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

Jackie Silverman

Alixandra Barasch

**Web Appendix**

Table of Contents

*Part A: Supplemental Analyses and Materials*

[**Additional Models (Study 1)** 2](#_Toc92535352)

[**App Screenshot and Demographic Information (Study 1)** 4](#_Toc92535353)

[**Indicators of Goal Pursuit (Studies 4, 5, and 7)** 6](#_Toc92535354)

[Items 6](#_Toc92535355)

[Analyses 7](#_Toc92535359)

[**Information about Attrition (Studies 5 and 6)** 13](#_Toc92535363)

[**Full Sample Analyses (Study 6)** 14](#_Toc92535366)

*Part B: Supplemental Studies and Post-Test*

[**Supplemental Study 1** 15](#_Toc92535367)

[**Supplemental Studies 2a and 2b** 19](#_Toc92535368)

[Study S2a 19](#_Toc92535369)

[Study S2b 21](#_Toc92535370)

[**Supplemental Studies 3a and 3b** 25](#_Toc92535371)

[Study S3a 26](#_Toc92535372)

[Study S3b 29](#_Toc92535373)

[**Supplemental Study 4** 35](#_Toc92535374)

[**Post-Test** 41](#_Toc92535375)

# **Additional Models (Study 1)**

We ran several additional models to further investigate if app users were more likely to meet the challenge on a given day following an intact logged streak, and were more likely to meet the challenge on a given day following a broken logged streak. First, we found that the effects held when we used a continuous measure to represent the length of the (broken) streaks, rather than defining the streak by a specific cut-off value (*b* = 0.04, SE = 0.01, *Z* = 4.22, *p* < .001).

Additional analyses demonstrated further robustness of the effects of intact and broken streaks. In particular, as shown in Table S1a, individuals were more likely to meet the challenge on a given day when they had an intact streak, versus broken streak, for a wide variety of streak length cut off values (from 2 to 20; *Z*s > 2.95, *p*s < .01). Moreover, as shown in Tables S1b and S1c, having an intact (broken) streak significantly increased (decreased) an individual’s likelihood of meeting the challenge on a given day for a wide variety of streak lengths, relative to other patterns (from 2 to 20; *Z*s > 2.90, *p*s < .01). These effects held when we accounted for weekend effects (with a dummy variable indicating weekends (1) vs. weekdays (0)) and day-of-challenge effects (with a continuous variable indicating the day of the challenge; 1 to 30).Finally, the effect of a broken streak (vs. not) held when our model included a “state-dependence” parameter for a broken streak (i.e., controlling for any unobserved factors that could have reduced engagement in the target behavior on a given day, independent of whether it was preceded by a streak; *b* = -0.38, SE = 0.08, *Z* = 4.91, *p* < .001).

*Table S1a: Modelling the effects of intact vs. broken streaks*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | *Streak length cut-off (days)* | | | | | | |
| *2* | *3* | *4* | *5* | *10* | *15* | *20* |
| Intercept | -1.87\*\*\* | -2.31\*\*\* | -2.30\*\*\* | -2.31\*\*\* | -2.30\*\*\* | -2.32\*\*\* | -2.34\*\*\* |
| **Streak Indicator (1 = intact, -1 = broken)** | **0.29\*\*** | **0.39\*\*\*** | **0.37\*\*\*** | **0.32\*\*\*** | **0.56\*\*\*** | **0.75\*\*\*** | **0.92\*\*** |
| Rate of Meeting Goal (0 to 1) | 4.58\*\*\* | 5.68\*\*\* | 5.73\*\*\* | 5.79\*\*\* | 5.80\*\*\* | 5.85\*\*\* | 5.90\*\*\* |
| Weekend Dummy Variable | -0.88\*\*\* | -0.87\*\*\* | -0.88\*\*\* | -0.88\*\*\* | -0.86\*\*\* | -0.86\*\*\* | -0.85\*\*\* |
| Day variable (1 to 30) | -0.02\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* |

*Table S1b: Modeling the effects of intact streaks separately*

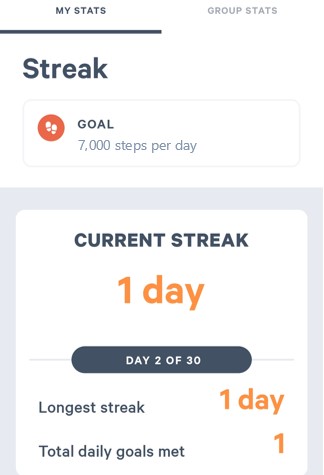
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | *Streak length cut-off (days)* | | | | | | |
| *2* | *3* | *4* | *5* | *10* | *15* | *20* |
| Intercept | -2.44\*\*\* | -2.31\*\*\* | -2.31\*\*\* | -2.33\*\*\* | -2.29\*\*\* | -2.31\*\*\* | -2.34\*\*\* |
| **Intact Streak Dummy Variable** | **0.81\*\*\*** | **0.36\*\*\*** | **0.30\*\*\*** | **0.22\*\*\*** | **0.51\*\*\*** | **0.75\*\*\*** | **0.90\*\*\*** |
| Rate of Meeting Goal (0 to 1) | 5.26\*\*\* | 5.66\*\*\* | 5.75\*\*\* | 5.82\*\*\* | 5.80\*\*\* | 5.84\*\*\* | 5.89\*\*\* |
| Weekend Dummy Variable | -0.85\*\*\* | -0.89\*\*\* | -0.90\*\*\* | -0.88\*\*\* | -0.87\*\*\* | -0.86\*\*\* | -0.85\*\*\* |
| Day variable (1 to 30) | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* |

*Table S1c: Modeling the effects of broken streaks separately*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | *Streak length cut-off (days)* | | | | | | |
| *2* | *3* | *4* | *5* | *10* | *15* | *20* |
| Intercept | -1.87\*\*\* | -2.40\*\*\* | -2.41\*\*\* | -2.42\*\*\* | -2.41\*\*\* | -2.40\*\*\* | -2.39\*\*\* |
| **Broken Streak Dummy Variable** | **-0.29\*\*\*** | **-0.81\*\*\*** | **-0.82\*\*\*** | **-0.79\*\*\*** | **-0.93\*\*\*** | **-0.90\*\*\*** | **-1.11\*\*** |
| Rate of Meeting Goal (0 to 1) | 4.58\*\*\* | 6.04\*\*\* | 6.04\*\*\* | 6.03\*\*\* | 6.00\*\*\* | 5.98\*\*\* | 5.98\*\*\* |
| Weekend Dummy Variable | -0.88\*\*\* | -0.80\*\*\* | -0.81\*\*\* | -0.84\*\*\* | -0.85\*\*\* | -0.86\*\*\* | -0.86\*\*\* |
| Day variable (1 to 30) | -0.02\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* | -0.03\*\*\* |

**Tables S1a-c.**Robustness tests (Study 1): model coefficients when accounting for day and weekend effects when predicting the effects of intact and broken streaks relative to each other (a) and separately (b and c) on subsequent stepping behavior.

# **App Screenshot and Demographic Information (Study 1)**



**Figure S1.** Screenshot from the fitness app, depicting a users’ streak of meeting the challenge (here, 1 day). When participants broke their streak or otherwise missed the challenge, the behavioral log simply informed them that they had a current streak of “0 days.”

|  |  |  |
| --- | --- | --- |
| **Demographic Variable** | **Options** | **Percentage of Users** |
| **Campus** | Main campus | 73% |
| Satellite campus 1 | 16% |
| Satellite campus 2 | 6% |
| Satellite campus 3 | 5% |
| **Reason for Joining** | “A colleague encouraged me to participate” | 52% |
| “I decided to participate myself” | 48% |
| **Recording Method** | Device or App | 87% |
| Manually | 10% |
| N/A (did not record any step data) | 3% |
| **Device or App Used**  **(of the 87% Using a Device or App)** | Apple product | 61% |
| FitBit product | 25% |
| Google product | 10% |
| Garmin product | 3% |
| Misfit app | <1% |
| Withings app | <1% |

**Table S2.** Additional demographics collected by the partner university’ wellness program and associated summary statistics.

# **Indicators of Goal Pursuit (Studies 4, 5, and 7)**

## Items

Feelings of achievement items

*(Studies 4, 5, and 7)*

* How much did you feel that you accomplished something? 1
* How much did you feel like you had achieved something?
* How proud did you feel?
* How successful did you feel?
* How effective did you feel?

### *Motivation items*

*(Studies 4, 5, and 7)*

* How motivated did you feel to continue [target behavior2]? 3
* How determined were you to continue [target behavior]?
* How much did your past behavior push you to stick with [target behavior] for the rest of this study?
* How much did you feel driven to continue [target behavior]?
* How much did you feel compelled to continue [target behavior]?

### Negative emotion

*(Studies 5 and 7)*

* How angry did you feel?
* How upset did you feel?
* How disappointed did you feel?
* How annoyed did you feel?

*Notes*

1. This item was inadvertently left out of Study 5.
2. The ‘target behavior’ in the question text differed across studies:
   1. *Study 4:* completing language learning questions
   2. *Study 5:* playing another game
   3. *Study 7:* playing another [Sum / Jumble]
3. This item was inadvertently left out of Study 7.

All items are on 11-point scales (from “Not at all” to “Extremely”/“A great deal”)

## Analyses

### Study 4

#### Effects of streak condition(s)

*Feelings of achievement.* A one-way ANOVA revealed a main effect of *streak* condition on feelings of achievement (*F*(2, 598) = 11.57, *p* < .001). Separate t-tests revealed that participants in the *intact logged streak* condition had greater feelings of achievement (*M* = 7.59, SD = 2.65) than participants in the *broken logged streak* condition (*M* = 6.42, SD = 2.61; *t*(398) = 4.45, *p* < .001, *d* = 0.45). Additionally, the *no log* condition fell between these two conditions (*M* = 7.32, SD = 2.38); participants in the *intact logged streak* condition had directionally greater feelings of achievement than participants in the *no log* condition (*t*(401) = 1.07, *p* = .29, *d* = 0.11), while participants in the *broken logged streak* condition felt significantly less achievement (*t*(397) = 3.62, *p* < .001, *d* = 0.36).

*Motivation.* A one-way ANOVA revealed a main effect of *streak* condition on motivation (*F*(2, 598) = 4.21, *p* = .015). Participants in the *intact logged streak* condition were more motivated (*M* = 7.44, SD = 3.05) than participants in the *broken logged streak* condition (*M* = 6.63, SD = 2.92; *t*(398) = 2.70, *p* = .007, *d* = 0.27). The *no log* condition fell between these two conditions (*M* = 7.25, SD = 2.74); participants in the *intact logged streak* condition felt directionally more motivated than participants in the *no log* condition (*t*(401) = 0.63, *p* = .53, *d* = 0.07), while participants in the *broken logged streak* condition felt significantly less motivated (*t*(397) = 2.20, *p* = .029, *d* = 0.22).

#### Mediation analyses

*Relative to the absence of a behavioral log.* In two exploratory models, we tested whether the extent to which participants reported adopting a streak maintenance goal mediated the effects of intact and broken logged streaks on subsequent behavior separately, relative to not having a behavioral log (i.e., the *no log* condition). We found that this *goal adoption* measure did not mediate the negative effect of having a broken streak of logging (Indirect effect = -.04, SE = .04, 95% CI = [-.14, .04]) or the positive effect of having an intact streak of logging (Indirect effect = .14, SE = .09, 95% CI = [-.03, .33]).

We also tested if participants’ sense of accomplishment (the combined measure of our *feelings of achievement* and *motivation* subscales) mediated the positive (negative) relationship between having an intact (broken) streak and not having a behavioral log (i.e., the *no log* condition) on subsequent behavior. We found that sense of accomplishment mediated the negative effect between having a broken logged streak (1) and not having a behavioral log on the app (0; Indirect effect = -.19, SE = .07, 95% CI = [-.36, -.07]), but not the effect between having an intact logged streak (1) and not having a behavioral log (0; Indirect effect = .09, SE = .09, 95% CI = [-.09, .26]).

*Serial mediation.* We conducted a serial mediation (PROCESS Model 6) to examine if the effect of an intact (vs. broken) logged streak on subsequent behavior was mediated by participants’ feelings of achievement (Mediator 1), which in turn influenced their motivation (Mediator 2). Consistent with our theory, the indirect effect of the mediation pathway through both mediators excluded zero (Indirect effect = .50, SE = .13, 95% CI = [.28, .79]).

### Study 5

#### Effects of streak condition(s)

*Feelings of achievement.* A one-way ANOVA revealed a main effect of *streak* condition on feelings of achievement (*F*(2, 449) = 15.39, *p* < .001). Separate t-tests revealed that participants in the *intact logged streak* condition had greater feelings of achievement (*M* = 8.80, SD = 1.89) than participants in the *broken logged streak* condition (*M* = 7.50, SD = 2.50; *t*(306) = 5.14, *p* < .001, *d* = 0.59). Additionally, the *no log* condition fell between these two conditions (*M* = 8.48, SD = 1.99); participants in the *intact logged streak* condition had directionally greater feelings of achievement than participants in the *no log* condition (*t*(291) = 1.43, *p* = .153, *d* = 0.17), while participants in the *broken logged streak* condition felt significantly less achievement (*t*(301) = 3.75, *p* < .001, *d* = 0.43).

*Motivation.* A one-way ANOVA revealed a main effect of *streak* condition on motivation (*F*(2, 449) = 4.42, *p* = .013). Participants in the *intact logged streak* condition were more motivated (*M* = 8.24, SD = 2.34) than participants in the *broken logged streak* condition (*M* = 7.41, SD = 2.68; *t*(306) = 2.89, *p* = .004, *d* = 0.33). The *no log* condition fell between these two conditions (*M* = 7.94, SD = 2.43); participants in the *intact logged streak* condition felt directionally more motivated than participants in the *no log* condition (*t*(291) = 1.06, *p* = .29, *d* = 0.13), while participants in the *broken logged streak* condition felt marginally less motivated (*t*(301) = 1.82, *p* = .070, *d* = 0.21).

*Negative emotion.* A one-way ANOVA revealed a main effect of *streak* condition on negative emotion (*F*(2, 449) = 21.18, *p* < .001). Participants in the *intact logged streak* condition felt less negative emotion (*M* = 2.81, SD = 2.87) than participants in the *broken logged streak* condition (*M* = 5.05, SD = 3.51; *t*(306) = 6.13, *p* < .001, *d* = 0.70). The *no log* condition fell between these two conditions (*M* = 3.25, SD = 3.22); participants in the *intact logged streak* condition felt directionally less negative emotion than participants in the *no log* condition (*t*(291) = 1.26, *p* = .210, *d* = 0.14), while participants in the *broken logged streak* condition felt significantly more negative emotion (*t*(301) = 4.64, *p* < .001, *d* = 0.53).

#### Mediation analyses

*Relative to the absence of a behavioral log.* Two exploratory models tested if participants’ sense of accomplishment mediated the positive (negative) relationship between having an intact (broken) streak and not having a behavioral log (i.e., the *no log* condition) on subsequent behavior. We found that sense of accomplishment mediated the negative effect between having a broken logged streak (1) and not having a behavioral log on the app (0; Indirect effect = -.42, SE = .16, 95% CI = [-.77, -.13]), but not the effect between having an intact logged streak (1) and not having a behavioral log (0; Indirect effect = .15, SE = .13, 95% CI = [-.08, .42]).

*Serial mediation.* We conducted a serial mediation (PROCESS Model 6) to examine if the effect of an intact (vs. broken) logged streak on subsequent behavior was mediated by participants’ feelings of achievement (Mediator 1), which in turn influenced their motivation (Mediator 2). Consistent with our theory, the indirect effect of the mediation pathway through both mediators excluded zero (Indirect effect = .45, SE = .13, 95% CI = [.24, .73]).

*Negative emotion.* In an additional mediation analysis, we did not find evidence that negative emotion mediated the effect of an intact (1) versus broken (0) logged streak on subsequent (Indirect effect = .11, SE = .10, 95% CI = [-.07, .32]).

### Study 7

#### Effects of streak condition(s)

*Feelings of achievement.* A one-way ANOVA revealed a main effect of *streak* condition on feelings of achievement (*F*(2, 598) = 10.92, *p* < .001). Participants in the *intact streak* condition had greater feelings of achievement (*M* = 8.68, SD = 2.01) than participants in the *broken streak* condition (*M* = 7.72, SD = 2.22; *t*(403) = 4.57, *p* < .001, *d* = 0.45) and participants in the *repairable broken streak* condition (*M* = 7.96, SD = 2.22; *t*(398) = 3.38, *p* = .001, *d* = 0.34). There was not a significant difference between the two *broken streak* conditions (*t*(395) = 1.11, *p* = .27, *d* = 0.11).

*Motivation.* A one-way ANOVA revealed a main effect of *streak* condition on motivation (*F*(2, 598) = 19.47, *p* < .001). Participants in the *intact streak* condition were more motivated (*M* = 8.40, SD = 2.20) than participants in the *broken streak* condition (*M* = 6.94, SD = 3.04; *t*(403) = 5.53, *p* < .001, *d* = 0.55). Participants in the *repairable broken streak* condition (*M* = 8.25, SD = 2.43) were more motivated than participants in the *broken streak* condition (*t*(395) = 4.75 *p* < .001, *d* = 0.48), and similarly motivated as participants in the *intact streak* condition (*t*(398) = 0.62, *p* = .54, *d* = 0.07).

*Negative emotion.* A one-way ANOVA revealed a main effect of *streak* condition on negative emotion (*F*(2, 598) = 12.62, *p* < .001). Participants in the *intact streak* condition felt less negative emotion (*M* = 2.11, SD = 2.33) than participants in the *broken streak* condition (*M* = 3.16, SD = 2.77; *t*(403) = 4.15, *p* < .001, *d* = 0.41) and participants in the *repairable broken streak* condition (*M* = 3.33, SD = 2.83; *t*(398) = 4.72, *p* < .001, *d* = 0.47). There was not a significant difference between the two *broken streak* conditions (*t*(395) = 0.59, *p* = .56, *d* = 0.06).

#### Mediation analyses

*Serial mediation.* We conducted a serial mediation (PROCESS Model 6) to examine if the effect of an intact (vs. broken) streak on subsequent behavior was mediated by participants’ feelings of achievement (Mediator 1), which in turn influenced their motivation (Mediator 2). Consistent with our theory, the indirect effect of the mediation pathway through both mediators excluded zero (Indirect effect = .33, SE = .09, 95% CI = [.17, .52]).

*Negative emotion.* In an additional mediation analysis, we did not find evidence that negative emotion mediated the effect of an intact (1) versus broken (0) streak on subsequent (Indirect effect = -.08, SE = .06, 95% CI = [-.21, .02]).

*Relative to the repairable broken streak condition.* In two additional models, sense of accomplishment also mediated a) the positive effect of an intact streak (Indirect effect = .13, SE = .07, 95% CI = [.02, .31]) and b) the negative effect of a broken streak (Indirect effect = -.16, SE = .06, 95% CI = [-.32, -.06]), relative to a broken but repairable streak.

# **Information about Attrition (Studies 5 and 6)**

## Study 5

Out of the 805 participants who were assigned to a condition at the start of the study, 353 participants (43.85%) chose to stop playing games before our manipulation at the fourth game, and thus could not be included in our final sample or analyses. Of these participants who did not reach our manipulation, 130 (36.83%) completed one game, 113 (32.01%) completed two games, 109 (30.88%) completed three games, and one (0.28%) completed zero games (i.e., left the study during the instructions). The number of games participants chose to play did not differ by condition (χ2 (df = 6) = 7.11, *p* = .31).

## Study 6

Out of the 802 participants assigned to condition, we excluded from our analyses 384 participants (47.88%) who do not correctly complete all three of the games before our manipulation at the fourth game, as preregistered. Of these participants, 48 (12.50%) completed zero correctly, 122 (31.77%) completed one correctly, and 214 (55.73%) completed two correctly. The number of correctly completed games did not differ by condition (χ2 (df = 6) = 5.48, *p* = .48).

# **Full Sample Analyses (Study 6)**

*Target behavior.* A binary logit revealed an effect of *streak* condition (*F*(2, 799) = 11.63, *p* < .001). Replicating previous studies, a chi-square analysis showed that participants in the *intact streak* condition were more likely to engage in the target behavior (44.53%) compared to participants in the *externally-attributed broken streak* condition (34.32%; *Χ2*(1) = 5.85, *p* = .016, *OR* = 1.54). Critically, this difference was even larger when comparing the *intact streak* condition to the *self-attributed broken streak* condition (24.44%; *Χ2*(1) = 23.73, *p* < .001, *OR* = 2.48). 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 (*Χ2*(1) = 6.31, *p* = .012, *OR =* 0.62). Thus, participants were even more likely to engage in the target behavior when they attributed the break to an external force, rather than to themselves.

Supplemental Study 1

Past work has shown that people perceive a streak as soon as there are at least three behaviors in a row (Carlson and Shu 2007). In addition, people perceive four, five, and six behaviors in a row to be just as “streaky” as three behaviors in a row. In this study, we explored consumers’ perceptions of streakiness across a wide variety of patterns of behavior to 1) replicate the prior findings in a consumer behavior context, 2) further understand how perceptions of streakiness change with streak length, 3) examine the effects of novel patterns on perceived streakiness, and 4) explore the effect of different “miss” timings on perceived streakiness.

**Methods**

We recruited 167 participants (*M* age = 19.98, 61.08% female) from a behavioral lab at a private northeastern university. Sample size was determined by how many participants were recruited for the lab session.

Participants first read some information about an app (modeled on Duolingo) that helps users learn a new language by providing interactive lessons of progressing difficulty in reading, writing, and speaking a selected language. Participants saw sample images from the app and read a few sentences about its purposes and features. In particular, participants were told that the app tracks and shows consumers which days they complete lessons via a behavioral log.

Then, participants were told to imagine that they had been using the app regularly for the past few months and had advanced through multiple lessons in the language they wanted to learn. Participants were then told they would see several patterns of their app use through the behavioral log on the app, and to imagine that the last day of each pattern was yesterday.

Participants then saw ten randomly-selected patterns of recent behavior as displayed on the app’s behavioral log, out of 25 patterns depicted below (see Table S3). These patterns included a wide variety of streak lengths (one, two, three, four, five, ten, and twenty behaviors in a row). They also featured different types of patterns: non-streaks (e.g., one day), intact streaks, and recently broken streaks. For each pattern, participants rated how much that pattern made them feel like they currently had a streak on a 7-point scale (from 1 “Not at all” to 7 “Very much”). Lastly, participants answered basic demographic questions.

**Results**

*Length of pattern*. Paired t-tests reveal that, consistent with previous work (Carlson and Shu 2007), participants viewed three behaviors in a row as more streaky (*M* = 5.40, SD = 1.65) than one behavior (*M* = 2.60, SD = 1.68; *t*(131) = 9.67, *p* < .001, *d* = 1.68) and two behaviors in a row (*M* = 4.01, SD= 1.81; *t*(133) = 4.63, *p* < .001, *d* = 0.80). Moreover, while four behaviors in a row was perceived as more streaky (*M* = 5.99, SD= 1.29) than three behaviors in a row (*t*(133) = 2.31, *p* = .022, *d* = 0.40), there was no significant difference between four and five behaviors in a row (*M* = 6.21, SD= 1.15; *t*(132) = 1.06, *p* = .29, *d* = 0.26), or between ten (*M* = 6.78, SD= 0.57) and twenty behaviors in a row (*M* = 6.93, SD = 0.40; *t*(135) = 1.72, *p* = .088, *d* = 0.31). However, ten behaviors in a row was perceived as significantly more streaky than five behaviors in a row (*t*(134) = 3.71, *p* < .001, *d* = 0.63). Figure S2 displays the perceived streakiness of all patterns in the study.

*Pattern type*. Independent t-tests revealed that all intact streaks (either with no misses or a miss at Day 1) of three or more behaviors in a row were perceived as significantly more streaky than the midpoint of the scale (*t*s > 3.40, *p*s < .001, *d*s > 0.80). All other patterns (i.e., non-streaks, broken streaks, and broken streaks with a miss at Day 1) were not perceived as significantly different from the midpoint of the scale (or were significantly lower than the midpoint), suggesting that they were not perceived as streaky patterns.

**Figure S2.** Perceived streakiness of different patterns of behavior, as a function of streak length and pattern type. The grey line indicates the midpoint of the scale (4).

**Discussion**

Consistent with prior work, people perceived a pattern of logged behaviors to be streaky as soon as there were at least three behaviors in a row. This was true both for intact streaks without any misses and for intact streaks preceded by a miss at Day 1. Furthermore, perceived streakiness increased at a decreasing rate after three behaviors in a row and reached a maximum relatively quickly (e.g., no difference in perceived streakiness between four and five behaviors in a row).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | *Number of Behaviors* | | | | | | | |
|  |  | *1* | *2* | *3* | *4* | *5* | *10* | *20* |
| *Type of pattern* | *Non-streak* | ✓ | ✓✓ |  | ✓✓X✓✓ |  |  |  |
| *Intact Streak* |  |  | ✓✓✓ | ✓✓✓✓ | ✓✓✓✓✓ | ✓✓✓✓✓✓✓✓✓✓ | ✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓ |
| *Broken Streak* |  | ✓✓X | ✓✓✓X | ✓✓✓✓X | ✓✓✓✓✓X | ✓✓✓✓✓✓✓✓✓✓X | ✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓X |
| *Streak with Miss at Day 1* |  | X✓✓ | X✓✓✓✓ | X✓✓✓✓ | X✓✓✓✓✓ | X✓✓✓✓✓✓✓✓✓✓ | X✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓ |
| *Broken Streak with Miss at Day 1* |  |  | X✓✓✓X | X✓✓✓✓X | X✓✓✓✓✓X | X✓✓✓✓✓✓✓✓✓✓X | X✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓X |

**Table S3.** Patterns of behavior presented to participants in Study S1.

Supplemental Studies 2a and 2b

Building on Study 3, Studies S2a and S2b further test the effects of intact and broken logged streaks separately. In these studies, we held constant participants’ most recent behavior and only manipulated the presence of preceding streak. Specifically, in Study S2a, all participants engaged in the target behavior in the most recent period, and we simply manipulated whether that behavior contributed to an intact logged streak or not. Similarly, in Study S2b, all participants did *not* engage in the target behavior in the most recent period (i.e., had a “miss” due to a quota), and we manipulated whether that miss broke a logged streak or not. Thus, these studies allow us to examine the effect of an intact logged streak, above and beyond any positive effects of successfully logging a behavior (Study S2a) and the effect of a broken logged streak, above and beyond any negative effect of not being able to log a behavior (Study S2b). Furthermore, Study S2b also tests the effect of a broken logged streak controlling for the presence of a quota message.

## Study S2a: The Effect of Intact Logged Streaks

**Methods**

We recruited 156 participants (*M* age= 23.88, 62.82% female) in a behavioral lab at a northeastern university. Sample size was determined by the number of participants who signed up for the lab session.

Participants first 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 were then told that they were assigned to start playing Number Sums but would have the option to switch to Word Jumbles later. Similar to Studies 2 and 3, for every Number Sum game participants played, a checkmark would be added to the behavioral log featured at the top of the app, which was again modeled on real-life apps. However, for every Word Jumble they attempted, no checkmark would be added to their tracker. After completing comprehension check questions about the instructions, participants proceeded to the app.

Participants were randomly assigned to one of two between-subjects conditions. In the *no streak* control condition, participants first completed one Number Sum, and then chose what to complete for their second game (Number Sums or Word Jumbles). In the *intact streak* condition, participants first completed three Number Sums in a row, after which the behavioral log stated that they had a streak of completing Number Sums. Then, they had the choice to complete either a Number Sum or Word Jumble as their fourth game.

Our primary dependent variable was participants’ choice to continue with Number Sums (i.e., the target behavior) versus switch to Word Jumbles. Importantly, this design controlled for participants’ most recent behavior – everyone completed a Number Sum and received a checkmark in their behavioral log in the most recent time period – while only manipulating whether they had a preceding logged streak or not.

After deciding which game to complete next (but before playing it), all participants answered the same 10 *sense of accomplishment* items as Studies 4, 5, and 7 (α = .94). Participants also answered three questions about their individual app usage (e.g., if they had any similar apps downloaded on their phones), as well as a free response question about their choice of which game to play. Lastly, participants completed a manipulation check to indicate the extent to which they were aware that they had a streak of playing games (1 “Not aware at all” to 11 “Extremely aware”) and demographics.

**Results**

*Manipulation check.* An independent t-test confirmed that participants in the *intact streak* condition were more aware of having a streak of playing games (*M* = 7.32, SD = 3.30) than participants in the *no streak* condition (*M* = 5.35, SD = 3.28; *t*(154) = 3.75, *p* < .001, *d* = 0.60).

*Target behavior.* A chi-square test revealed that more participants engaged in the target behavior (i.e., completed Number Sums) when they had an intact streak (69.95%) than when they did not have a streak (48.72%; *Χ2* (1) = 5.93, *p* = .015, *OR* = 2.46; see Figure S3).

*Indicator of goal pursuit.* An independent t-test revealed that participants with an intact streak felt a greater sense of accomplishment(*M* = 6.67; SD = 2.42) than participants without a streak (*M* = 5.13, SD = 2.42; *t*(154) = 3.98, *p* < .001, *d* = 0.64; see Figure S4).

*Mediation analysis*. In accordance with our preregistration, we conducted a mediation analysis using a bootstrap procedure with 10,000 samples (Hayes 2017). We included *streak* condition (*intact streak* = 1; *no streak* = 0) as the independent variable, *sense of accomplishment* as the mediator, and game choice as the dependent variable. As predicted, sense of accomplishment mediated the effect of having a streak on engagement in the target behavior (Indirect effect = .88, SE = .30, 95% CI = [.40, 1.57]).

## Study S2b: The Effect of Broken Logged Streaks

**Methods**

We recruited 218 participants (*M* age= 20.07, 60.19% female) in a behavioral lab at a northeastern university. Sample size was determined by the number of participants who signed up for the lab session.

Participants read the same instructions as in Study S2a: they would be testing out an app by playing two different types of games (Number Sums and Word Jumbles). In addition, like in Studies 3, 6, and 7, all participants were informed that at some point, they might see a “quota” message when enough participants had already completed a specific game, and that this would mean that they were not be needed to test that particular game. The instructions emphasized that the quota message was not an indication of their gaming abilities or any kind of failure with the app, and that they would see this message at most once during the study. After completing a comprehension check about the instructions, participants proceeded to the app.

Participants were randomly assigned to one of two between-subjects conditions. In the *no streak* control condition, participants completed one Number Sum, and then encountered the quota message in place of the second game. In the *broken streak* condition, participants completed three Number Sums in a row and were notified that they had a streak of completing Number Sums. Then, they encountered the quota message in place of the fourth game.

After encountering the quota message, participants then chose which type of game they would like to complete next: either Number Sums (i.e., the target behavior) or Word Jumbles. As in Study S2a, this choice served as our primary dependent variable. Notably, all participants were able to complete at least one game successfully, and all participants saw the same quota message, giving them a “miss” on their behavioral log. Thus, this design controlled for participants’ most recent behavior, as well as the experience of seeing the quota message, and simply manipulated whether that miss broke a streak within their behavioral log or not.

After deciding but before continuing with the game of their choice, all participants answered the same *sense of accomplishment* measure as in Study S2a (α = .95). Participants also answered an attention-check question regarding how many games they had played before the quota message (83.94% answered correctly) and basic demographics.

**Results**

*Target behavior.* A chi-square test revealed that fewer participants engaged in the target behavior in the *broken streak* condition (37.61%) than in the *no streak* condition (64.22%; *Χ2* (1) = 14.39, *p* < .001, *OR* = 0.34; see Figure S3).

✓ ✓✓✓

✓X ✓✓✓X

Study S2b: Effect of Broken Streaks

**Figure S3.** Percent of participants who chose to engage in the target behavior as a function of their recent pattern of logged behavior. Error bars are 1 standard error.

Study S2a: Effect of Intact Streaks

*Indicator of goal pursuit.* An independent t-test revealed that participants in the *broken streak* condition felt a lower sense of accomplishment (*M* = 4.51, SD = 1.90) than participants in the *no streak* condition (*M* = 5.26, SD = 2.27; *t*(216) = 2.66, *p* = .008, *d* = 0.36; see Figure S4).

✓ ✓✓✓

✓X ✓✓✓X

Study S2b: Effect of Broken Streaks

**Figure S4.** Participants’ sense of accomplishment (an indicator of goal pursuit) as a function of their recent pattern of logged behavior. Error bars are 1 standard error.

Study S2a: Effect of Intact Streaks

*Mediation analysis*. Sense of accomplishment mediated the negative effect of a broken streak (versus no streak) on participants’ engagement in the target behavior (Indirect effect = -.21, SE = .11, 95% CI = [-.49, -.05]).

**Discussion**

These studies showed that participants were more likely to engage in a behavior when it contributed to a logged streak, relative to when it did not (S2a), and that participants were less likely to engage in a behavior when it followed a broken logged streak, relative to when it did not (S2b). Notably, we demonstrated these effects when all participants engaged in the same logged behavior before our primary dependent variable. Thus, these results suggest the presence of both a positive effect of an intact logged streak and a negative effect of a broken logged streak above and beyond any recency effects. Moreover, by controlling for participants’ most recent behavior, Study S2b’s findings indicate that broken logged streaks decrease subsequent behavior above and beyond any possible effects of our “quota” manipulation.

Additionally, these studies provided support for our proposed mechanism. Relative to not having a streak, having an intact streak within a behavioral log led to a greater sense of accomplishment, and this drove the positive effect of intact streaks on subsequent behavior. Furthermore, relative to having a recent miss, having a broken logged streak led to a decreased sense of accomplishment, which drove the negative effect of a broken streak.

Supplemental Studies 3a and 3b

In the studies reported in the manuscript, we 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). However, these findings cannot address how the length of a logged streak might affect consumers’ motivation to maintain it and shape their subsequent behaviors. Thus, Studies S3a and S3b both test the effects of intact versus broken streaks of varying lengths. Specifically, Study S3a examined the effects of short (three in a row) and long (six in a row) streaks of logging the completion of real language-learning questions in an app, while Study S3b examined the effects of short (four in a row) and long (twenty in a row) streaks of logging beers on a beer-tasting app via a scenario.

We anticipated one of two possible effects of longer streaks that are in line with our theory that consumers regard the maintenance of logged streaks as a goal. One possibility is that the effects of logged streaks will be similar regardless of length, which would suggest that the activation of a streak goal (from an intact logged streak) is categorical and insensitive to streak length (cf. Sheeran, Webb and Gollwitzer 2005). Alternatively, it is possible that the effects of logged streaks will be stronger for longer streaks compared to shorter streaks, since longer streaks require more effort to achieve, and may thus seem like more of an accomplishment worth maintaining (or more aversive to break: cf. Locke and Latham 1990).

As in Study 4, both Studies S3a and S3b also measured the extent to which participants adopted a goal of maintaining their logged streaks. Study S3a also measured participants’ behavioral intentions to continue using and recommend the app to a friend (as in Studies 4 and 6). In addition, building on the finding from Study 4 that participants were willing to sacrifice the opportunity to engage in a more enjoyable behavior to maintain their streaks, Study S3a investigated participants’ willingness to watch an advertisement to repair or maintain their streaks.

## Study S3a: The Effect of Streak Length on Real Language Learning

**Methods**

We recruited 596 MTurk participants (*M* age = 36.03, 51.16% female). As in Study 4, all participants were told that they would be testing out Portuguese questions for a language-learning app. All participants were informed that, just like on many other apps, they would see a behavioral log that would track all of the language-learning questions they attempted by adding a checkmark.

Participants were randomly assigned to condition in a 2(*logged streak:* *intact* or *broken*) by 2(*streak* *length: short* or *long*) between-subjects design. In the *short streak length* condition, participants completed four Portuguese questions in a row. In the *long streak length* condition, participants completed seven questions in a row. Similar to Study 4, we also manipulated whether the last question (i.e., the fourth question in the *short* condition and the seventh question in the *long* condition) was successfully recorded in the behavioral log. Specifically, in the *intact logged streak* condition, this last question was logged successfully, leaving participants’ logged streaks intact. However, in the *broken logged streak* condition, participants saw an error message after completing the last question that informed them that there was an issue with the log for this question due to the app still being in the testing phase. Again, participants were informed that this error would not occur again. In this study, participants in the *broken logged streak* condition were simply told that their streak on their behavioral log was broken, rather than seeing an “X” on the log itself.

Participants then chose what type of activity they would like to do next: either continue with the target behavior (complete another language-learning question on the app) or switch to watching a fun video.[[1]](#footnote-1)

Before engaging in their chosen activity, participants responded to the same *streak maintenance goal* measure and behavioral intention items from Study 4 (their likelihood to continue using the app and recommend it to a friend). Additionally, participants reported whether they would be willing to watch an advertisement to maintain their logged streak (*intact logged streak* condition) or to repair their broken logged streak (*broken logged streak* condition). Finally, participants did their chosen activity and answered demographics.

**Results**

*Target behavior.* A binary logit with *logged streak* condition, *streak* *length* condition, and their interaction as factors revealed a main effect of *logged streak* condition (*F*(1, 592) = 28.82, *p* < .001); participants were more likely to engage in the target behavior when they had an intact streak (65.76%) than when they had a broken streak (43.85%; *OR* = 2.46). There was a marginal effect of *streak length* condition (*F*(1, 592) = 2.98, *p* = .085); participants in the *short streak* condition (57.91%) were more likely to continue the target behavior than participants in the *long streak* condition (51.51%; *OR* = 1.30). However, the interaction between the *logged streak* and *streak length* conditions was not significant (*F*(1, 592) = 0.03, *p* = .87; see Figure S5). Participants with an intact logged streak were more likely to engage in the target behavior than participants with a broken logged streak, both following a shorter streak (*intact:* 69.44% vs. *broken:* 47.06%; *Χ*2 (1) = 15.25, *p* < .001; *OR* = 2.56) and a longer streak (*intact:* 62.25% vs. *broken:* 40.54%; *Χ*2 (1) = 14.11, *p* < .001; *OR* = 2.42).

**Figure S5.** Percent of participants who chose to engage in the target behavior as a function of their logged streak being intact or broken, and its length (short or long). Error bars are 1 SE.

*Streak maintenance goal.* A two-way ANOVA with *logged streak* condition, *streak length* condition, and their interaction as factors revealed a main effect of the *logged streak* condition (*F*(1, 592) = 48.70, *p* < .001), such that participants with an intact logged streak reported viewing streak maintenance as an explicit goal (*M* = 6.94, SD = 3.35) to a greater extent than participants with a broken streak (*M* = 4.99, SD = 3.45, *d* = 0.66). There was no effect of *streak length* condition (*F*(1, 592) = 0.30, *p* = .59) nor a significant interaction (*F*(1, 592) = 0.07, *p* = .79). The effects held when we considered each *streak length* condition separately (*short: Mintact* = 6.98, SD = 3.38 vs. *Mbroken* = 5.10, SD = 3.57; *t*(295) = 4.64, *p* < .001, *d* = 0.54; *long*: *Mintact* = 6.90, SD = 3.34 vs. *Mbroken* = 4.88, SD = 3.34; *t*(297) = 5.24, *p* < .001, *d* = 0.61).

*Mediation analysis*. We conducted a mediation analysis that included *logged streak* condition (*intact* = 1; *broken* = 0) as the independent variable, the *streak maintenance goal* item as the mediator, and choice as the dependent variable. The extent to which participants reported adopting a streak maintenance goal mediated the effect of an intact versus broken streak on engaging in the target behavior (Indirect effect = .62, SE = .11, 95% CI = [.42, .85]). This mediation held when we considered each *streak length* condition separately (*short*: Indirect effect = .49, SD = .13, 95% CI = [.28, .80]; *long*: Indirect effect = .77, SE = .18, 95% CI = [.46, 1.18]).

*Behavioral intentions.* As in Study 4, participants in the *intact logged streak* condition were more likely to continue using the app and recommend it to a friend compared to participants in the *broken logged streak* condition (*F*s > 30, *p*s < .001, *d*s > 0.45). There was no effect of *streak length* condition nor an interaction for these measures (*F*s < 3.60, *p*s > .05). Furthermore, 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 streaks.

## Study S3b: The Effect of Streak Length on Beer-Logging

**Methods**

We recruited 507 participants via MTurk (*M* age = 35.79, 47.14% female). Recruitment was targeted towards individuals who were interested in craft beer and had a smartphone, so that the scenario would be relevant and familiar to participants.

Participants first read about a beer-logging app that allows users to track and rate beers that they try (similar to the Untappd app). All participants were told to imagine that they had become regular users of this app, and that they had logged at least 50 beers recently on the app. Next, they were told to imagine that they had been to a bar the previous weekend, where they had tried a new beer.

Participants were randomized to condition in a 2(*logged streak: intact or* *broken*) by 2(*streak length: short* or *long*) between-subjects design. In the *intact* *logged* *streak* condition, participants were told that they had successfully logged the beer they tried last weekend. In the *broken logged streak* condition, participants were told that they wanted to log this beer, but they were unable to do so because of poor cell service, a factor outside of their control. Participants in the *short streak length* condition were told that they had logged four beers in a row before last weekend, while participants in the *long streak length* condition were told that they had logged twenty beers in a row before last weekend.

Then, all participants were then told to imagine they were enjoying a beer today, and that this was the first beer they had drunk since the one they had at the bar last weekend. They then rated their likelihood to log this beer on their beer app on a 1 “Extremely unlikely” to 11 “Extremely likely” scale. This served as our primary dependent variable.

Next, participants answered several questions about their thought processes and feelings in the scenario, all on 1 “Not at all” to 11 “Extremely”/“A great deal” scales. Participants answered two items similar to the direct process measure used in Studies 4 and S3a (“How much did you think about streaksyou may have in your log?” and “How much did you think about consistency in your logging behavior?”; *r* = .59), which we combined to create one *streak maintenance goal* measure. Participants also answered several questions meant to capture other indicators of goal adoption and progress. Specifically, like in Studies 6 and 7, participants also answered four *negative emotion* items (how annoyed, upset, angry, and disappointed they felt; α = .94). In addition, participants answered two *positive emotion* items (how happy and excited they felt; *r* = .47), as well as one item similar to the *motivation* subscale items in Studies 4, 5, and 7 (“How much did you feel motivated to continue logging?”).

Additionally, participants answered several exploratory items. Two items pertained to feelings of automaticity (“How automatic was your decision about whether to log your beer today?” and “How much did you carefully consider your decision about whether to log your beer today?”, reverse coded; *r* = .09), and were averaged together to create one *automaticity* measure. Five items concerned participants’ feelings of momentum (“How much did you think about the "momentum"you had in logging?”; “How much did your decision to log feel effortless?”, “How much did you feel that your logging behavior was self-sustaining?”, “How much did your past behavior push you to log?”, and “How much did you feel like you were ‘on a roll’?”; α = .76), and were averaged together to create one *perceived momentum* measure.

Participants then answered two manipulation check questions about whether they had logged their beer from last weekend and how long their streak was before last weekend (90.14% and 95.46% participants, respectively, answered these questions correctly). Lastly, participants answered basic demographics.

**Results**

*Likelihood of logging*. A two-way ANOVA with *logged streak* condition, *streak length* condition, and their interaction as factors revealed a main effect of *logged streak* condition on likelihood of logging today’s beer (*F*(1, 503) = 36.65, *p* < .001). Participants with an intact streak in their behavioral log were more likely to log their next beer (*M* = 9.97, SD= 1.76) than participants who had a broken streak (*M* = 8.92, SD= 2.10, *d* = 0.54). There was no significant effect of streak length on likelihood of logging today’s beer (*F*(1, 503) = 0.04, *p* = .85). There was also no significant interaction between streak and length (*F*(1, 503) = 0.57, *p* = .45; see Figure S6). Furthermore, the effect of an intact versus broken streak held in both *streak length* conditions (*short:* *Mintact* = 10.03, SD = 1.55 vs. *Mbroken* = 8.84, SD = 2.22; *t*(249) = 4.78, *p* < .001, *d* = 0.62; *long*: *Mintact* = 9.93, SD = 1.94 vs. *Mbroken* = 9.01, SD = 1.97; *t*(254) = 3.78, *p* < .001, *d* = 0.47).

**Figure S6.** Participants’ likelihood of continuing to use a beer-logging app, as a function of their logged streak being intact or broken, and its length (short or long). Error bars are 1 SE.

*Streak maintenance goal.* A two-way ANOVA with *logged streak* condition, *streak length* condition, and their interaction as factors revealed a main effect of *logged streak* condition (*F*(1, 503) = 43.47, *p* < .001); participants with an intact logged streak reported considering their streak and consistency in logging (*M* = 8.92, SD = 2.09) to a greater extent than participants who had a broken logged streak (*M* = 7.56, SD= 2.47, *d* = 0.59). There was no significant effect of streak length (*F*(1, 503) = 0.41, *p* = .52), nor a significant interaction (*F*(1, 503) < 0.01, *p* = .98). The effect of an intact versus broken streak held in both *length* conditions (*short*: *Mintact* = 8.85, SD = 2.07 vs. *Mbroken* = 7.50, SD = 2.39; *t*(249) = 4.72, *p* < .001, *d* = 0.60; *long*: *Mintact* = 8.99, SD = 2.12 vs. *Mbroken* = 7.63, SD = 2.57; *t*(254) = 4.62, *p* < .001, *d* = 0.58).

*Mediation analysis*. We conducted a mediation analysis that included *logged streak* condition (*intact* = 1; *broken* = 0) as the independent variable, the *streak maintenance goal* measure as the mediator, and likelihood of continuing to log as the dependent variable. The confidence interval for the indirect effect excluded zero (Indirect effect = .33, SE = .08, 95% CI = [.19, .53]). Results were similar when we considered each *streak length* condition separately (*short*: Indirect effect = .49, SD = .13, 95% CI = [.28, .80]; *long*: Indirect effect = .77, SE = .18, 95% CI = [.46, 1.18]).

*Indicators of goal pursuit.* Two-way ANOVAs revealed that participants in the *intact logged streak* condition felt less negative emotion (*M* = 1.83, SD = 1.89), more positive emotion (*M* = 6.58, SD = 2.54), and more motivation (*M* = 9.49, SD = 1.75) than participants in the *broken logged streak* condition (*Mnegative emotion* = 8.18, SD = 2.10; *Mpositive emotion* = 5.19, SD= 2.35; *Mmotivation* = 8.52, SD= 2.27; *F*s > 28.00, *p*s < .001, *d*s > 0.47). There were no significant effects of streak length (*F*s < 0.50, *p*s > .60), nor any significant interactions (*F*s < 3.90, *p*s > .05) for these measures. These effects persisted in both *streak length* conditions (*t*s > 3.10, *p*s < .002, *d*s > 0.39).

*Exploratory measures.* Participants in the *intact logged streak* condition felt their decision to continue logging was more automatic (*M* = 7.44, SD = 2.07) and that they had more momentum in logging (*M* = 8.74, SD = 1.66) than participants who had a broken streak (*Mautomaticity* = 6.82, SD= 1.95; *Mmomentum* = 7.57, SD= 1.81; *F*s > 12.00, *p*s < .001, *d*s> 0.30). There were no significant effects of streak length (*F*s < 0.50, *p*s > .60), nor any significant interactions (*F*s < 0.50, *p*s > .75). These effects persisted in both *length* conditions (*t*s > 2.40, *p*s < .020, *d*s > 0.39).

**Discussion**

Replicating previous studies, participants were more likely to engage in a target behavior (Study S3a) and were more likely to continue logging (Study S3b) when they had an intact (versus broken) streak in their behavioral log. Notably, this effect was similar for both shorter and longer streaks. This finding suggests that whether a streak goal is activated plays a more important role in subsequent behavior than the length of the streak at that moment, which is in line with our theory.

These studies also replicated multiple key results that provide further evidence of our theory. First, both studies again found evidence that the effect of logged streaks on behavior is driven by the extent to which consumers have a goal of maintaining their logged streaks. Second, Study S3b provided additional process evidence via additional indicators of goal pursuit; participants with an intact logged streak felt more positive emotion, less negative emotion, and more motivation than participants with a broken logged streak. Third, Study S3a again found that this effect occurred even when the outside option was more enjoyable than the target behavior, indicating that consumers are willing to make a sacrifice to maintain their logged streaks. Moreover, half of participants in Study S3a were willing to engage in a completely separate undesirable behavior – watching an advertisement – to maintain or repair their logged streaks.Finally, Study S3a again demonstrated that logged streaks influence other behavioral intentions: participants in the *intact logged streak* condition were more likely to recommend the app to a friend and were themselves more likely to continue using the app than participants in the *broken logged streak* condition.

Supplemental Study 4

This study investigates a potential moderator of the effect of intact versus broken logged streaks on subsequent behavior: whether behavioral logs are made public or kept private. Our theory could make two potential predictions about the effects of this factor. On the one hand, making streaks more public could amplify the effects of logged streaks, as doing so may amplify consumers’ sense of accountability (e.g., Rogers et al. 2015) or their perceived status (Moldovanu, Sela and Shi 2007), which would increase their motivation to maintain their streaks. On the other hand, to the extent that consumers’ motivation for maintaining logged streaks primarily stems from *personal* feelings of accomplishment (e.g., Bandura 1977), making behavioral logs more public might not have any effect on how logged streaks impact subsequent behavior. In this study, we tested these possibilities by manipulating whether participants were informed that their behavioral logs – and the intact or broken streaks therein – would be made public to other users of the app, or kept private.

This study also builds on our studies reported in the manuscript in several ways, with both theoretical and practical implications. First, to manipulate participants’ logged streaks, this study leverages a similar categorization operationalization as Study 5. That is, we hold participants’ actual behavior constant, and simply manipulate whether the last behavior contributes to an intact streak (when the category of logged behaviors is broader) or breaks a streak (when the category of logged behaviors is narrower). Second, we again examine participants’ willingness to watch an advertisement to repair or maintain their streaks, as well as other behavioral intentions (i.e., likelihood to continue using the app or recommend it to a friend). Moreover, we test whether participants are willing to sacrifice the opportunity to engage in a more enjoyable behavior (here, viewing funny online content) to maintain their logged streaks. Third, we further examine the proposed mechanism by measuring indicators of goal pursuit (negative and positive emotion).

**Methods**

We recruited 604 MTurk participants (*M* age = 40.84, 52.81% female, 0.99% other/prefer not to say). As in Studies 5, 6, and 7, all participants were told that they would be testing out a gaming app that had both Word Jumbles and Number Sums available. Here, they started with the Word Jumble games.

Participants were randomly assigned to condition in a 2(*logged streak:* *intact* or *broken*) by 2(*public* or *private*) between-subjects design. All participants were informed that, just like on many other apps, they would see a behavioral log that would track all of the games they attempted. We manipulated whether participants’ streaks displayed in this behavioral log would remain private or be made public. Specifically, at the start of the study, participants in the *private* condition read, “your activity on the app – including your progress on the game tracker – will be private. That is, how you progress through the games will NOT be saved or displayed anywhere.” Participants in the *public* condition instead read, “your activity on the app – including your progress on the game tracker – will be public. That is, how you progress through the games will be saved and displayed to people who use the app in the future.” The majority (93.87%) of participants correctly answered a manipulation check at the end of the study regarding whether their behavior on the app would be made public or kept private. Moreover, an additional question at the end of the study confirmed that participants in the *public* condition thought more about what other users would think about their activities on the app (*M* = 2.84, SD = 1.88) than participants in the *private* condition (*M* = 2.24, SD = 1.71; *t*(602) = 4.08, *p* < .001, *d* = 0.33).

Like in Study 5, we also manipulated participants’ logged streak by varying the types of games that counted on the behavioral log. Participants 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 participants in the *broken logged streak* condition read that only completing Word Jumbles would give them a checkmark on the behavioral log. All participants then completed three Word Jumbles followed by one Number Sum. Thus, all participants completed the exact same series of games, but their fourth game either maintained or broke their logged streak, depending on condition.

Participants then chose what type of activity they would like to do next: either continue with another Word Jumble on the app (the target behavior) or switch to viewing funny online content (Buzzfeed’s Comics Pinterest board). This served as our primary dependent variable.

Before engaging in their chosen activity, participants answered an item intended to measure their adoption of a streak maintenance goal*:* “To what extent did you have a goal of maintaining a streak of playing games?” (from 1 “Not at all” to 11 “Very much”). However, we realized post hoc that there was a typo in our question wording, as *all* participants had a streak of playing games (but not *logging* those games).[[2]](#footnote-2) Participants also answered two *negative emotion* items (how sad and upset they felt; *r* = .80) and two *positive emotion* items (how happy and excited they felt; *r* = .84). Additionally, participants responded to the same two behavioral intention items from Studies 4 and 6 (their likelihood to continue using the app and recommend it to a friend). Also, as in Study 6, participants reported whether they would be willing to watch an advertisement to maintain their logged streak (*intact logged streak* condition) or to repair their broken logged streak (*broken logged streak* condition). Finally, participants did their chosen activity and answered demographics.

**Results**

*Target behavior.* A binary logit with *logged streak* condition, *public* condition, and their interaction as factors revealed a main effect of *logged streak* condition (*F*(1, 600) = 8.54, *p* = .004); participants were more likely to engage in the target behavior when they had an intact streak (55.78%) than when they had a broken streak (43.85%; *OR* = 1.61). There was no significant effect of *public* condition (*F*(1, 600) = 0.73, *p* = .39), nor was there a significant interaction (*F*(1, 600) = 0.14, *p* = .71; see Figure S7). Moreover, the effect of an intact versus broken logged streak was significant when participants’ behavioral logs would be made public (*intact:* 58.28% vs. *broken:* 44.81%; *Χ*2 (1) = 5.54, *p* = .019; *OR* = 1.72) and marginally significant when their behavioral logs would be kept private (*intact:* 53.29% vs. *broken:* 42.86%; *Χ*2 (1) = 3.26, *p* = .071; *OR* = 1.52).

**Figure S7.** Percent of participants who engaged in the target behavior as a function of their logged streak being intact or broken, and whether it would be kept private or made public. Error bars are 1 SE.

*Indicators of goal pursuit. P*articipants in the *intact logged streak* condition felt more positive emotion (*M* = 7.17, SD = 2.58) than participants in the *broken logged streak* condition *(M* = 5.05, SD = 2.71; *F*(1, 600) = 95.99, *p* < .001, *d* = 0.80)*.* There was neither an effect of the behavioral log being public or private, nor a significant interaction (*F*s < 1.50, *p*s > .25).

Additionally, participants in the *intact logged streak* condition felt less negative emotion (*M* = 2.03, SD = 2.11) than participants in the *broken logged streak* condition *M* = 4.47, SD = 32.901; *F*s > 90.00, *p*s < .001, *d*s > 0.80)*.* While there was no effect of the behavioral log being public versus private (*F*(1, 600) = 0.22, *p* = .64), there was a significant interaction (F(1, 600) = 6.67, *p* = .010). The effect of an intact versus broken logged streak on negative emotion was larger in the *public* condition (*Mintact* = 1.81, SD = 1.66 vs. *Mbroken* = 4.78, SD = 3.02; *t*(303) = 10.58, *p* < .001, *d* = 1.22) than in the *private* condition (*Mintact* = 2.25, SD = 2.46 vs. *Mbroken* = 4.15, SD = 2.74; *t*(297) = 6.32, *p* < .001, *d* = 0.73).

*Behavioral intentions.* Participants in the *intact logged streak* condition were more likely to continue using the app and to recommend it to a friend compared to participants in the *broken logged streak* condition (*F*s > 15.00, *p*s < .001, *d*s > 0.30). There was no effect of the behavioral log being public versus private, nor was there an interaction for these measures (*F*s < 1.50, *p*s > .20). Furthermore, 38% of participants were willing to watch an advertisement to maintain their logged streaks, and 48% of participants were willing to watch an advertisement to repair their broken logged streaks.

**Discussion**

This study replicated the effect of intact versus broken logged streaks on subsequent behavior. This effect held both when participants’ behavioral logs would be made public and when they would be kept private. Interestingly, the effect was directionally larger for public behavioral logs, suggesting that public recognition could potentially amplify the effects of logged streaks, perhaps even more so in settings where consumers are naturally quite invested in feedback from others (e.g., on social media).

Moreover, consistent with our theory, participants with an intact streak in their behavioral log felt more positive emotion and less negative emotion than participants with a broken streak. As further evidence that consumers value their logged streaks, participants were also willing to forgo a more enjoyable outside option (here, viewing funny online content) to keep their logged streaks intact, and approximately 40% of participants were even willing to watch an advertisement to maintain (or avoid breaking) their logged streaks. Finally, as in Study 4, we found that participants with an intact streak in their behavioral log were more likely to continue using the app and to recommend it to a friend than participants with a broken streak.

Post-Test

We conducted a post-test to determine if participants made any unintended inferences from seeing the quota message. We recruited 374 participants from a (*M* age = 20.09, 56.64% female) from a behavioral lab at a private northeastern university. Sample size was determined by how many participants were recruited for the lab session.

Earlier in the lab session, participants completed a pilot similar to studies featured in our manuscript, where they completed Portuguese questions on a language learning app. Importantly, this pilot had no mention of a quota of any kind. After several filler tasks, participants were asked to imagine that they were testing out a language learning app like the one they had used earlier in the session, and were told that this scenario was unrelated to any choices or responses they had made in the previous study. They then saw the same instructions about the quota message that were used in Studies 3, 6, 7, and S2b (see OSF for exact instructions).

Participants were randomly assigned to one of two between-subjects conditions. In the *quota* condition, participants were told that, “After completing a few questions, you see this quota message in place of a vocabulary question.” In the *no quota* condition, participants were told that, “After completing a few questions, you do NOTsee this quota message.”

Then, participants answered four questions intended to measure any inferences participants may have made from seeing (or not seeing) the quota message (all on 1 to 7 scales; see Table S4). There were no significant effects of the *quota* condition on participants’ responses to these measures (*t*s < 1.60, *p*s > .10, *d*s < .20). More specifically, in this post-test, the quota message manipulation did not significantly affect participants’ perceived likelihood of seeing another quota message, their perceptions that it would help the researchers to respond in a certain way, their perceptions of app functionality, or their liking of the app. Hence, we find no evidence that our quota message manipulation triggers systematic inferences, at least regarding these dimensions.

|  |  |  |  |
| --- | --- | --- | --- |
| Question | No Quota condition M (SD) | Quota condition  M (SD) | Statistic |
| How likely is it that you would see the quota message in place of a future question?  (1 “Not at all”, 7 “Extremely”) | 3.99 (1.35) | 3.81 (1.29) | *t*(372) = 1.34  *p* = .18  *d* = .14 |
| How much would you feel like it would be more helpful to the app developers if you stopped learning Portuguese and switched to a different activity (e.g., learning a different language or watching a video)?  (1 “Not at all”, 7 “Extremely”) | 3.63 (1.47) | 3.87 (1.49) | *t*(372) = 1.55  *p* = .121  *d* = .16 |
| How well do you feel the app is functioning?  (1 “Very poorly”, 7 “Very well”) | 4.32 (1.29) | 4.18 (1.31) | *t*(372) = .41  *p* = .69  *d* = .11 |
| How much do you like the app?  (1 “Not at all”, 7 “Extremely”) | 4.15 (1.38) | 4.09 (1.32) | *t*(372) = .41  *p* = .68  *d* = .04 |

**Table S4.** Measures of participants’ inferences about the quota message manipulation in the post-test.

1. Similar to Study 4, a question at the end of the survey confirmed that participants preferred watching fun videos over completing language-learning questions (*t*(595) = 2.53, *p* = .012, *drm* = 0.13). [↑](#footnote-ref-1)
2. Because of this mistake, it is not surprising that there was no effect of *logged streak* condition on this measure, nor was there an effect of *public* condition or a significant interaction (*F*s < 1.00, *p*s > .75). [↑](#footnote-ref-2)