rather than a final outcome.

More broadly, we contribute to a large body of work on how consumers' past decisions affect their subsequent

9 This finding also helps to rule out the possibility that a broken streak creates a "partition" that allows consumers to reconsider their future behavior (Cheema and Soman 2008), as this alternative would not predict moderation by the opportunity for streak repair.

FIGURE 6

PERCENTAGE OF PARTICIPANTS IN STUDY 7 WHO ENGAGED IN THE TARGET BEHAVIOR AS A FUNCTION OF HAVING AN INTACT, BROKEN, OR REPAIRABLE BROKEN STREAK IN THEIR BEHAVIORAL LOG. ERROR BARS REPRESENT 95% CONFIDENCE INTERVALS



behavior. Previous work has demonstrated that consumers tend to act consistently over time, especially with regard to their most recent behavior (e.g., shopping "momentum": Dhar et al. [2007]; bingeing: Schweidel and Moe [2016]) or a high overall frequency of behavior (e.g., loyalty: Nunes and Drèze 2006). We build on this research by examining the effect of *highlighting* a specific, commonly occurring pattern of behavior—a streak—on consumers' subsequent decisions. Additionally, we contribute to past work showing that consumers continue certain behaviors after several past instances (e.g., collecting two or three items: Gao et al. [2014]; acquiring a medium like loyalty points: Hsee et al. [2003]) by demonstrating the effects of highlighting multiple *consecutive* past behaviors on subsequent behavior.

Finally, our work contributes to growing research on how technology affects consumers' lives (Deighton et al. 2017; Schmitt 2019). In particular, while recent work on behavioral tracking has explored its experiential consequences (e.g., effects on enjoyment: Etkin [2016]), we investigate how streaks made salient through such tracking influence subsequent behavior. Furthermore, our findings inform our understanding of how the symbols used in behavioral tracking (e.g., the checkmarks in behavioral logs) can have a special influence on consumers' behavior. Building on past work showing the value of such symbols (Hamari et al. 2014; Hsee et al. 2003), we demonstrate how a specific pattern represented within a behavioral log can be reinforcing in itself.

Practical Implications

Our work provides substantive insights for companies seeking to increase and maintain consumer engagement through logging tools. While firms have no direct control over the sequences of behavior consumers choose to engage in, they have almost total discretion over whether, when, and how to employ tracking technologies that log and display information about those behaviors to consumers. Our findings suggest several ways that companies can leverage these technologies to improve consumer engagement.

First and foremost, companies may need to use different communication strategies depending on consumers' recent patterns of behavior. While highlighting intact streaks via behavioral logs can encourage consumers to maintain them, companies might want to avoid highlighting consumers' broken streaks (e.g., by not sending notifications to alert them when they break their streak, a current practice in many apps). Companies could further boost engagement through what they display to consumers within their behavioral logs, even when consumers engage in the same series of behaviors. In particular, companies may want to incorporate flexibility into what counts toward a streak (e.g., by broadly defining which behaviors count toward streaks, rather than using narrower subcategories, as shown in studies 5 and 6). They could also consider greater flexibility in the time periods that count toward a streak (i.e., whether streaks are calculated at the daily or weekly level).

Larger units of time could allow irregular or inconsistent behavior to be portrayed as an intact streak within the consumer's log, thus making it easier for them to maintain their goal progress. However, while defining streak-contributing behaviors more broadly may keep less-motivated consumers engaged, it may also decrease the value and meaning of maintaining a streak.

Additionally, our findings indicate several potential solutions that companies could implement when consumers (inevitably) break their logged streaks. For one, the results of study 6 suggest that rather than allowing consumers to attribute broken streaks to themselves, companies could lessen the negative impact by shouldering some of the responsibility (e.g., by communicating to users that the company is partially responsible for the break). Companies could also "retarget" consumers who have broken their streaks to get them back on track. As demonstrated in study 7, providing opportunities for streak repair (e.g., by allowing consumers to fill in the "miss") can attenuate the negative effect of a broken (vs. intact) streak. Alternatively, study 6 suggests that some consumers would even engage in an unrelated task, like watching an advertisement, to restore their logged streaks. In fact, several companies have started to implement similar interventions that allow consumers to keep their streaks alive after a miss while simultaneously generating revenue. For example, Duolingo allows users to buy a "streak freeze" with in-app currency so they can pre-emptively keep their streak intact if they ever miss a day of language learning.

Beyond the company benefits of increased consumer engagement, offering behavioral logs that highlight streaks may improve consumer welfare as well. Many consumers use apps to track and motivate behaviors that are challenging to achieve, such as exercising (Fox 2013) and budgeting (Malcolm 2015). Helping consumers stay (or get back) on track by highlighting streaks should increase their persistence toward personal goals and thus represents a win–win situation, where companies profit from long-term consumer retention and consumers benefit from enhanced well-being.

Future Directions

To our knowledge, our experiments are the first to examine the effects of logged streaks on subsequent behavior, and as such open a number of avenues for further investigation. First, we primarily examined the effects of streaks defined by three logged behaviors in a row, consistent with prior work on how people classify streaks (Carlson and Shu 2007; see also study S1 in the web appendix). Since these findings cannot address how the length of a logged streak might affect consumers' motivation to maintain it, we ran two initial studies to investigate this question. Study S3a (web appendix; $N = 596$) examined the effects of shorter (three in a row) versus longer (six in a row) streaks of logging language-learning questions using our app paradigm, and study S3b (web appendix; $N = 507$) examined the effects of shorter (4 in a row) versus longer (20 in a row) streaks of logging beers in a beer-tasting app using a scenario. Interestingly, both studies replicated the effects of intact versus broken logged streaks on subsequent behavior, regardless of streak length (study S3a in the web appendix: $ps < .001$, $ORs > 2.40$; study S3b in the web appendix: $ps < .001$, $ds > 0.45$), and neither study revealed a significant interaction between the *streak* and *length* conditions ($ps > .40$). These results are consistent with our theory, suggesting that the activation of a streak goal (from an intact logged streak) may be somewhat categorical and insensitive to streak length (Sheeran, Webb, and Gollwitzer 2005). However, though these studies find similar effects for streaks of up to 20 logged behaviors in a row, they cannot speak to any potential differences in these effects for even longer streaks (e.g., 100 consecutive behaviors). Future work should investigate this possibility. On the one hand, at least anecdotally, consumers seem quite proud of and motivated by incredibly long streaks and are especially devastated when they are broken (Lorenz 2017), so it is possible that the effects of logged streaks could be amplified after a certain length. On the other hand, consumers could eventually become satiated with the repeated behavior and lose interest in it (Galak and Redden 2018); their desire for variety could outweigh their desire to continue pursuing their streak goal, thus attenuating the effect of logged streaks after a certain length.

Second, while our research has focused on the effects of highlighting intact versus broken streaks via behavioral logs, open questions remain about the possible independent effects of intact and broken streaks. In particular, studies 3, 4, and 5 found some evidence for the positive effect of an intact logged streak and the negative effect of a broken logged streak, relative to the absence of a behavioral log, but the statistical significance of these contrasts was not consistent across studies. Future work should investigate the presence and magnitude of distinct effects of intact and broken streaks more closely, and the mechanisms that give rise to stronger effects in certain situations (e.g., losses looming larger than gains: Tversky and Kahneman [1979]).

Third, future work should explore how various features of behavioral logs themselves might influence the effects of streaks on subsequent behavior. For instance, behavioral tracking can require different levels of effort from the consumer, ranging from automatically tracking consumers' behaviors for them (as in studies 1, 5, 6, and 7) to requiring consumers to actively log their behaviors through a unique action (as in studies 2, 3, and 4). Because more effortful behaviors are a stronger self-signal of consumers' own attitudes and intentions (Ajzen and Fishbein 1977), it is possible that more active logging would increase the effects of streaks on behavior compared to automatic tracking.

Additionally, the extent to which logged streaks affect behavior might depend on whether behavioral logs are made public or kept private. For the most part, our studies examined behaviors that were completed and logged privately, or at least not made explicitly public. But consumers often share information about their logged behaviors in person and on social media, and many apps encourage users to interact with others. It is possible that making logged streaks more public might amplify their effects by increasing consumers' sense of accountability (Rogers et al. 2015) or perceived status (Moldovanu, Sela, and Shi 2007). As an initial investigation of this possibility, we ran an additional preregistered study (study S4 in the web appendix, $N = 604$), which manipulated whether participants intact (vs. broken) logged streaks would be kept private or shared with future users of the app. Replicating our central finding, participants with an intact logged streak were more likely to continue the target behavior (55.78%) than those with a broken logged streak (43.85%; X^2 (1) = 8.58, p = .003, OR = 1.61). Moreover, although the interaction between the *streak* and *public* conditions was not significant ($F(1, 600) = 0.14$, p = .71), the effect of intact versus broken streaks was directionally stronger when the log would be made public (OR = 1.72) versus kept private (OR = 1.52). Future work might systematically test the influence of various features that make behavioral logs more versus less public, such as when there are opportunities to post and comment on others' streaks, as consumers may be especially invested in such feedback.

Finally, in certain cases, consumers may be intuitively aware of their streaks of behavior, even without behavioral trackers. Future research should explore the factors that might increase the natural salience of streaks and therefore the extent to which they affect consumer decisions without being emphasized via behavioral logs. For example, complicated or unique types of behavior might make consumers more likely to notice and remember their streaks (Kausler and Hakami 1983; Thompson 1982). Individual differences might also matter; consumers high in need for cognition (Cacioppo and Petty 1982) or structure (Neuberg and Newsom 1993) may think more critically about patterns in their behavior, thus boosting the salience of streaks even without behavioral logs highlighting them.

Conclusion

The study of how logged streaks affect consumer behavior is still in its infancy, and many potentially fruitful questions remain. Given the growing prevalence of apps that track consumers' behaviors, as well as the evidence presented here that consumers view maintaining their logged streaks as a goal in and of itself, research in this area has important implications for practitioners and academics alike. Such insights can only become more vital as advances in technology further increase the salience of behavioral streaks over time.

DATA COLLECTION INFORMATION

The authors collected data through Amazon's Mechanical Turk and behavioral labs at the University of Delaware and New York University between Winter 2017 and Winter 2021. The first author collected and analyzed all data. All data, stimuli, and preregistrations, as well as our web appendix, can be found in our OSF repository: https://osf.io/kpjh9/.

REFERENCES

Ajzen, Icek and Martin Fishbein (1977), "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research," *Psychological Bulletin*, 84 (5), 888–918.

Anderson, Simon P. and Joshua S. Gans (2011), "Platform Siphoning: Ad-Avoidance and Media Content," *American Economic Journal: Microeconomics*, 3 (4), 1–34.

Atkinson, John W. (1957), "Motivational Determinants of Risk-Taking Behavior," *Psychological Review*, 64 (6, Pt.1), 359–72.

Austin, Patrick Lucas (2019), "Need Some Help Reaching Your Goals? Try These Five Habit-Tracking Apps." Last Accessed May 29, 2022. https://time.com/5621109/best-habit-tracking-apps/

Bagozzi, Richard P. and Rik Pieters (1998), "Goal-Directed Emotions," *Cognition & Emotion*, 12 (1), 1–26.

Bagozzi, Richard P. and Utpal Dholakia (1999), "Goal Setting and Goal Striving in Consumer Behavior," *Journal of Marketing*, 63 (4_Suppl1), 19–32.

Bandura, Albert (1977), "Self-Efficacy: Toward a Unifying Theory of Behavioral Change," *Psychological Review*, 84 (2), 191–215.

Bandura, Albert and Edwin A. Locke (2003), "Negative Self-Efficacy and Goal Effects Revisited," *The Journal of Applied Psychology*, 88 (1), 87–99.

Barasz, Kate, Leslie K. John, Elizabeth A. Keenan, and Michael I. Norton (2017), "Pseudo-Set Framing," *Journal of Experimental Psychology. General*, 146 (10), 1460–77.

Belk, Russell W. (1988), "Possessions and the Extended Self," *Journal of Consumer Research*, 15 (2), 139–68.

Bitterly, T. Bradford, Robert Mislavsky, Hengchen Dai, Katherine L. Milkman. (2015), "Want-Should Conflict: A Synthesis of past Research," in *The Psychology of Desire*, ed. W. Hoffman and L. Nordgren, New York, NY: Guilford Press, 244–64.

Cacioppo, John T. and Richard E. Petty (1982), "The Need for Cognition," *Journal of Personality and Social Psychology*, 42 (1), 116–31.

Carlson, Kurt A. and Suzanne B. Shu (2007), "The Rule of Three: How the Third Event Signals the Emergence of a Streak," *Organizational Behavior and Human Decision Processes*, 104 (1), 113–21.

Cheema, Amar and Dilip Soman (2008), "The Effect of Partitions on Controlling Consumption," *Journal of Marketing Research*, 45 (6), 665–75.

Cochran, Winona and Abraham Tesser (1996), “The ‘What the Hell’ Effect: Some Effects of Goal Proximity and Goal Framing on Performance,” in *Striving and Feeling: Interactions among Goals, Affect, and Self-Regulation*, ed. L. L. Martin and A. Tesser, Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 99–120.

Dai, Hengchen, Katherine L. Milkman, and Jason Riis (2014), “The Fresh Start Effect: Temporal Landmarks Motivate Aspirational Behavior,” *Management Science*, 60 (10), 2563–82.

de León, Riley (2020), “The Education App that is Making Equals of Bill Gates and the World’s Masses,” Last Accessed May 29, 2022. https://www.cnbc.com/2020/11/19/the-education-app-making-equals-of-bill-gates-and-the-worlds-masses.html

Deighton, John, Jacob Goldenberg, and Andrew T. Stephen (2017), “Introduction to Special Issue: The Consumer in a Connected World,” *Journal of the Association for Consumer Research*, 2 (2), 137–39.

Dhar, Ravi and Itamar Simonson (1999), “Making Complementary Choices in Consumption Episodes: Highlighting versus Balancing,” *Journal of Marketing Research*, 36 (1), 29–44.

Dhar, Ravi, Joel Huber, and Uzma Khan (2007), “The Shopping Momentum Effect,” *Journal of Marketing Research*, 44 (3), 370–78.

Etkin, Jordan (2016), “The Hidden Cost of Personal Quantification,” *Journal of Consumer Research*, 42 (6), 967–84.

Evers, Ellen, Yoel Inbar, and M. Marcel Zeelenberg (2014), “Set-Fit Effects in Choice,” *Journal of Experimental Psychology. General*, 143 (2), 504–9.

Fishbach, Ayelet, Rebecca K. Ratner, and Ying Zhang (2011), “Inherently Loyal or Easily Bored?: Nonconscious Activation of Consistency versus Variety-Seeking Behavior,” *Journal of Consumer Psychology*, 21 (1), 38–48.

Fox, Susannah (2013), “The Self-Tracking Data Explosion,” Last Accessed May 29, 2022. http://www.pewinternet.org/2013/06/04/the-self-tracking-data-explosion/

Fredrickson, Barbara L. (2001), “The Role of Positive Emotions in Positive Psychology: The Broaden-and-Build Theory of Positive Emotions,” *American Psychologist*, 56 (3), 218–26.

Galak, Jeff and Joseph P. Redden (2018), “The Properties and Antecedents of Hedonic Decline,” *Annual Review of Psychology*, 69, 1–25.

Gao, Leilei, Yanliu Huang, and Itamar Simonson (2014), “The Influence of Initial Possession Level on Consumers’ Adoption of a Collection Goal: A Tipping Point Effect,” *Journal of Marketing*, 78 (6), 143–56.

Gelman, Andrew and Hal Stern (2006), “The Difference between Significant and Not Significant Is Not Itself Statistically Significant,” *The American Statistician*, 60 (4), 328–31.

Gilovich, Thomas, Robert Vallone, and Amos Tversky (1985), “The Hot Hand in Basketball: On the Misperception of Random Sequences,” *Cognitive Psychology*, 17 (3), 295–314.

Hamari, Juho, Jonna Koivisto, and Harri Sarsa (2014), “Does Gamification Work? A Literature Review of Empirical Studies on Gamification,” in *2014 47th Hawaii International Conference on System Sciences*, ed. R. H. Sprague Jr, Waikoloa, HI: IEEE, 3025–3034.

Hayes, Andrew (2017), *Introduction to Mediation, Moderation, and Conditional Process Analysis. Second Edition, a Regression-Based Approach*, New York, NY: Guilford Press.

Heath, Chip, Richard P. Larrick, and George Wu (1999), “Goals as Reference Points,” *Cognitive Psychology*, 38 (1), 79–109.

Hollenbeck, John R., Charles R. Williams, and Howard J. Klein (1989), “An Empirical Examination of the Antecedents of Commitment to Difficult Goals,” *Journal of Applied Psychology*, 74 (1), 18–23.

Hoyer, Wayne D. (1984), “An Examination of Consumer Decision Making for a Common Repeat Purchase Product,” *Journal of Consumer Research*, 11 (3), 822–29.

Hsee, Christopher K., Fang Yu, Jiao Zhang, and Yan Zhang (2003), “Medium Maximization,” *Journal of Consumer Research*, 30 (1), 1–14.

Hsee, Christopher K., Yang Yang, and Bowen Ruan (2015), “The Mere-Reaction Effect: Even Nonpositive and Noninformative Reactions Can Reinforce Actions,” *Journal of Consumer Research*, 42 (3), ucv022.

Karapanos, Evangelos, Rúben Gouveia, Marc Hassenzahl, and Jodi Forlizzi (2016), “Wellbeing in the Making: Peoples’ Experiences with Wearable Activity Trackers,” *Psychology of Well-Being*, 6 (1), 4–17.

Kausler, Donald H. and Malekeh K. Hakami (1983), “Memory for Activities: Adult Age Differences and Intentionality,” *Developmental Psychology*, 19 (6), 889–94.

Keinan, Anat and Ran Kivetz (2011), “Productivity Orientation and the Consumption of Collectable Experiences,” *Journal of Consumer Research*, 37 (6), 935–50.

Kivetz, Ran, Oleg Urminsky, and Yuhuang Zheng (2006), “The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention,” *Journal of Marketing Research*, 43 (1), 39–58.

Latham, Gary P. and Edwin A. Locke (2006), “Enhancing the Benefits and Overcoming the Pitfalls of Goal Setting,” *Organizational Dynamics*, 35 (4), 332–40.

Leskin, Paige (2019), “Snapchat Users Are So Upset About Losing Their Streaks That They Email the Company to Get Them Back,” Last Accessed May 29, 2022. https://www.businessinsider.com/snapchat-streaks-how-to-get-snapstreak-back-2019-7

Liberman, Nira, Yaacov Trope, and Elena Stephan (2007), “Psychological Distance,” *Social Psychology: Handbook of Basic Principles*, 2, 353–83.

Locke, Edwin A. and Gary P. Latham (1990), *A Theory of Goal Setting & Task Performance*, Englewood Cliffs, NJ: Prentice-Hall, Inc.

Locke, Edwin A., Norman Cartledge, and Jeffrey Koeppel (1968), “Motivational Effects of Knowledge of Results: A Goal-Setting Phenomenon?,” *Psychological Bulletin*, 70 (6, Pt.1), 474–85.

Lorenz, Taylor (2017), “Teens Explain the World of Snapchat’s Addictive streaks, Where Friendships Live or Die,” Last Accessed May 29, 2022. https://www.businessinsider.com/teens-explain-snapchat-streaks-why-theyre-so-addictive-and-important-to-friendships-2017-4

Lynkova, D. (2020), “Top Gamification Statistics of 2020: Next Level Gaming,” https://review42.com/gamification-statistics/

Malcolm, Hadley (2015), “Budgeting Becoming Popular Thanks to Apps,” https://www.usatoday.com/story/money/2015/04/27/budgeting-apps-affect-spending-habits/26190991/

Maltseva, K. and C. Lutz (2018), “A Quantum of Self: A Study of Self-Quantification and Self-Disclosure,” *Computers in Human Behavior*, 81, 102–14.

Markus, Hazel and Elissa Wurf (1987), “The Dynamic Self-Concept: A Social Psychological Perspective,” *Annual Review of Psychology*, 38 (1), 299–337.

Mazar, Nina, On Amir, and Dan Ariely (2008), “More Ways to Cheat-Expanding the Scope of Dishonesty,” *Journal of Marketing Research*, 45 (6), 633–53.

McClelland, David C. (1961), *The Achieving Society*, Princeton, NJ: Van Nostrand.

Menon, Satya and Barbara E. Kahn (1995), "The Impact of Context on Variety Seeking in Product Choices," *Journal of Consumer Research*, 22 (3), 285–95.

Moldovanu, Benny, Aner Sela, and Xianwen Shi (2007), "Contests for Status," *Journal of Political Economy*, 115 (2), 338–63.

Neuberg, Steven L. and Jason T. Newsom (1993), "Personal Need for Structure: Individual Differences in the Desire for Simpler Structure," *Journal of Personality and Social Psychology*, 65 (1), 113–31.

Nield, David (2019), "The Best Apps for Changing Your Habits," Last Accessed May 29, 2022. https://www.popsci.com/best-apps-tracking-habits/

Nunes, Joseph C. and Xavier Drèze (2006), "Your Loyalty Program Is Betraying You," *Harvard Business Review*, 84 (4), 124–31; 150.

Ouellette, Judith and Wendy Wood (1998), "Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior," *Psychological Bulletin*, 124 (1), 54–74.

Redden, Joseph P. (2008), "Reducing Satiation: The Role of Categorization Level," *Journal of Consumer Research*, 34 (5), 624–34.

Reed, Americus, Mark R. Forehand, Stefano Puntoni, and Luk Warlop (2012), "Identity-Based Consumer Behavior," *International Journal of Research in Marketing*, 29 (4), 310–21.

Rogers, Todd, Katherine L. Milkman, Leslie K. John, and Michael I. Norton (2015), "Beyond Good Intentions: Prompting People to Make Plans Improves Follow-Through on Important Tasks," *Behavioral Science & Policy*, 1 (2), 33–41.

Rogers, Michelle (2020), "Ten Exercises You Can Do from Your Couch." Last Accessed May 29, 2022. https://blog.bcbsnc.com/2020/04/10-exercises-you-can-do-from-your-couch/

Schmall, Tyler (2019), "This Is Why Most Americans Don't Exercise More," https://nypost.com/2019/01/13/this-is-why-most-americans-dont-exercise-more/

Schmitt, Bernd (2019), "From Atoms to Bits and Back: A Research Curation on Digital Technology and Agenda for Future Research," *Journal of Consumer Research*, 46 (4), 825–32.

Schneider, Walter and Richard M. Shiffrin (1977), "Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention," *Psychological Review*, 84 (1), 1–66.

Schunk, Dale H. (1989), "Social Cognitive Theory and Self-Regulated Learning," in *Self-Regulated Learning and Academic Achievement*, ed. B. J. Zimmerman and D. Schunk, New York, NY: Springer, 83–110.

Schweidel, David A. and Wendy W. Moe (2016), "Binge Watching and Advertising," *Journal of Marketing*, 80 (5), 1–19.

Scott, Maura L. and Stephen M. Nowlis (2013), "The Effect of Goal Specificity on Consumer Goal Reengagement," *Journal of Consumer Research*, 40 (3), 444–59.

Sheeran, Paschal, Thomas L. Webb, and Peter M. Gollwitzer (2005), "The Interplay between Goal Intentions and Implementation Intentions," *Personality & Social Psychology Bulletin*, 31 (1), 87–98.

Singer, Jerome E. (1966), "Motivation for Consistency," in *Cognitive Consistency: Motivational Antecedents and Behavior Consequents*, ed. S. Feldman, New York: Academic Press, 47–73.

Sjöklint, Mimmi, Ioanna D. Constantiou, and Matthias Trier (2015), "The Complexities of Self-Tracking-An Inquiry into User Reactions and Goal Attainment," Available at SSRN 2611193.

Soman, Dilip and Amar Cheema (2004), "When Goals Are Counterproductive: The Effects of Violation of a Behavioral Goal on Subsequent Performance," *Journal of Consumer Research*, 31 (1), 52–62.

Spence, Kenneth Wartenbee (1956), *Behavior Theory and Conditioning*. New Haven, CT: Yale University Press, 35.

Spiller, Stephen A. (2011), "Opportunity Cost Consideration," *Journal of Consumer Research*, 38 (4), 595–610.

Steele, Claude M. (1988), "The Psychology of Self-Affirmation: Sustaining the Integrity of the Self," in *Advances in Experimental Social Psychology*, Vol. 21, ed. L. Berkowitz, Hillsdale, NJ: Erlbaum, 261–302.

Taylor, Shelley E. and Susan T. Fiske (1978), "Salience, Attention, and Attribution: Top of the Head Phenomena," *Advances in Experimental Social Psychology*, 11, 249–88.

Thompson, Charles P. (1982), "Memory for Unique Personal Events: The Roommate Study," *Memory & Cognition*, 10 (4), 324–32.

Tolli, Adam P. and Aaron M. Schmidt (2008), "The Role of Feedback, Causal Attributions, and Self-Efficacy in Goal Revision," *The Journal of Applied Psychology*, 93 (3), 692–701.

Tversky, Amos and Daniel Kahneman (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47 (2), 263–92.

Tykocinski, Orit E., Thane S. Pittman, and Erin E. Tuttle (1995), "Inaction Inertia: Foregoing Future Benefits as a Result of an Initial Failure to Act," *Journal of Personality and Social Psychology*, 68 (5), 793–803.

Venkatesan, Rajkumar and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (4), 106–25.

Verplanck, William S. (1956), "The Operant Conditioning of Human Motor Behavior," *Psychological Bulletin*, 53 (1), 70–83.

Vogels, Emily A. (2020), "About One-in-Five Americans Use a Smart Watch or Fitness Tracker." Last Accessed May 29, 2022. https://www.pewresearch.org/fact-tank/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/

Wallace, Scott G. and Jordan Etkin (2018), "How Goal Specificity Shapes Motivation: A Reference Points Perspective," *Journal of Consumer Research*, 44 (5), 1033–51.

Yang, Haiyang, Antonios Stamatogiannakis, and Amitava Chattopadhyay (2015), "Pursuing Attainment versus Maintenance Goals: The Interplay of Self-Construal and Goal Type on Consumer Motivation," *Journal of Consumer Research*, 42 (1), 93–108.

Zepeda, Lydia and David Deal (2008), "Think before You Eat: photographic Food Diaries as Intervention Tools to Change Dietary Decision Making and Attitudes," *International Journal of Consumer Studies*, 32 (6), 692–8.

Zuckerman, Miron (1979), "Attribution of Success and Failure Revisited, or: The Motivational Bias Is Alive and Well in Attribution Theory," *Journal of Personality*, 47 (2), 245–87.