The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention

Ran Kivetz, Oleg Urminsky, and Yuhuang Zheng
Columbia University, Graduate School of Business

Forthcoming in the *Journal of Marketing Research*, February 2006

Ran Kivetz is the Sidney Taurel Associate Professor of Business, Graduate School of Business, Columbia University (rk566@columbia.edu). Oleg Urminsky (opu1@columbia.edu) and Yuhuang Zheng (yz143@columbia.edu) are both doctoral candidates at Columbia University. The authors are indebted to David Katz, the manager of Columbia University Café Cappuccino, and to MoodLogic, Inc. for their cooperation. The authors are also grateful for helpful comments and suggestions received from Pradeep Chintagunta, Sunil Gupta, Raghu Iyengar, Gita Johar, Yifat Kivetz, Rajeev Kohli, Donald Lehmann, Oded Netzer, P. B. Seetharaman, Itamar Simonson, Andrea Vag, participants in seminars at Columbia University, Stanford University, the University of Florida, the University of Pennsylvania, and the Society for Judgment and Decision-Making, and the JMR Editor and reviewers. The research reported in this article was supported by the Lang Faculty Research Fellowship in Entrepreneurship.

The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention

ABSTRACT

The goal-gradient hypothesis (Hull 1934) denotes the classic finding from behaviorism that animals expend more effort as they approach a reward (e.g., hungry rats run faster as they near cheese). Building on this hypothesis, we generate new propositions for the human psychology of rewards. We test these propositions using a variety of methods, data, and modeling approaches, including field experiments, paper-and-pencil problems, and secondary customer data as well as hazard rate, Tobit, and logit models. Some of the key findings indicate that: (a) participants in a real café reward program ("buy ten coffees, get one free") purchase coffees more frequently the closer they are to earning a free coffee; (b) Internet users who rate songs in return for reward certificates visit the rating website more often, rate more songs per visit, and persist longer in the rating effort as they approach the reward goal; (c) the *illusion of progress toward the goal* likewise induces purchase acceleration; for example, customers who receive a "12-stamp" coffee card with two pre-existing "bonus" stamps complete the 10 required purchases faster than customers who receive a "regular" 10-stamp card; and (d) a stronger tendency to accelerate toward the reward predicts greater retention and faster reengagement in the program. Our conceptualization and empirical findings are captured by a parsimonious *goal-distance model*, in which effort investment is a function of the proportion of original distance remaining to the goal (i.e., psychological distance). In addition, using statistical and experimental controls, we show that the observed goal-gradients cannot be explained by habituation, expiration concerns, other time-trend effects, or heterogeneity bias. We discuss the theoretical implications of the findings and their managerial implications for incentive systems, promotions, and customer retention.

"… [R]ats in a maze … run faster as they near the food box than at the beginning of the path."

- Hull 1934

The goal-gradient hypothesis, originally proposed by the behaviorist Clark Hull in 1932, states that the tendency to approach a goal increases with proximity to the goal. In a classic experiment testing this hypothesis, Hull (1934) found that rats in a straight alley ran progressively faster as they proceeded from the starting box to the food. While the goal-gradient hypothesis has been investigated extensively with animals (e.g., Anderson 1933; Brown 1948; for a review see Heilizer 1977), its implications for human behavior and decision-making are understudied. Further, this issue has important theoretical and practical implications for intertemporal consumer behavior in reward programs and other types of motivational systems (e.g., Deighton 2000; Hsee, Yu, and Zhang 2003; Kivetz 2003; Lal and Bell 2003; van Osselaer, Alba, and Manchanda 2003).

In this article, we build on the behaviorist goal-gradient hypothesis and generate new propositions in the context of two real reward programs (hereafter "RPs"). In the interdisciplinary spirit of bridging the consumer behavior and marketing science fields (Winer 1999; Wittink 2004), we investigate these propositions using a variety of methods, data, and modeling approaches (e.g., field experiments, paper-and-pencil problems, and secondary customer data; hazard rate, Tobit, and logit models). Consistent with the goal-gradient hypothesis, its corollaries, and their adaptation to the human psychology of rewards, some key findings indicate that:

1. Members of a café RP ("buy ten coffees, get one free") purchase coffee more frequently the closer they are to earning a free coffee (on average, inter-purchase times decrease by 20% or 0.7 days throughout the program).

2. The findings generalize beyond coffee purchases to effort involving repeatedly rating music over the Internet and goal-gradients operationalized via acceleration in inter-visit times, lift in rating quantities, and enhanced persistence closer to the reward threshold.

3. The illusion of progress toward the goal likewise induces purchase acceleration. For example, customers who receive a "12-stamp" coffee card with two pre-existing "bonus" stamps complete the 10 required purchases faster than customers who receive a "regular" 10-stamp card. Process experiments show that the illusionary goal progress effect cannot be explained by rival accounts such as sunk cost.

4. Consistent with the notion that a steeper goal-gradient indicates a stronger motivation to earn rewards, individuals' tendency to accelerate toward their first reward predicts a higher probability of retention and faster reengagement in the program.

5. All of the findings are captured by a parsimonious goal-distance model, in which effort investment is a function of the proportion of original distance remaining to the goal (i.e., psychological distance).

6. The observed purchase and effort acceleration cannot be explained by habituation, expiration concerns, other time-trend effects, or heterogeneity bias. For example, goal-motivated acceleration is observed after accounting for weekly sales and other time-varying covariates, and a majority of accelerators are found after accounting for unobserved heterogeneity in both the hazard rate and the tendency to accelerate. Importantly, purchase and effort rates reset (to a lower level) after the first reward is earned and then re-accelerate toward the second reward goal.

The paper is organized as follows: Section 1 briefly reviews the behaviorist goal-gradient hypothesis and considers its relevance for the context of consumer RPs. Section 2 proposes a theoretical goal-distance model that incorporates the goal-gradient hypothesis; in subsequent sections, we employ this model to generate and test new propositions that highlight the intriguing consequences of the goal-gradient hypothesis for the human psychology of rewards. Section 3 discusses a real café RP and a discrete-time proportional hazard rate model used to test for purchase acceleration. Section 4 reports field and questionnaire experiments that test the effect of illusionary goal progress. Section 5 reports data from a second real incentive system, which generalizes the findings to acceleration and persistence in effort involving repeatedly rating music. Section 6 explores the implications of the goal-gradient for customer retention. Finally, Section 7 discusses the theoretical and managerial implications of this research.

1. THE GOAL-GRADIENT HYPOTHESIS IN BEHAVIORISM

Originally formulated by Hull (1932) and refined by Miller (1944), the goal-gradient hypothesis states that the tendency to approach a goal increases with proximity to the goal. The strongest evidence for this hypothesis has been obtained in the context of animal learning, consistent with Hull's (1932) prediction “[t]hat animals in traversing a maze will move at a progressively more rapid pace as the goal is approached” (p. 42). Hull (1934) constructed a straight runway with electrical contacts placed so that the time consumed by rats to cross each of several six-foot sections could be precisely measured. The key finding - replicated in several variations of the procedures and apparatus - indicated that the animals ran faster the closer they were to the food reward. Figure 1 displays a typical set of results from Hull (1934) that reveal this pattern. In another widely quoted study, Brown (1948) attached rats to an apparatus that recorded the force in grams with which the rats pulled (toward the region of reinforcement) when stopped either close to or far from the food. Consistent with the goal-gradient hypothesis, rats that were stopped closer to the food pulled more strongly than those stopped farther away (see also Anderson 1933).

A review of the literature (Heilizer 1977) reveals that most goal-gradients were obtained with rats and physical responses (e.g., speed-of-locomotion). A considerable portion of goal-gradient studies examined issues that are not directly related to the current investigation, such as contrasting approach versus avoidance gradients (e.g., Förster, Higgins, and Idson 1998). The limited research conducted with humans employed mostly physiological measurements (e.g., galvanic skin response, heart rate, arm pressure) that were non-instrumental for goal achievement. As detailed in Heilizer (1977), it is difficult to interpret non-instrumental behaviors and physiological measurements as evidence supporting (or refuting) the goal-gradient hypothesis. To the best of our knowledge, there is no published study that provides unequivocal evidence of a systematic, behavioral goal-gradient in humans. Thus, one of the primary goals of this article is to test for behavioral (approach) goal-gradients and their various operationalizations (timing, quantity, and persistence of effort).

In this research, we employ real RPs as an empirical context for investigating the goal-gradient hypothesis. Such programs share a common underlying structure, whereby people need to invest a stream of efforts to earn future rewards. This general effort-reward structure applies to many decision contexts and life domains, including consumer loyalty programs, employee incentive systems, sales-force compensation plans, patient compliance programs, and even academic tenure tracks. Nevertheless, it is noteworthy that the empirical impact of RPs on actual customer behavior is still largely undetermined (cf. Dowling and Uncles 1997; Lewis 2004; Sharp and Sharp 1997).

2. THEORY AND MODEL

The notion that progress and distance to the goal affect consumer motivation is supported by theories of social-cognition and human decision-making. Dynamic models of motivation (e.g., Atkinson 1957; Lewin 1951; Miller 1944) propose that people possess a strong achievement drive, which is heavily influenced by goals. Carver and Scheier's (1990) cybernetic control model suggests that comparisons of the rate of progress toward the goal with a relevant criterion generate affect; when progress exceeds [falls short of] the criterion, positive [negative] affect arises (see also Soman and Shi 2003). Researchers have also highlighted the impact on decision-making and behavior of individuals' *psychological distance* from their outcomes and goals (Lewin 1951; Trope and Liberman 2003). Additionally, Heath, Larrick, and Wu (1999) propose that, due to the diminishing sensitivity of prospect theory's value function, people should exert more effort as they near their (self-imposed) goals. In summary, prior theorizing regarding human motivation, affect, and cognition supports the relevance of the goal-gradient hypothesis for the human psychology of rewards.

What are the implications, then, of the goal-gradient hypothesis for RPs? As originally proposed by Kivetz (2000), the notion that achievement motivation increases with smaller goal-distance suggests that customers will accelerate and persist in their efforts as they near the program's incentive threshold (i.e., the reward requirement or goal). The operationalization of "effort acceleration" depends on the specifics of the particular RP. When the program requirements involve discrete purchases or incidents (e.g., "stay ten nights and earn a reward"), the acceleration will be manifested by more frequent activity (shorter inter-purchase or inter-visit times). When the RP is structured so that more intense activity (e.g., a larger purchase or more units of effort) in any single visit earns more credits toward the reward (e.g., "earn one point for each dollar spent"), acceleration may be detected through both shorter inter-purchase times and increased purchase (or effort) quantities. In the present research, we investigate both temporal and quantity operationalizations of goal-motivated acceleration. We also examine various forms of RP effort, including real purchases (of coffee) and actual work (rating music online). Finally, we generalize the goal-gradient hypothesis by examining whether consumers persist longer in their effort as a function of smaller goal-distance. Next, building on prior research, we develop a parsimonious *goal-distance model* that incorporates the aforementioned and other predictions.

The Goal-Distance Model

A great deal of research in psychophysics and judgment and decision-making has shown that perception and preference are sensitive to relative rather than absolute dimensions. For example, Stevens (1957) and his predecessors demonstrated that sensory experiences reflect ratio (proportionality) judgments rather than absolute magnitude differences. Preference and choice have also been shown to be relative rather than absolute, depending on such factors as the salient reference point (Kahneman and Tversky 1979; Prelec and Loewenstein 1998); the relative positions of other alternatives in the choice set (Huber, Payne, and Puto 1982; Simonson and Tversky 1992); the relative accuracy and effort of decision strategies (Payne, Bettman, and Johnson 1992); the preferences of other people (Kivetz and Simonson 2003); and the relative (proportional) value of the choice options (Herrnstein and Prelec 1991).

The sensitivity to relative and reference values suggests that consumers will spontaneously consider their distance to a goal incorporating the total distance as a reference point, leading to an evaluation of relative goal-distance. Accordingly, we conceptualize and model the psychological (or perceived) goal-distance as the proportion of the total (original) distance remaining to the goal. We define this distance as $d_t = (r-n_t)/r$, where r is the perceived total effort requirement of the reward (i.e., the starting distance to the goal), and n_t is the amount of the requirements already fulfilled by the individual at time t. The observed measure d_t has a possible range of 1 to 0, with 1 occurring when no progress toward the goal has yet been made and 0 occurring when the goal is achieved. The goal-gradient hypothesis implies that the latent (unobserved) motivation at time t to achieve the goal (m_t^*) is a decreasing function of d_t, that is, $\partial m_t^*/\partial d_t < 0$.[1]

Because the underlying motivation to achieve the reward is unobserved, we model the customer's observed behavior (or effort investment). Since the observed effort should increase with stronger goal motivation, we expect greater effort with smaller proportional goal-distance (d_t). We employ different operationalizations of the goal-distance model that capture observed effort behavior as a function of d_t.

The discussion leads to the following hypothesis:

[1] Of course, consumers can still be sensitive to absolute magnitude, and the conceptualization of psychological goal-distance in proportional terms is likely to apply only within a reasonable empirical range. It is noteworthy that Hull (1934) reported that the goal-gradient of a 20-foot runway resembled a foreshortened (proportionally contracted) gradient from a 40-foot runway. This finding can be captured by modeling the rats' behavior as a function of proportional but not absolute goal-distance.

H1: Consumers will accelerate their efforts to earn a reward as the psychological distance (d_t) to the reward goal decreases.

Next, we test H1 using a real café RP. We subsequently extend the goal-gradient hypothesis to the particularities of consumer behavior, by exploring its implications for customer retention and investigating the effects of illusionary goal progress. The latter allows for a direct test of the effect of proportional versus absolute goal-distance and for ruling out rational accounts of the goal-gradient effect such as time-discounting.

3. THE CAFÉ REWARD PROGRAM

To allow for a strong and realistic test of intertemporal behavior, we employed a field study in which customers made real coffee purchases in the context of an actual café RP. By tracking purchases, we were able to test for purchase acceleration toward the reward goal (i.e., H1). The study included two control groups: (a) members from whom we "bought back" incomplete cards and (b) customers participating in an experimental control program that was identical to the actual RP except that purchasing coffees did not earn rewards. The inclusion of these control groups allowed us to compare the intertemporal purchase behavior of redeemers and non-redeemers (i.e., "loyals" and "defectors") and examine differences between reinforced and non-reinforced behavior. We also investigate alternative explanations using a variety of methodologies, including testing a key corollary termed "post-reward resetting," exploring the behavior of the two aforementioned control groups, and incorporating unobserved heterogeneity in the tendency to accelerate.

Method

The participants in the field study were customers of a café located within the campus of a large East Coast university. At the time of data collection, the café had several on-campus locations. Customers were offered to enroll for free in a café RP, in which they had to make ten coffee purchases in order to earn a reward. To allow tracking of their purchases, members were required to carry a frequent coffee buyer card (see Figure 2, left panel). They received one stamp on the card for each coffee purchase they made (only one stamp per visit was permitted). Stamps were printed using six-wheel automatic numbering machines that, unbeknownst to customers, sequentially numbered each stamp given out (these numbers did not resemble dates). Once members accumulated ten such stamps from any combination of the café locations, they were eligible for a free coffee redeemable on their next visit to one of the café locations.[2] Members were asked to indicate their name and e-mail address on the back of the card, which enabled us to match cards redeemed by the same member. Overall, we obtained 949 completed (i.e., redeemed) 10-stamp cards, recording nearly ten thousand coffee purchases.

Buyback of incomplete cards. The design of the café RP allowed us to collect only those cards that were completed and redeemed for a reward. Therefore, to sample from the broader member population (i.e., including those members that would otherwise fail to complete or redeem their card) we instituted a card buyback offer. Research assistants posing as café employees approached individual card-holding members and offered them to return their cards back to the café (regardless of the number of stamps on them) for a cash award of $4 per card and a 1% chance to win $100. Members were told that the cards were needed for the café's customer research. Overall, we acquired 73 buyback cards.

Recruitment of experimental control group using "transparent cards." We recruited 42 customers into an experimental control condition in which they carried "transparent cards." These cards were similar to the regular 10-stamp card but were marked on the back so they could not be redeemed for a reward. The control customers were randomly sampled from the population of program members. Research assistants (posing as café employees) intercepted customers who requested a regular program card and offered instead to enroll them in a "purchase habit" study designed to help the café management better understand their customers. Participants were asked to carry a "transparent" card and have it stamped every time they made a qualifying purchase at the café. They received $5 upon agreeing to participate in the study, and were told that they would receive $15 more when they surrendered their cards six weeks later, regardless of how many coffee purchases they made during that time. We verified that control participants did not use the regular RP cards during the study.

[2] Members could also earn a free baked good of equal monetary value (biscotti, croissant, or muffin). However, the majority (85%) of reward redemptions were for coffee and the results did not differ based on the redeemed reward.

Results

A plot of the raw mean inter-purchase times, aggregated across all redeemed cards (excluding "transparent" and buyback cards), demonstrates purchase acceleration as a function of smaller goal-distance (see Figure 3)[3]. Consistent with H1, as members accumulated more stamps on their card, the average length of time before their next coffee purchase decreased. The mean difference between the first and last observed inter-purchase times was 0.7 days ($t = 2.6$; $p < .05$), representing an average acceleration of 20% from the first to the last inter-purchase time. As an estimate of the overall effect of acceleration on the average card, one can compare the mean observed time to complete a card, which was 24.6 days, with the number of days it would have taken to complete a card at the rate of the first observed inter-purchase time, namely, 29.4 days. This yields a difference of nearly 5 days (16%) in card completion time.

Although the analysis of raw data provides preliminary support for purchase acceleration (H1), it does not account for a variety of important factors. Accordingly, we employ more sophisticated data analysis, using a discrete-time proportional hazard rate model. This modeling approach incorporates time-varying covariates and controls (e.g., weekly number of issued stamps) intended to rule out alternative explanations such as time-trend effects. Our modeling approach also permits accounting for unobserved heterogeneity (i.e., individual differences in base purchase rates and acceleration tendencies).

Hazard rate modeling methodology. In the main dataset, each row represented one day per customer on a card on which the customer could have made a purchase; we included the days after the first stamp was received up to the day of the last stamp. There are variables in the dataset at the customer level, day level, and card level. Overall, this dataset had 29,076 rows of data, based on 949 completed (i.e., redeemed) 10-stamp cards, capturing nearly ten thousand coffee purchases.

Hazard rate models (Cox 1972) are an important method to model inter-purchase times. In these models, the instantaneous probability of purchase (called the hazard function, $h(t)$) is estimated, conditional on the amount of time since the prior purchase. In the discrete-time model

[3] We excluded days on which the café was closed from the analysis.

(Gupta 1991; Helsen and Schmittlein 1993), the hazard model likelihood is decomposed into probabilities of purchase within given time intervals.

In a hazard rate model, the baseline continuous survival function $S(t)$ represents the probability of no purchase having occurred after time t has elapsed since the last purchase:

$$S(t) = \exp[-\int_0^t h(u)du] \tag{1}$$

We employed a discrete-time proportional hazard model parameterized as the discretized survival function. Following the derivation in Seetharaman and Chintagunta (2003), the full discretized survival function can be expressed as a function of the baseline hazard function $h(t)$, time-varying covariates $\boldsymbol{X}_t$ (including the proportional distance to the goal) and estimated covariate coefficients $\boldsymbol{\beta}$ (including a constant term):

$$S(t, \boldsymbol{X}_t) = \exp[-\sum_{u=1}^{t} e^{\boldsymbol{X}_u \boldsymbol{\beta}} \int_{u-1}^{u} h(w)dw] \tag{2}$$

In our application, we decompose the survival function into day-specific components, and our dependent variable is then the probability of purchase on a given day, conditional on no purchase having yet occurred, which is given by:

$$\Pr(t, \boldsymbol{X}_t) = 1 - \frac{S(t, \boldsymbol{X}_t)}{S(t-1, \boldsymbol{X}_{t-1})} = 1 - \exp[-e^{\boldsymbol{X}_t \boldsymbol{\beta}} \int_{t_l}^{t_u} h(u)du] = 1 - \left(\frac{S(t_u)}{S(t_l)} \right)^{e^{\boldsymbol{X}_t \boldsymbol{\beta}}} \tag{3}$$

Our analysis is done at the day level: we assume that each day is a potential purchase occasion, with the exception of days on which the café was closed. Since each day's probability is estimated as the difference in survival probability from start of day to end of day, it was necessary to code each day t as a range of continuous times between the lower bound t_l and the upper bound t_u when applying Equation 3. Purchases made on the same day were coded as occurring between time $t_l = 0$ and time $t_u = 0.5$; purchases on the subsequent day were coded as occurring between $t_l = .5$ and $t_u = 1.5$, and so forth. In the following likelihood function for an observed purchase, we denote the observed number of days elapsed at time of purchase by T, and code an indicator function δ_v representing whether a purchase did occur on a given day v ($\delta_v = 1$) or did not occur on day v ($\delta_v = 0$):

$$L = \prod_{v=0}^{T} \Pr(v, \boldsymbol{X}_v)^{\delta_v} [1 - \Pr(v, \boldsymbol{X}_v)]^{1-\delta_v} \quad (4)$$

The full likelihood is the product of all the purchase-specific likelihoods across cards and individuals (Seetharaman and Chintagunta 2003).

We determine the best-fitting base hazard function using the Schwarz BIC measure. The BIC measure trades off improvements in the log-likelihood against increases in the number of parameters. We employ latent classes to account for unobserved heterogeneity in the hazard rate parameters (Kamakura and Russell 1989), which is important in order to rule out heterogeneity bias. Since the unit of analysis in the latent class model is the customer, we are taking into account the common error variance when multiple cards belong to a single individual and we are specifically accounting for cross-customer unobserved heterogeneity.[4] The latent class modeling can be thought of as a nonparametric multivariate distribution on the hazard rate parameters across individuals, with each latent class representing a support point in the distribution. While we do find significant unobserved heterogeneity in the hazard rate parameters, when the models are run without latent class segmentation all of the results still hold. In all models, we used GAUSS software to implement Newton-Raphson maximum likelihood estimation and determined the number of latent classes using the BIC criterion.

Analyses of acceleration with common parameters across consumers. In this subsection, the primary focus of our modeling is the effect of goal-distance on inter-purchase times. Recall that the distance to the reward goal is captured via the measure $d_t = (r-n_t)/r$. In the café reward program, n_t is the number of stamps accumulated on the card at time t, and r is the total number of required stamps. Given that in the main dataset we model probability of purchase when there are between 1 and 9 stamps accumulated on the card (and $r = 10$), the measure d_t ranges between 0.9 and 0.1.

To test the goal-gradient hypothesis (H1) in this and the subsequent empirical applications, we constructed the goal-distance model (hereafter "GDM"). The GDM includes linear and quadratic parameters that capture the effect of goal-distance on observed behavior[5].

[4] All of the results replicate when we exclude from the analyses the subsequent cards of members who redeemed more than one card.
[5] In this and subsequent empirical applications, we examined higher order parameters using orthogonal contrast codes (Fisher and Yates 1957) but found these to be not significant.

Here, the GDM is added as a covariate in the proportional hazard model. The model is parameterized by defining the probability of purchase for customer i on a given day t as:

$$\Pr_i(t, \boldsymbol{X}_{it}) = 1 - \left(\frac{S(t_u)}{S(t_l)}\right)^g \text{ where } g = exp[\beta_0 + \beta_1 d_{it} + \beta_2 (d_{it} - \bar{d}_i)^2 + \gamma \boldsymbol{X}_{it}] \tag{5}$$

$S(t)$ = the baseline survival function,

d_{it} = the proportion of total distance remaining to the goal for individual i at time t,

$d_{it} - \bar{d}_i$ = the mean-centered proportion of total distance remaining to the goal,

β_1 and β_2 = the linear and quadratic goal-distance parameters, respectively,

$\boldsymbol{X}_{it}$ = the vector of covariates (i.e., control variables), and

γ = the corresponding vector of coefficients.

The parameters for estimation in the GDM are the intercept β_0, the goal-distance parameters β_1 and β_2, and the vector of coefficients γ. Consistent with H1, we expect the parameter β_1 to be smaller than zero, capturing the predicted increase in the probability of purchase (hazard rate) as a function of smaller goal-distance (d_{it}). It is noteworthy that if β_1 is greater than zero, we get "effort *de*celeration" (i.e., lower probability of purchase as a function of smaller goal-distance). Among the time-varying covariates, we included both the weekly number of issued stamps and a code for whether or not a given day was after the end of the spring classes in order to control for alternative explanations based on time trend.[6]

We accounted for unobserved heterogeneity in the base hazard rate parameters (but not in the goal-distance or the covariate parameters) using the methodology described earlier. Table 1 displays the estimated parameters for the log-logistic hazard rate function and the GDM shown in Equation 5.[7] The linear goal-distance parameter $\hat{\beta}_1$ was smaller than zero ($p < .01$). This result supports H1 and demonstrates that members accelerated their coffee purchases as they were closer to earning a free coffee. Consistent with the goal-gradient curve shown in Figure 3, the negative quadratic parameter $\hat{\beta}_2$ ($p < .01$) implies a diminishing rate of acceleration. A nested likelihood

[6] The control variables also included card type (indicating whether the member purchased American or Italian coffees) and additional time-varying covariates: midterm break (a code for whether or not the day was during the midterm break), day of week (linear trend from Monday through Thursday), and dummy codes for Friday and Saturday/Sunday. None of the covariates were allowed to vary across the latent classes.

[7] All covariates were normalized before model-fitting in this and the subsequent model estimations.

ratio test indicated that the GDM provided an improvement in fit over a "naïve" model in which β_1 and β_2 were restricted to zero ($\chi^2 = 18.8$, $df = 2$, $p < .01$).

Alternative Explanations

Although the observed and estimated purchase acceleration is consistent with H1 and the existence of a goal-gradient in RPs, several alternative explanations for this finding must be examined. One such rival account is that an unidentified time-trend effect led to a decrease in inter-purchase times. For example, members could have developed a card usage routine or an addiction to coffee (i.e., habituation). Relatedly, although café customers had no reason to expect the RP to expire, such concerns may have motivated members to accelerate their purchases.

Post-reward resetting. As previously discussed, we included in our models two time-varying covariates intended to statistically control for time-trend effects, namely the weekly number of issued stamps (essentially a control for sales trend) and whether or not the day was after the end of spring classes (when some students graduate). Nevertheless, to directly examine the time-trend and habituation accounts, we analyzed the post-redemption purchase behavior of 110 members who completed a first card and then reengaged in the program to complete a second card. These members demonstrated strong goal-gradients (i.e., faster inter-purchase times as a function of lower d_{it}) on both of their cards ($\hat{\beta}_1$ = -.06 and -.09, for 1st and 2nd card, respectively; both p's < .05).

According to the goal-gradient hypothesis, the motivation to invest effort increases with progress toward the reward threshold. Therefore, a corollary of this hypothesis is that after customers earn their first reward they should exhibit a post-reward resetting (i.e., a slowdown) in their purchase rates when starting to work toward their second reward, followed by a second pattern of purchase acceleration. In contrast, time-trend and habituation accounts predict monotonic acceleration in coffee purchases across the two cards, at least until some plateau or ceiling effect is reached. Therefore, according to these rival accounts, the inter-purchase times on the second card should be a direct continuation of the trend on the first card.

To contrast the resetting corollary with the time-trend and habituation accounts, we calibrated a non-parametric model with individual dummy codes representing each different inter-purchase time across the two cards. The advantage of using this non-parametric model is that we can separately estimate the 18 inter-purchase times for the two sequential cards, while controlling for the covariates $\boldsymbol{X}_{it}$ as well as for the unobserved heterogeneity in the base hazard rates. Figure 4 reveals a clear overlap between the plots of the inter-purchase times estimated for the two cards. The figure also shows that the first two inter-purchase times on the second card ($\overline{X}$ = 3.1 and 2.7 days, respectively) are substantially slower than the last two inter-purchase times on the first card ($\overline{X}$'s = 2.2 and 2.1 days, respectively), and are in fact similar to the first two inter-purchase times on the first card ($\overline{X}$'s = 3.2 and 2.8 days, respectively).

To statistically test for post-reward resetting while controlling for time-varying covariates and heterogeneity in the base hazard rates, we modeled just the last two inter-purchase times on the first card and the first two inter-purchase times on the second card. Instead of linear and quadratic goal-distance parameters, we included a contrast code for first versus second card. The first two inter-purchase times on the second card were slower than the last two inter-purchase times on the first card ($p < .01$). The test was in the same direction and significant when we compared only the first inter-purchase time on the second card with the last inter-purchase time on the first card ($p < .05$).

In summary, consistent with the goal-gradient hypothesis, purchase rates revealed a clear post-reward resetting. Members accelerated their coffee purchases toward their first reward (a free coffee) and then slowed down when starting to work toward a second similar reward; the same members subsequently re-accelerated their purchases as they approached the second reward. These findings rule out the habituation account as well as other forms of time-trend (e.g., graduation or expiration concerns). Next, we examine the purchase behavior of two control groups, which allows us to compare reinforced and non-reinforced behavior and further rule out alternative explanations.

Analysis of non-reinforced behavior (*"transparent cards"*). According to the goal-gradient hypothesis (H1), card-holding customers accelerate their purchases due to enhanced motivation closer to the goal (i.e., the reward threshold). A corollary of this hypothesis is that customers will fail to exhibit acceleration when carrying a transparent card (similar to the regular 10-stamp card but unredeemable for a reward), when their purchase behavior is not reinforced with any purchase-contingent reward. In contrast, if purchase acceleration reflects other factors, such as a time-trend in sales, then we would expect to observe similar acceleration among customers enrolled in the transparent card control group. The intertemporal purchase behavior of customers carrying transparent cards can serve as a benchmark or control for assessing the acceleration detected in the main dataset.

Unlike the main dataset, which included only complete (redeemed) 10-stamp cards, some of the transparent cards were incomplete. We therefore first analyzed all transparent cards that included at least three stamps. We estimated the GDM (shown in Equation 5) using the same log-logistic hazard function and accounting for unobserved heterogeneity in the base hazard rates. All the significant coefficients for the control variables have the same sign and same interpretation as in the model of redeemed cards. We added, however, a coefficient capturing the effect of the final number of stamps on the card in order to account for observed differences in base hazard rates between cards collected with fewer or more stamps.

In the GDM for transparent cards, we find *de*celeration ($\hat{\beta}_1 = .3$, $p < .01$) and no curvature ($\hat{\beta}_2 = .006$, $p > .1$). The inter-purchase times on the transparent cards steadily increased as customers neared the non- reinforced (unrewarded) completion of the card. It should be noted that this model was fit using data (i.e., days) only up to the last purchase on the transparent card. We find an even stronger deceleration effect ($\hat{\beta}_1 = .4$, $p < .01$; $\hat{\beta}_2 = -.1$, $p > .1$) when we include all observed days up to the collection of the transparent card (i.e., including days after the last purchase on incomplete transparent cards). Finally, we calibrated the GDM on the sub-sample of completed transparent cards, and again found deceleration ($\hat{\beta}_1 = .15$, $p < .05$; $\hat{\beta}_2 = .05$, $p > .1$).

Overall, the analysis of the transparent cards supports the notion that the purchase acceleration found in the main dataset was driven by goal motivation rather than time-trend effects or habituation. Although participants in the transparent card program were sampled from the population of RP members, the fact that their reward ($20) was not contingent on their purchase behavior reversed the tendency to accelerate toward the 10-stamp threshold.

Analysis of incomplete (*"buyback"*) *cards.* We believe that the sample of buyback cards differs from the sample of redeemed cards (used in the main analyses) in that members from whom we bought back cards exhibited a lack of goal-motivation. Buyback cards were in circulation for a period of time that was longer than that of redeemed cards ($\overline{X}$ = 65 days vs. $\overline{X}$ = 25 days, $t = 7.6$; $p < .01$), which suggests that without our intervention such buyback cards would have resulted in "breakage" (i.e., non-redeemed stamps or cards). This allows for comparing the intertemporal purchase behavior of non-redeemers ("defectors") and redeemers ("loyals").

We calibrated the GDM on the sample of buyback cards using the log-logistic hazard function and accounting for unobserved heterogeneity in the base hazard rates. We included buyback cards with at least three stamps and added the final number of stamps on the card as a covariate to account for observed differences in the base hazard rates of cards with fewer or more stamps. Again, all the significant coefficients for the control variables had the same sign and same interpretation as in the model of redeemed cards.

In the GDM for incomplete "buyback" cards, we found no linear effect of goal-distance ($\hat{\beta}_1 = .02$, $p > .1$) but there was a quadratic effect ($\hat{\beta}_2 = -.1$, $p < .05$). This pattern suggests that, unlike redeemers, buyback customers do not accelerate their purchases as a function of progress toward the reward. Moreover, when we model all observed days up to the buyback of the incomplete card (i.e., including days after the last purchase on the card), we find an increasing *de*celeration effect ($\hat{\beta}_1 = .3$ and $\hat{\beta}_2 = -.1$, respectively; both p's < .01). Overall, then, customers from whom we bought back incomplete cards ("defectors") differed from redeemers ("loyals") in that the former did not accelerate and even decelerated their purchases as a function of accumulated stamps.

Analysis of unobserved heterogeneity in goal-motivated acceleration. A final alternative explanation that must be considered is heterogeneity bias. Specifically, although we did account for unobserved heterogeneity in the base hazard rates, it is possible that the estimation of homogenous goal-distance parameters gave rise to an apparent goal-gradient that actually did not exist among a majority of individual members. We therefore calibrated the GDM on the main dataset using simultaneous (i.e., joint) estimation of unobserved heterogeneity in both the hazard rate and goal-distance parameters. Table 2 displays the estimated segment (class) sizes and segment-level parameters. Segment 1, the largest segment (58%), has significant linear and quadratic goal-distance parameters; this pattern of acceleration is similar to that obtained earlier with the homogenous goal-distance parameters. Segment 2 (29%) has the same linear goal-distance parameter as the largest segment, but the coefficient does not reach statistical significance ($p = .13$) due to the smaller segment size. Segments 3 and 4 (7% and 6%, respectively) both had a linear goal-distance parameter near zero. Overall, the segmentation analysis is *in*consistent with the heterogeneity bias rival account; a majority of significant accelerators are found after accounting for unobserved heterogeneity in the tendency to accelerate.

Evidence for the Goal-Gradient Hypothesis in the Café Reward Program: Discussion

Consistent with the goal-gradient hypothesis (H1), the findings from the café RP indicate that customers accelerated their purchases as a function of smaller goal-distance. The decrease in inter-purchase times was both observed in the raw data and estimated using a discrete-time proportional hazard rate model. The 20% (0.7 day) decrease in average inter-purchase times from the first to the last stamp on the card implies that in a typical month, on average, members purchased two more coffees than they would have without an RP in order to earn one free coffee.[8]

We ruled out several alternative explanations for the observed purchase acceleration using an experimental control (i.e., the transparent cards) and statistical controls. We also found that members exhibited a markedly similar acceleration pattern on two sequential cards, with a slowdown in purchase rates after earning the first reward and beginning to work toward the second. Such post-reward resetting is inconsistent with the rival accounts based on time-trend or habituation, whereas it is consistent with the notion that the motivation to expend effort depends on goal-distance.

The analysis of the incomplete (buyback) cards revealed that defectors were less likely to exhibit a goal-gradient than members who redeemed at least one card. This finding suggests that insufficient motivation to earn rewards underlies both defection (churn) and *de*celeration in purchase rates. In a subsequent section, we employ individual-level acceleration estimates to more systematically examine the relation between goal-motivation and customer retention. Next,

[8] The first inter-purchase time was 3.3 days, while the total card time yielded an average rate of 2.7 days per purchase. This is equivalent to 9 purchases per month based on the first inter-purchase time, compared to 11 purchases per month based on the observed rate that includes purchase acceleration.

we extend the goal-gradient hypothesis to the particularities of the human psychology of rewards, by exploring the effect of illusionary progress toward the goal.

4. THE ILLUSION OF PROGRESS TOWARD THE GOAL

Building on the behaviorist goal-gradient hypothesis, we predicted, and found, that customers accelerate their purchases closer to the reward threshold. Although this result is consistent with our conceptualization that goal proximity increases motivation, it could also be explained based on rational, cost-benefit grounds. In particular, as the distance to the reward is diminished, any additional unit of effort reduces a greater percentage of the remaining discrepancy to reward attainment. Additionally, time-discounting theories imply that (temporal) goal proximity enhances the value of rewards. Thus, the perceived benefit from an incremental unit of effort may increase closer to the reward threshold.

The rational explanations for purchase acceleration rely on the absolute distance to the reward. In contrast, our conceptualization suggests that the key determinant of goal motivation is the proportion of original distance remaining to the goal. The GDM captures this psychological quantity through the measure d_t, which is influenced not only by the absolute distance remaining to the reward r-n_t, but also by the perception of the original goal-distance r (i.e., the total a priori effort requirement for the reward). Thus, we posit that, ceteris paribus, goal motivation is influenced by the perceived rather than real progress to the goal. The perceived and real progress are distinct when the perception of the original goal-distance can be systematically manipulated without affecting the real, absolute distance remaining to the goal.

Marketers (or researchers) can create an "illusionary progress" toward the goal by increasing the total original distance to the reward (i.e., increasing r) while simultaneously increasing the perception of the distance (requirements) already completed (i.e., increasing n_t by the same quantity). Such a manipulation reduces the psychological (or proportional) distance to the reward, d_t, while holding constant the real, absolute remaining distance (i.e., the actual remaining requirements, r-n_t).[9] Accordingly, in the subsequent tests, we create illusionary progress by increasing the total requirements of a baseline RP, while simultaneously awarding consumers with an equivalent, yet bogus "head start" (i.e., bonus credit or points in the amount of the incremental requirements).

A manipulation of illusionary progress distinguishes between our goal-gradient conceptualization and the rational accounts of purchase acceleration. If motivation is influenced by the psychological distance to the reward, as defined by the proportion of original distance remaining to the goal (d_t), then illusionary progress should enhance goal motivation and consequently lead to increased efforts to earn the reward. In contrast, because illusionary progress does not affect the absolute (real) distance from the reward, cost-benefit calculations and time-discounting cannot account for the predicted effort acceleration.

The discussion leads to the following hypothesis:

H2: Illusionary progress toward the reward goal will motivate consumers to accelerate their efforts to earn the reward.

Hypothesis 2 predicts that illusionary progress will lead to faster completion of the reward requirements. We begin with a strong and realistic (field) test of H2, in which we examined actual purchase behavior. We then report the results of process tests, intended to rule out alternative explanations based on the idiosyncratic fit heuristic (Kivetz and Simonson 2003) and sunk cost (Thaler 1980).

A Field Experiment of Illusionary Goal Progress

Method. The participants were 108 customers of the café described earlier. They were randomly assigned to either a control condition or an experimental (illusionary goal progress) condition. Specifically, research assistants posing as café employees randomly offered customers either a 10-stamp or a 12-stamp coffee card (see Figure 2, left vs. right panel, respectively). The 10-stamp and the 12-stamp cards indicated that members were required to accumulate 10 and 12 coffee

[9] This is easy to verify algebraically by investigating the change in the function $d_t = (r\text{-}n_t)/r$ after an addition of Δ to both r and n_t. Specifically, $d_t^{+\Delta} = ([r+\Delta] - [n_t+\Delta])/(r+\Delta) = (r - n_t)/(r+\Delta) < (r - n_t)/r = d_t$. That is, unlike absolute goal-distances, proportional distances are affected by a common addition, such that $d_t^{+\Delta} < d_t$.

purchases (respectively) to earn one free coffee. However, customers assigned to the 12-stamp experimental condition received two pre-existing “bonus” stamps, described as an offer to anyone who opted to join the program. Thus, while the two groups faced identical effort requirements when joining the program (i.e., $r\text{-}n_t$ = accumulating 10 coffee purchases), the experimental group started with a lower proportion of original distance remaining to the goal than did the control group (i.e., $d_t^{+2} = 0.83$ vs. $d_t = 1.0$, respectively). All other aspects of the program were held constant across the two conditions and were identical to those described earlier for the café RP.

Results. Consistent with H2, the results indicate that illusionary goal progress led to faster completion of the reward requirement. On average, customers in the control condition completed the 10 required purchases (for the 10-stamp card) in 15.6 days. In contrast, customers in the experimental (illusionary progress) condition completed the 10 required purchases (for the corresponding “12-stamp” card) in a shorter time, only 12.7 days, nearly three days or 20% faster ($t = 2.0$, $p < .05$; *medians* = 15 vs. 10 days, $Z = 2.1$, $p < .05$ by a Mann-Whitney U test).

Process Tests of Illusionary Goal Progress

In these questionnaire-based experiments, we tested H2 and the alternative explanations using travelers who were waiting for trains at sitting areas in a major train station. Sixty-five travelers were randomly assigned to either a control or an experimental condition of a hypothetical frequent diner program offered by their favorite pizza chain. In the control condition, respondents were told that they would have to carry an eight-stamp card (shown in a picture), on which they would receive one stamp for each pizza meal they bought at the chain. Once they accumulated eight stamps, they would earn a free medium-size pizza of their choice. In the experimental condition, respondents were asked to evaluate a similar frequent diner program, except that they had to carry a ten-stamp card (i.e., the program supposedly required purchasing 10 pizza meals). In this condition, respondents were exposed to an illusionary progress to the reward goal. In particular, they were told that as a special offer for joining the program, they would receive two free bonus stamps (they were shown a 10-stamp card with the first two stamp-slots already checked). Thus, while, de facto, the control and the experimental groups faced identical effort requirements (i.e., $r\text{-}n_t = 8$ pizza meals), the proportion of original distance remaining to the goal was lower for the experimental compared to the control group (i.e., $d_t^{+2} = 0.8$ vs. $d_t = 1.0$, respectively).

Respondents in both conditions were first asked to rate the likelihood that they would join the program, using an 11-point scale ranging from 0 ("Definitely would not join") to 10 ("Definitely would join"). Then, respondents were told to assume that they actually joined and were asked to estimate how many weeks it would take them to complete the program. Consistent with H2 and the results of the field experiment, respondents in the experimental as opposed to control condition estimated completing the eight required purchases in fewer weeks ($\overline{X} = 11$ vs. $\overline{X} = 16$ weeks, $t = 1.6$; $p < .1$; *medians* = 8 vs. 12 weeks, $\chi^2 = 6.5$; $p < .05$ by a non-parametric median test).

Respondents' joining likelihood was elicited in order to rule out an alternative explanation – based on the idiosyncratic fit heuristic (Kivetz and Simonson 2003) – for the predicted illusionary progress effect. In particular, according to the idiosyncratic fit heuristic, consumers decide whether to join reward and other promotional programs on the basis of their individual fit (relative to typical other consumers) with the program. Therefore, in both the aforementioned field experiment and in the current test, we deliberately used a manipulation of illusionary progress that was not expected to affect respondents' idiosyncratic fit with the RP. Specifically, the two free bonus stamps offered in the experimental conditions were described as an offer to *anyone* who opted to join the program. Indeed, consistent with the notion that idiosyncratic fit was not affected by the manipulation of illusionary progress, we found no effect on the likelihood of joining the frequent diner program ($\overline{X} = 5.9$ vs. $\overline{X} = 5.8$ in the experimental vs. control condition, respectively, $t = .1$; *n.s.*)

Another alternative explanation for the observed illusionary progress effect is that the two bonus stamps were considered a (virtual) sunk cost (e.g., Thaler 1980). Relatedly, the bonus stamps may have enhanced the perceived value of the card, thereby leading to estimations of faster completion time. To examine this rival account, we randomly assigned 118 new respondents (sampled from the same population of travelers) to one of three conditions: (1) the previous illusionary progress experimental condition (i.e., a "10-stamp" card with two pre-existing "bonus" stamps); (2) the previous control condition (i.e., an 8-stamp card with no stamps yet); and (3) a "sunk cost" condition. The experimental and control conditions were identical to the corresponding conditions described earlier, except that respondents were asked to imagine that they had recently joined the program. In the "sunk cost" condition, respondents were asked to imagine that they had recently joined a 10-stamp frequent diner program, had made two pizza meal purchases, and therefore, had two stamps on their card. Thus, the sunk cost condition was identical in all aspects to the experimental condition (including the picture of a 10-stamp card with two stamps already on it), except that the two stamps were due to the respondent's own purchase effort. It should also be emphasized that all three conditions entailed the same absolute remaining distance to the reward goal (eight additional pizza purchases).

In all three conditions, respondents were asked to imagine that they lost their current frequent diner card. They were then asked to rate on four scales how sad, mad at themselves, upset, and disappointed they would feel due to the loss of the card. Ratings were made on four seven-point scales, ranging from 1 (e.g., "Not at all sad") to 7 (e.g., "Very sad"). The scales were averaged into a single measure of feeling valence ($\alpha = .88$).

Respondents' ratings indicated that they felt worse about losing their card in the "sunk cost" condition ($\bar{X} = 2.1$) than in either the experimental (illusionary progress) condition ($\bar{X} = 1.7$; $t = 1.9$, $p < .05$) or the control condition ($\bar{X} = 1.7$; $t = 1.9$, $p < .05$). Importantly, there was no difference in feeling valence between experimental and control respondents ($t = .2$; *n.s*). These results are *in*consistent with the sunk cost alternative explanation. In particular, if illusionary progress gives rise to a faster purchase rate because bonus credits are construed as sunk cost, then we would expect respondents to feel a greater sense of loss after losing the experimental rather than the control card. The fact that the "sunk cost" respondents did feel worse about losing their card rules out the possibility that a measurement problem gave rise to the similarity in (good) feelings between the experimental and the control respondents.

The Illusion of Progress toward the Goal: Discussion

The effect of illusionary goal progress provides direct support for our proposition that psychological goal-distance (d_t) is a key determinant of achievement motivation and willingness to invest effort. Whereas this effect is consistent with the GDM and the conceptualization of proportional goal-distance, it is inconsistent with the rational accounts of purchase acceleration. That is, illusionary goal progress does not reduce the absolute distance to, or delay of, the reward and should therefore have no effect, according to cost-benefit and time-discounting explanations.

5. THE JABOOM! MUSIC-RATING INCENTIVE PROGRAM

So far, our test of the goal-gradient (H1) was operationalized using acceleration in inter-purchase times. In this section, we generalize the findings to effort involving repeatedly rating music and to goal-gradients operationalized through acceleration in both inter-visit times and rating quantities, as well as through increased persistence closer to the goal. We analyze secondary data obtained from a real RP, in which participants earned rewards for rating songs over the Internet. Next, we describe the music-rating program in detail.

The Methodology of the Music-Rating Incentive Program

The music-rating program was launched by MoodLogic, Inc., a technology company that develops and sells music organization software. The company initiated the program to build a database of music perceptions and tastes (required for its music organizers and preference engines). The program, labeled "Jaboom!," was operated on a dedicated website (members were addressed on the site as "Jaboomers!"). Internet users, recruited through an e-mail marketing campaign, were offered free enrollment in the RP, in which they had to rate 51 songs on the Jaboom! website to earn a $25 Amazon.com certificate. The RP, which was presented to participants as an ongoing program, continued for a period of 24 months after our observation period. Thus, expiration concerns should not have affected the behavior of program members.

Upon joining, members were asked to provide a valid e-mail address and select a unique login name and password. This information was used to determine the dates of each member's site visits and the number of songs the individual member rated on each such visit. There were no constraints on the number of songs that could be rated in a single visit or the number of certificates that could be earned by a single member.

Members could rate songs from one of six genres of their choice (e.g., rock, country, jazz) and could skip a given song or terminate their rating session at any point. Each song was rated on approximately 50 scales, while the member repeatedly heard the same 30-second song snippet. The scales elicited subjective perceptions and tastes (e.g., mood and likeability of the song) and more objective judgments (e.g., predominating instruments). On average, it took about four minutes to rate a typical song. A screen shot of the Jaboom! music-rating interface is shown in Figure 5.

Results

The dataset includes the rating behavior of 148 members, who rated a total of 14,866 songs in 472 separate website visits. Given that during most website visits (96%) members rated multiple songs, we examined both the quantity of ratings in each visit and the inter-visit times. Inter-visit times were modeled using the discrete-time proportional hazard rate model described earlier. The findings were similar to those of the café RP and are briefly described here. We then proceed to report the analyses and results of the quantity and persistence operationalizations of the goal-gradient hypothesis.

Tests of the Goal-Gradient Hypothesis Using Inter-Visit Times

In our dataset, there were 371 attempted reward certificates (i.e., with at least one song rated toward the certificate) and 262 earned certificates (i.e., with 51 songs rated). Of the 262 earned certificates, 114 were completed in a single visit. While such single-visit certificates arguably signify strong goal motivation, they need to be excluded from the analyses of inter-visit times, since they do not permit testing for acceleration (or deceleration). The hazard rate version of the GDM was therefore calibrated on the data from the remaining 148 certificates, which were earned in two or more visits. The dataset was constructed such that each row represented a unique day on which a particular member could have visited the Jaboom! website; we included the days after the first visit had occurred up to the day of the last visit.[10]

We parameterized the survival function using the GDM shown in Equation 5 and employed the likelihood function shown in Equation 4. Recall that the variable capturing the hypothesized acceleration is $d_t = (r\text{-}n_t)/r$, that is, the proportion of total distance remaining before the goal. In this empirical application, r = 51 songs (i.e., the original total distance to the

[10] In this and the subsequent sub-sections, we obtain similar results when we exclude from the analyses the data from subsequent certificates of members who earned more than one certificate.

goal) and n_t is the number of song-ratings accumulated at time t toward earning the reward certificate. Given that we model the probability of visiting when there are between 0 and 50 song ratings accumulated, the measure d_t ranges between 1 (when no goal progress has yet been made) and 0.02 (when the goal is almost achieved), respectively.

Table 3 displays the estimated parameters for the GDM, using the Weibull hazard parameterization. The linear goal-distance parameter was smaller than zero ($p < .01$). This result supports H1 and demonstrates that members visited the Jaboom! website more frequently as they were closer to earning the reward (i.e., inter-visit times decreased as a function of goal proximity). Consistent with Hull's findings with rats and the results of the café RP, the rate of acceleration was diminishing, as indicated by the negative quadratic goal-distance parameter ($p <$ .01). It is noteworthy that the model includes two variables as controls: (1) the total number of visits to the Jaboom! website at the day level, intended to control for time trend effects, and (2) the number of visits it took to complete each certificate, intended to rule out aggregation bias. We obtained similar results when either or both of these controls were excluded from the model. Additionally, the linear and quadratic goal-distance parameters remained significant and did not vary in magnitude across models with different numbers of latent classes in the base hazard rate.

We also calibrated the GDM using simultaneous estimation of unobserved heterogeneity in the acceleration and hazard rate parameters. Although the BIC criterion favors a one-class solution, we report the two-class solution, in order to rule out heterogeneity bias. In the larger class (65%), we find linear acceleration ($\hat{\beta}_1 = -.6, p < .1;\ \hat{\beta}_2 = -.1, p > .1$), while in the smaller class (35%) we find non-significant linear acceleration ($\hat{\beta}_1 = -.1, p > .1;\ \hat{\beta}_2 = -.3, p < .1$). The finding that both segments demonstrate acceleration is inconsistent with the heterogeneity bias rival account.

Finally, we recalibrated the same GDM on the data obtained from "incomplete" certificates (i.e., with less than 51 songs rated toward the unattained certificate). There were 36 such incomplete certificates with at least two rating visits (i.e., at least one inter-visit time that could be modeled as a function of goal-distance). Consistent with the analysis of the incomplete cards in the café RP, we found *de*celeration ($\hat{\beta}_1 = .7, p < .1;\ \hat{\beta}_2 = .01, p > .9$). Further, we calibrated the GDM on a dataset combining both complete and incomplete certificates and added an interaction term β_{INT} between completion (yes vs. no) and linear acceleration. We found significantly stronger linear acceleration for complete than incomplete certificates ($\hat{\beta}_{INT}$ = -.5; Wald $\chi^2 = 16.2$, $p < .01$). That is, while acceleration is related to retention and goal attainment, deceleration is associated with program defection and goal abandonment.

Tests of the Goal-Gradient Hypothesis using Rating Quantity

So far, we have examined how goal-distance influences inter-purchase and inter-visit times. In many cases, however, customers can accelerate their efforts by increasing the quantity of credits earned in a given interaction with the RP. Accordingly, in this subsection, we test the goal-gradient hypothesis by examining whether the quantity of songs rated per visit increases with goal proximity. We report separate tests of pooled- and segment-level quantity acceleration, behavior on incomplete certificates, and post-reward resetting.

Figure 6 presents the raw data obtained from the 148 complete certificates (with at least two visits) that we observed. Consistent with the goal-gradient hypothesis, members rated more songs in later visits, as they approached the 51 song goal.

To model the quantity of songs rated in a visit, we employed a Type I Tobit model (Tobin 1958). This approach allows us to capture the underlying tendency to rate songs while taking into account the constraint of 51 songs per certificate. Specifically, in the final visit of each certificate, the number of song-ratings is censored at 51 - n_t, where n_t is the number of song-ratings accumulated toward the certificate at the start of visit t. By definition, any additional songs rated in visit t beyond the 51 - n_t limit would be credited toward the next certificate rather than the goal of earning the present 51 song certificate. Relatedly, as 51 - n_t approaches zero, the inherent motivation to rate songs may increase (because $d_t \rightarrow 0$), but the 51 song constraint (i.e., the reward threshold) would suppress such an effect.

For uncensored data, the Tobit model is equivalent to maximum likelihood regression. However, because applying a regression to censored data leads to biased parameter estimates (Breen 1996), the Tobit models the censored data as the probability of rating 51 - n_t or more songs.

Thus, we model the quantity Qit rated in each visit t by participant i, incorporating the GDM into a Type I Tobit model:

$$Q_i(t, \boldsymbol{X}_{it}) = \begin{cases} g & \text{if } g < 51\text{-}n_{it} \\ 51 - n_{it} & \text{if } g > 51\text{-}n_{it} \end{cases} \quad \text{where } g = \beta_0 + \beta_1\, d_{it} + \beta_2\,(d_{it} - \bar{d}_i)^2 + \boldsymbol{\gamma X}_{it} + \varepsilon,\ \varepsilon \sim \mathrm{N}(0, \sigma^2) \quad (6)$$

Here, g can be thought of as a latent variable representing the tendency to rate songs, which can be directly observed in the data Q_{it} only when sufficient songs remain on the certificate (i.e., when $g < 51 - n_{it}$). This model generalizes the GDM to the domain of effort quantity and employs the operationalization of goal-distance (d_{it}) used earlier to test for temporal goal-gradients. We expect the parameter β_1 to be negative, indicating that a smaller goal-distance leads to a greater quantity of song-ratings. We also included a visit-level control that captures daily variations in the total number of songs rated on the Jaboom! website, as well as a variable representing certificates already earned in the visit by the participant.

Following the exposition of Breen (1996), the likelihood function is defined as:

$$L_i = \sum\nolimits_1 (2\pi\sigma^2)^{-.5} \exp\left(\frac{Q_{it} - \beta_0 - \beta_1 d_{it} - \beta_2 (d_{it} - \bar{d}_i)^2 - \boldsymbol{\gamma X}_{it}}{2\sigma^2}\right)$$
$$+ \sum\nolimits_0 1 - \Phi\left(\frac{\beta_0 + \beta_1 d_{it} + \beta_2 (d_{it} - \bar{d}_i)^2 + \boldsymbol{\gamma X}_{it} - (51 - n_{it})}{\sigma}\right) \quad (7)$$

Here, the first sum Σ_1 is the likelihood function for the linear regression summed over all the non-censored cases (i.e. visits in which the certificate is not completed). The second sum Σ_0 is the probability of observing at least the censored amount $51 - n_{it}$, defined by the standard normal cdf Φ and summed over the visits in which a certificate was completed (i.e. the censored cases). Note that in this model, the error variance σ^2 is a separate parameter to be estimated.

This model was calibrated on the data from the 148 certificates that were earned in two or more visits. Consistent with the goal-gradient hypothesis, we found a negative linear goal-distance parameter ($p < .01$, see Table 4), indicating that members rated more songs per visit when they were closer to the goal. There was also a quadratic effect ($p < .01$), indicating a diminishing rate of acceleration. In addition, we found a negative effect of the number of certificates already earned during the same visit ($p < .01$), consistent with satiation or fatigue. We obtained similar results when we forced the constant and error variance to vary across latent classes.

We also estimated the GDM using data from the 36 "incomplete" certificates (certificates with less than 51 song ratings but at least two visits). We found a weaker (although significant) effect of goal-distance (d_{it}) in this data ($\hat{\beta}_1$ = -2.5, $p < .01$; $\hat{\beta}_1$ = 2.9, $p < .01$). Further, we calibrated the GDM on a dataset combining both complete and incomplete certificates and added an interaction term β_{INT} between completion (yes vs. no) and linear acceleration. Consistent with the findings from the analyses of inter-visit times, we found that completed certificates exhibited a significantly stronger quantity goal-gradient than did incomplete certificates ($\hat{\beta}_{INT}$ = -8.5, $p < .01$).

Post-reward resetting. An alternative explanation for the observed quantity acceleration is learning (or habituation). For example, it is possible that with repeated visits to the Jaboom! website members learned to rate songs faster, thereby rating more songs in later visits. To examine the learning alternative explanation, we tested whether the 34 Jaboomers! who earned two multi-visit certificates exhibited post-reward resetting. The resetting corollary predicts that members should rate more songs on the last visit of their first certificate (when $d_{it} < 1$) than on the first visit of their second certificate (when $d_{it} = 1$). These members demonstrated strong goal-gradients (i.e., greater rating quantities as a function of lower d_t) on both of their earned certificates ($\hat{\beta}_1$ = -11.6 and -9.0, for 1st and 2nd certificate, respectively; both p's < .01). Moreover, consistent with post-reward resetting and inconsistent with the learning explanation, these members rated an average of 24 songs on the last visit of their first certificate compared to 16 songs on the first visit of their second certificate (pairwise $t = 2.3$; $p < .05$).

Unobserved heterogeneity in goal-motivated acceleration. To rule out heterogeneity bias as an explanation for the quantity goal-gradient, we calibrated the GDM on the data from completed certificates using simultaneous estimation of unobserved heterogeneity in the goal-distance parameters, the constant, and the error variance. Although the BIC criterion favored a one-class solution, we report the two-class solution. The size of the larger class was 97% ($\hat{\beta}_1$ = 12.4, $p < .01$; $\hat{\beta}_2$ = -7.4, $p < .01$) and the smaller class was 3% ($\hat{\beta}_1$ = -13.6, $p < .01$;

$\hat{\beta}_2$ = 1.5, $p > .7$). Thus, inconsistent with heterogeneity bias, both segments demonstrate acceleration in rating quantities as a function of goal proximity.

<u>Tests of the Goal-Gradient Hypothesis using Persistence in Effort</u>

In this subsection, we generalize the goal-gradient hypothesis to the domain of effort persistence. Based on the notion that the motivation to achieve a goal increases with its proximity, we predict that consumers will be more likely to persist in their effort when the reward is proximal, and equivalently, will be more likely to cease their efforts when the reward is distal. In the context of the music-rating program, this prediction implies that the more songs the member has accumulated toward the 51 song goal (i.e., the smaller is d_t), the less likely the member would be to end an active song-rating visit.

To test the persistence version of the goal-gradient hypothesis, we model the probability of terminating a visit at any point in the song-rating process as a function of the distance to the goal, d_t, as well as other covariates. We employ a logit model that accounts for unobserved heterogeneity in the baseline probability of ending a visit. The probability of participant i terminating the visit t after rating q songs in that visit is given by:

$$\Pr_i(t, \boldsymbol{X}_{itq}) = \exp(g)/(1 + \exp(g)) \text{ where } g = \beta_0 + \beta_1 d_{itq} + \beta_2 (d_{itq} - \bar{d}_i)^2 + \beta_3 C_{itq} + \boldsymbol{\gamma X}_{itq} \quad (8)$$

Equation 8 generalizes the GDM to the domain of effort persistence and employs the definition of proportional goal-distance used earlier. In particular, $d_{itq} = (r\text{-}n_{itq})/r$, where $r = 51$ song-ratings and n_{itq} is the number of songs accumulated by individual i toward earning the reward after rating q songs in visit t ($d_{itq} \in [0.02, 1.0]$). The parameter β_1 captures the effect of goal-distance on the probability of terminating a visit. We expect β_1 to be positive, indicating that a greater goal-distance (d_{itq}) leads to a higher likelihood of defection (and equivalently, a smaller d_{itq} leads to enhanced effort persistence). We used a dummy variable C_{itq} that was coded as 1 if the previous rating in the visit earned a certificate (i.e., if $d_{itq} = 1$) and 0 otherwise. The goal-gradient hypothesis predicts an increased likelihood of visit termination (or RP defection) when d_{itq} reverts to 1.0, and therefore, β_3 was expected to be positive. We estimated the effect of $d_{itq} = 1.0$ separately in order to guarantee that the hypothesized goal-gradient effect (captured via

β_1) could not be explained solely on the basis of an increased likelihood of defection at $d_{itq} = 1.0$, that is, just after reward attainment. We also included a day-level control that captures daily variations in the total number of songs rated on the Jaboom! website.

The GDM was calibrated on the entire dataset including ratings (and participants) that did not eventually earn a reward certificate. Thus, we jointly model visit termination and overall program defection. Table 5 displays the model estimates. Consistent with the goal-gradient hypothesis, the linear goal-distance parameter $\hat{\beta}_1$ was positive ($p < .01$), indicating that members were more likely to defect when they were farther away from the reward goal (or equivalently, more likely to persist when they were closer to the goal). Further, as predicted, $\hat{\beta}_3$ was also positive ($p < .05$), indicating that the highest probability of terminating a visit occurred just after goal achievement, when members were the farthest away from the (new) goal (at $d_{itq} = 1.0$). Figure 7 illustrates these results and shows that a majority (17% or 80/473) of all visit terminations occurred at $d_{itq} = 1.0$. The figure also shows that the percentage of visit terminations continues to decrease as a function of smaller distance to the goal. For example, among the 473 terminations observed, 8.7% (41/473) occurred when only one song was accumulated toward the next reward (i.e., when $d_{itq} = 0.98$), whereas only 0.2% of visit terminations (1/473) occurred when as many as 50 songs were accumulated (i.e., when $d_{itq} = 0.02$).

The Music-Rating Program: Discussion

The Jaboom! music-rating program allowed us to test the goal-gradient hypothesis in an empirical context that was very different from the café RP, using an incentive system that is akin to a freelance employment contract. Using this empirical application, we were able not only to replicate the finding of goal-motivated acceleration in visit rates (in the context of website rather than café visits), but also to extend the goal-gradient effect to the domain of quantity decisions and effort persistence. The various operationalizations of the goal-gradient hypothesis were examined using several generalizations of the GDM, which all relied on a common measure of goal-distance, namely d_t. These varieties of the GDM included a hazard rate model of the timing of visiting the Jaboom! website, a Tobit model of the quantity of ratings per visit, and a logit model of the probability of effort termination and program defection.

We found significant goal-gradients even after we accounted for unobserved heterogeneity and statistically controlled for time-trends in visit and song-rating frequencies. Moreover, we observed the phenomenon of post-reward resetting, whereby members accelerated toward each of two subsequent rewards, but exhibited a drop in their rating quantities after earning the first reward and starting to work toward the second (i.e., when d_t reverts to 1.0). The finding of post-reward resetting is a key corollary of the goal-gradient hypothesis; it demonstrates that effort expenditure is a function of goal-distance and rules out learning and other time-trend effects as well as a self-selection (or survivor) rival account. Self-selection is also inconsistent with the analysis of participants' entire sequence of ratings (including ratings that did not eventually lead to reward), which revealed significant goal-gradients.[11] Next, we employ individual differences to explore the relation between the goal-gradient and retention.

6. IMPLICATIONS OF THE GOAL-GRADIENT FOR CUSTOMER RETENTION

Prior research with animals showed that a steeper goal-gradient was generated by an increased drive (e.g., hunger) to attain the reward (Hull 1934). This finding suggests that RP members who exhibit enhanced acceleration possess a stronger motivation to earn free rewards (due to a higher achievement motivation or subjective reward valuation). If indeed the motivation to earn free rewards is related to the steepness of the goal-gradient, then individual differences in the tendency to accelerate (holding constant the overall program effort) should predict customer retention after attainment of the first reward. Specifically, we expect members who accelerate more strongly toward their first reward to be more likely to reengage in the program and earn a second reward. Relatedly, stronger acceleration is expected to lead to faster reengagement in the RP.

To examine these predictions, we recalibrated the GDM with unobserved heterogeneity in both the hazard rate and the goal-distance parameters, using a sub-sample of the café RP data that excluded subsequent cards. We used this model to obtain individual-level linear acceleration

[11] Detailed analyses that rule out the self-selection (survivor) account are available upon request from the authors.

estimates, based on the member's first coffee card. These estimates were calculated by multiplying the latent class parameters of goal-distance (i.e., acceleration) by the individual-level probabilities of class membership. The individual-level acceleration estimates were then used as independent variables in predicting member reengagement in the café RP. Importantly, the tests reported below included covariates that statistically controlled for the length of time it took to complete the first card (i.e., the individual's overall program effort and product liking) and the date of completion of this card. That is, we investigated the effect of individual differences in the slope of the goal-gradient (the linear goal-distance parameter), holding constant the base hazard rate (i.e., the average inter-purchase time) and possible seasonality (or right-censoring) effects.

Retention probability. We used a logistic regression to test the prediction that members who accelerate more strongly toward their first reward will be more likely to subsequently earn a second reward. The (dummy) dependent variable received a value of 1 if the member earned a second reward and 0 otherwise. The effect of the first card individual-level acceleration estimate for the member was in the hypothesized direction ($\hat{\beta}$ = -.5, *Wald*-χ^2 = 5.3; p < .05). In particular, members who accelerated more strongly toward their first reward were more likely to earn a second reward. In order to visually demonstrate this effect, we also split the sample, based on first card estimated linear acceleration, into three equally sized groups: decelerators (mean $\hat{\beta}$ = .1), accelerators (mean $\hat{\beta}$ = -.05) and strong accelerators (mean $\hat{\beta}$ = -.1). Figure 8 (left panel) depicts the probability of completing a second card (based on the raw data) for each of these three groups.

Reengagement time. To test the hypothesis that steeper acceleration predicts faster reengagement, we analyzed the sub-sample of 110 members who completed both a first and a second card. Reengagement time was computed as the period between the last purchase on the first card and the first purchase on the second card. We fit a new (Weibull) hazard rate model predicting the reengagement times using the individual-level estimates of first card acceleration as an independent variable. We also included the covariates used earlier in the hazard rate model of inter-purchase times and controlled for the duration and completion date of the first card. The effect of accelerating toward the first reward on the time to reengage in the program was in the hypothesized direction ($\hat{\beta}$ = -2.6, *Wald*-χ^2 = 13.8; p < .01). That is, members who accelerated more strongly toward their first reward were faster to begin working toward their second reward (holding constant the base hazard rate, or the average inter-purchase time, on the first card). Figure 8 (right panel) illustrates this result using the aforementioned tertiary split.

Implications for customer retention: Discussion. We posited that individual differences in the goal-gradient capture variations in the motivation to earn free rewards. Consistent with this argument, we found that customers who accelerated more strongly toward their first reward subsequently exhibited greater retention and faster reengagement in the café RP. These effects were replicated in the context of the Jaboom! music-rating program.[12] Given the previously reported findings of post-reward resetting, the present results cannot be explained as a simple continuation of the increased purchase rates of accelerators. Overall, the findings underscore the importance of incorporating the goal-gradient construct in the modeling and analysis of RPs.

7. GENERAL DISCUSSION

The goal-gradient is one of the classic phenomena discovered in the animal learning and behaviorism literature of the early twentieth-century. It has important implications for achievement motivation and goal pursuit, but nevertheless has been understudied in humans. This is particularly surprising given that the goal-gradient hypothesis provides considerable insights into the psychology of rewards and the optimal design of customer and employee incentive systems. In this research, we extended the goal-gradient hypothesis to the domain of consumer behavior and investigated its consequences for illusionary goal progress and customer retention.

The present research can be viewed as part of the ongoing (fruitful) attempt to bridge the consumer behavior and marketing science disciplines (e.g., Bell and Lattin 2000; Hardie, Johnson, and Fader 1993; Kivetz, Netzer, and Srinivasan 2004; Simonson and Winer 1992; Wertenbroch 1998; Winer 1986; for related discussion see Wittink 2004). Such intra-disciplinary endeavors often test behavioral theories using econometric modeling, secondary data, and/or field studies. In the current research, we built on prior analyses in behaviorism, social-cognition, and behavioral decision research and employed a variety of modeling frameworks and empirical

[12] Detailed results are available upon request from the authors.

tests in the context of two real incentive systems. We primarily relied on field experiments and econometric analyses of actual multi-period customer behavior. Such methodologies are crucial for studying dynamic goal pursuit and intertemporal responses to RPs and other promotions (see, e.g., Gupta 1988; Simonson 1990; Van Heerde, Leeflang, and Wittink 2000). The alternative approach, whereby respondents are asked to assume a hypothetical state (e.g., "imagine that you have accumulated *x* points") provides an adequate test of lay theories and self-perception, but not of the actual evolution of goal motivation and behavior.

Key Findings and their Implications

We found that members of a café RP accelerated their coffee purchases as they progressed toward earning a free coffee. The goal-gradient effect also generalized to a very different incentive system, in which shorter goal-distance led members to visit a song-rating website more frequently, rate more songs during each visit, and persist longer in the rating effort. Importantly, in both incentive systems we observed the phenomenon of post-reward resetting, whereby customers who accelerated toward their first reward exhibited a slowdown in their efforts when starting to work (and subsequently accelerating) toward their second reward. To the best of our knowledge, this article is the first to demonstrate unequivocal, systematic behavioral goal-gradients in the context of the human psychology of rewards.

For marketers, the goal-gradient may provide profitable opportunities. In addition to facilitating segmentation, targeting, and promotions (discussed later), the goal-gradient may lead to a sales lift that exceeds the cost of the reward. For example, the results of the café RP imply that, to earn one free coffee, customers bought two more coffees than they would have otherwise.[13] At the same time, consumers may derive pleasure from working toward future goals. This idea is consistent with the findings of an observational study, in which we had research assistants unobtrusively record the behavior and affect of the café customers. The results indicated that customers who participated in the RP -- as opposed to customers who did not -- were more likely

[13] Given the exclusivity of the café on campus, the observed purchase acceleration is likely a consequence of increased consumption rather than brand switching. Brand switching was not a possibility in the case of the music-rating program.

to smile when buying coffee (3.8 vs. 3.4 on a 5-point scale; $p < .05$), chat for a few minutes with café employees (26% vs. 7%; $p < .05$), say “thank you” (95% vs. 87%; *n.s.*), and leave a tip (21% vs. 3%; $p < .01$). Although these results should be interpreted with caution because customers self-select into the RP, they suggest that goal striving is intrinsically motivating above and beyond extrinsic rewards.

We posited that people are influenced by the proportional (or psychological) distance to the goal (i.e., $d_t = (r-n_t)/r$). Accordingly, we proposed that the illusion of progress toward the goal would enhance achievement motivation by reducing the perceived proportion of distance remaining to the goal. One test of this hypothesis involved a field experiment, in which customers who received a 12-stamp card with two pre-existing “bonus” stamps completed the 10 required coffee purchases faster than customers who received a “regular” 10-stamp card. It is noteworthy that the goal-gradient and illusionary progress effects can be captured by a mathematically-equivalent GDM in which effort depends on the proportion of original goal-distance *already* accomplished (i.e., $d_t = n_t/r$). Future research could explore the impact of framing goal progress in terms of completed versus remaining effort.

The illusion of goal progress and its boundary conditions merit further research. Beyond its theoretical importance, this phenomenon has substantial managerial implications for the design of RPs and other incentive systems. Currently, many RPs award “bonus points” to new members (e.g., *AMEX Membership Rewards Program* and *Hyatt Gold Passport*). Given the rich, complex structure of such programs, it is easy for managers to increase the point-requirements of rewards by an amount equivalent to the bonus, effectively creating illusionary goal progress.

Consistent with the notion that a steeper goal-gradient implies a greater drive to achieve the reward, we found that stronger accelerators reengaged in the program faster and were more likely to earn a second reward. Relatedly, failure to persist in the effort stream and fulfill the requirements was associated with weaker acceleration and even *de*celeration. The relation between the goal-gradient and retention was also evident in the finding that, just after reward attainment (when goal-distance regressed to 100%), customers exhibited a drop in activity (post-reward resetting) and were also most likely to defect. These findings have important implications for customer segmentation and the design of marketing interventions aimed at reducing churn. For example, it is particularly important to communicate with, and motivate, customers immediately after they earn a reward.

Extending the goal-distance model. Using a common measure of goal-distance d_t and logit, hazard rate, and Tobit frameworks, the GDM captured three forms of goal-gradients, namely increased persistence, rate, and quantity of effort closer to the reward threshold. These three goal-gradients predict increases in the recency of the last transaction, the average frequency of all transactions, and the average monetary value of these transactions; thus, the goal-gradient and its modeling have important implications for the widely-used RFM approach.

Given the robustness and generality of the GDM, we believe that it can be applied to a broad range of goal-based motivational systems, including more complex incentive systems. Future research can employ the GDM to account for consumer and employee behavior in sophisticated incentive systems, such as those used by airlines, retailers, and sales organizations; in such programs, r and n_t are often expressed in terms of miles, points, and dollars or units sold. Such RPs employ rich, complex structures that offer a multitude of different rewards at varying requirement levels. Customers can exhibit goal-gradients in a variety of ways, including purchase timing and quantity acceleration and increased retention and lock-in. However, the goal-gradient effect may be harder to detect in such complex RPs, for the following reasons. First, a priori, the researcher cannot identify (or observe) the consumer's goal. Second, the observed ex post goal (based on the actual reward redeemed) is self-selected by the consumer, and thus, it is difficult to draw causal inferences about differences in the behavior of consumers who redeem different rewards. Third, in the presence of multiple effort-reward combinations, the consumer's chosen goal may change during the program, thereby complicating the investigation of the goal-gradient hypothesis. Finally, the issue of right-censoring in observed behavior (relative to underlying motivation) that arises in the test of the music-rating quantity acceleration applies to complex RPs as well. This last problem can be solved by using the Tobit version of the GDM. Despite the various challenges, we believe that capturing the goal-gradient in more complex situations is a worthy endeavor, which the GDM can facilitate.

The implications of the goal-gradient for promotion, pricing, and competition. The goal-gradient effect has important implications, which merit further research, for key marketing variables. The findings suggest that goal proximity increases customers' responsiveness to credit-earning promotions and offers. For example, frequent flyers' willingness to purchase miles may increase closer to program goals. This hypothesis is consistent with the results of an unpublished study, in which we asked 329 respondents to imagine that they participated in a frequent flyer program that offered a free domestic round-trip ticket for accumulating 25,000 miles. Respondents were told that they had already accumulated either 13,000 or 23,000 miles (distant vs. near goal, respectively; manipulated between subjects); they were asked to indicate whether they would agree to receive weekly marketing e-mails in return for 1,000 bonus frequent flyer miles. As predicted, respondents in the near- as opposed to distant-reward condition were more likely to accept the promotional offer (56% vs. 38%, $\chi^2 = 10.7$; $p < .001$). It is important to note that an increase in promotion sensitivity due to goal proximity could be attenuated when (a) the promotional effort is identical to the main program effort (e.g., "fly next week and earn triple miles") *and* (b) effort acceleration is subject to behavioral ceiling effects (e.g., frequent flyers cannot accelerate their flights beyond a certain point).

The goal-gradient effect has important implications for price sensitivity and competition. It suggests that the own- and cross-price elasticities of the RP sponsor will be lower for members who are closer to the program's goals. Compared to non-members of the RP, members may be willing to pay a price premium or forgo convenience (e.g., purchase more expensive and/or layover flights), particularly when they approach a program goal. In addition, the increased motivation to achieve RP goals may reduce competition and price wars by escalating customer lock-in and switching costs and enhancing consumption rates (i.e., expanding the category).

Conclusion

Building on the behaviorist goal-gradient hypothesis, we proposed that people working toward future rewards would accelerate their effort as they near the reward threshold. Using a wide range of empirical and modeling approaches, the findings reported in this article provide converging evidence for the impact and importance of goal-gradients in the human psychology of rewards. Customers not only accelerate toward rewards (in terms of timing, quantity, and persistence of effort), but their acceleration also predicts loyalty and future engagement with similar goals. The goal-distance model unifies these results as well as anticipates the finding of the illusion of progress toward the goal. On the basis of this research, we propose that the goal-gradient and its modeling have important theoretical and practical implications for achievement motivation and goal behavior as well as for incentive systems and marketing promotions.

Table 1: The Goal-Distance Model with Unobserved Heterogeneity in Base Hazard Rates (Café Reward Program)

Latent Class-level parameters:	Class 1	Class 2	Class 3	Class 4	Class 5
Segment Size	30%	28%	7%	32%	3%
γ (hazard rate)	0.5^{***}	0.7^{***}	1.5^{***}	0.06^{***}	1.8^{**}
α (hazard rate)	1.4^{***}	2.3^{***}	3.7^{***}	1.15^{***}	1.9^{***}
β_0	0.4^{**}	-0.3^{**}	-0.7^{***}	1.5^{***}	-0.02
Parameters not varying by latent class					
Goal-distance parameters	*Estimate*				
Linear effect of goal-distance, $\hat{\beta}_1$	-0.03^{***}				
Quadratic effect of goal-distance, $\hat{\beta}_2$	-0.04^{***}				
Covariate parameters (i.e. control variables)					
Weekly number of stamps	-0.005				
End of semester	0.04^{***}				
Midterm break	-0.17^{***}				
Card type (American vs. Italian)	-0.10^{***}				
Day of week (M-Th)	-0.02				
Friday	-0.18^{***}				
Saturday/Sunday	-0.39^{***}				

****** p < .01; ** p < .05; * p < .1 based on Wald test (two-tailed)*

Table 2: Parameter Estimates with Unobserved Heterogeneity in Goal-Distance and Hazard Rate Parameters (Café Reward Program)[14]

Latent Class-level parameters:	Class 1	Class 2	Class 3	Class 4
Segment Size	58%	29%	7%	6%
γ (hazard rate)	0.6^{***}	0.07^{***}	1.5^{***}	0.9
α (hazard rate)	1.7^{***}	1.2^{***}	3.5^{***}	1.3^{***}
β_0	0.09	1.4^{***}	-0.7^{***}	0.5^{***}
Goal-distance parameters				
Linear effect of goal-distance, $\hat{\beta}_1$	-0.04^{**}	-0.04	-0.003	0.007
Quadratic effect of goal-distance, $\hat{\beta}_2$	-0.06^{***}	0.005	-0.10^{*}	-0.009

****** p < .01; ** p < .05; * p < .1 based on Wald test (two-tailed)*

[14] The model included the same control variables reported in Table 1. The covariate estimates were nearly identical and are not reported here.

Table 3: The Hazard Rate Goal-Distance Model of Inter-Visit Times (Music-Rating Program)

Latent Class-level parameters:	Class 1	Class 2
Segment Size	71%	29%
α (hazard rate)	0.5^{***}	2.1^{***}
β_0	1.4^{**}	$-.9^{***}$
Parameters not varying by latent class		
Acceleration Parameters	*Estimate*	
Linear effect of goal-distance, $\hat{\beta}_1$	-0.25^{***}	
Quadratic effect of goal-distance, $\hat{\beta}_2$	-0.22^{**}	
Covariate parameters (i.e. control variables)		
Daily number of visits	0.1	
Total visits to earn certificate	-0.01	
Saturday/Sunday	0.03	

**** p < .01; ** p < .05; * p < .1 based on Wald test (two-tailed)*

Table 4: The Tobit Goal-Distance Model of Quantity

Model parameters	***Estimate***
Intercept	29.9^{***}
Variance (σ^2)	253.4^{***}
Linear effect of goal-distance, $\hat{\beta}_1$	-13.0^{***}
Quadratic effect of goal-distance, $\hat{\beta}_2$	-7.5^{***}
Daily number of songs	0.8
Total visits to earn certificate	-10.9^{***}
Saturday/Sunday	-0.2
Number of certificates already earned during the visit	-4.0^{***}

**** p < .01; ** p < .05; * p < .1 based on Wald test (two-tailed)*

Table 5: The Logit Goal-Distance Model of Effort Persistence and Defection

Latent Class-level parameters:	Class 1	Class 2
Segment Size	94%	6%
Constant	-3.5^{***}	-4.5^{***}
Parameters not varying by latent class		
Acceleration Parameters	*Estimate*	
Linear effect of goal-distance, $\hat{\beta}_1$	0.49^{***}	
Quadratic effect of goal-distance, $\hat{\beta}_2$	0.12^{**}	
Effect of certificate completed ($d_{tq} = 1$), $\hat{\beta}_3$	0.27^{***}	
Covariate parameters (i.e. control variables)		
Daily number of songs	-0.05	
Total visits to earn certificate	0.40^{***}	
Saturday/Sunday	0.05	
Number of certificates already earned during the visit	0.17^{**}	

**** p < .01; ** p < .05; * p < .1 based on Wald test (two-tailed)*

Figure 1: Typical Findings from Hull's (1934) Experiments with Rats[15]

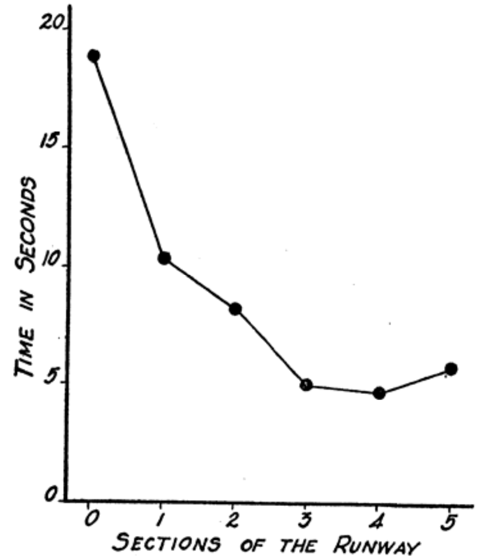


Figure 2: The Café Reward Program Cards

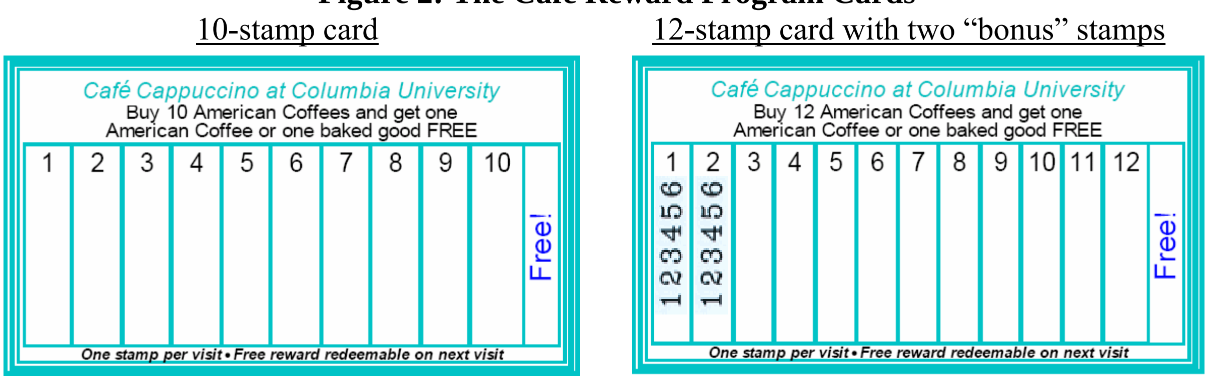
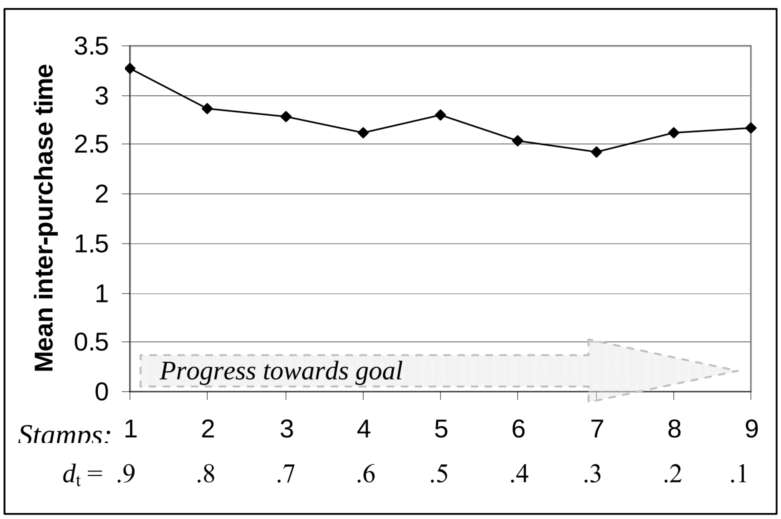


Figure 3: Purchase Acceleration as a Function of Smaller Goal-Distance



[15] Composite graph from 11 blind rats, showing the length of time consumed in traversing the several sections of a straight runway extending from the starting box, at the beginning of Segment 0, to the food, some 21 inches beyond Segment 5. The several points on the curve represent means from approximately 160 measurements.

Figure 4: Average Inter-purchase Times on First and Second Cards

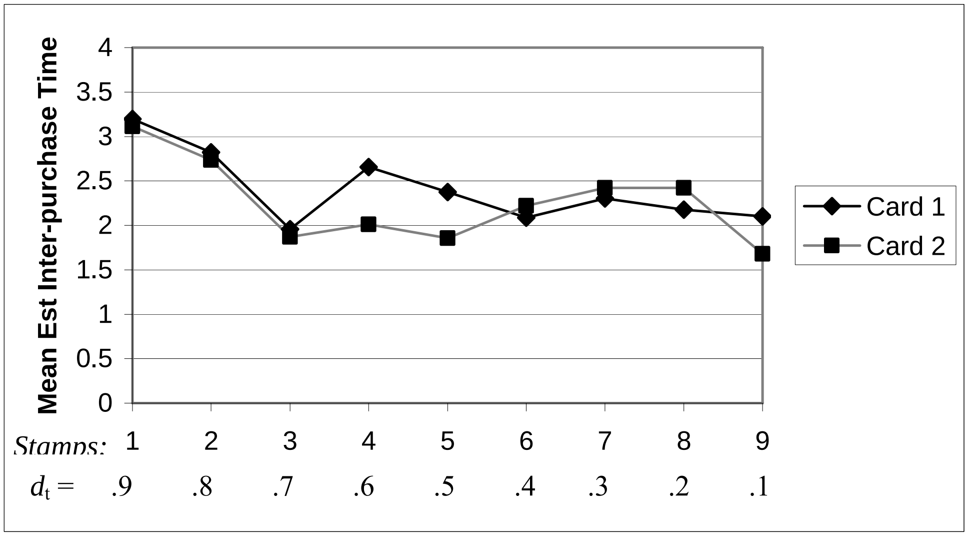


Figure 5: The Jaboom! Music-Rating Web-Interface

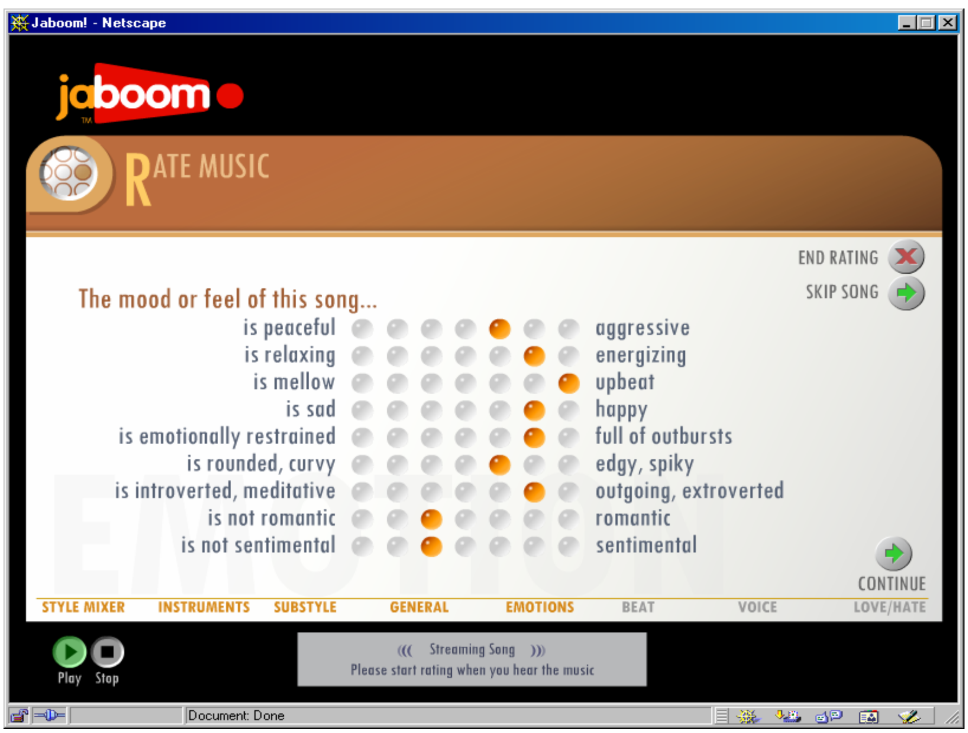


Figure 6: Number of Songs Rated as a Function of Goal Progress

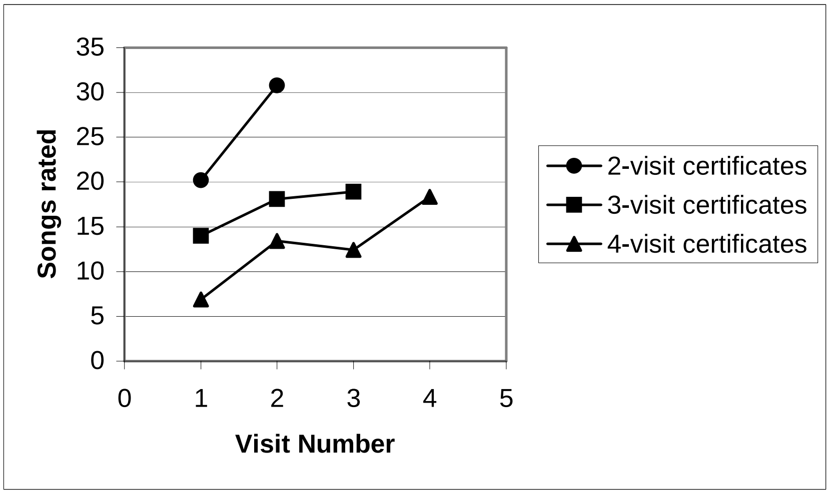


Figure 7: Number of Visit Terminations as a Function of Goal-Distance

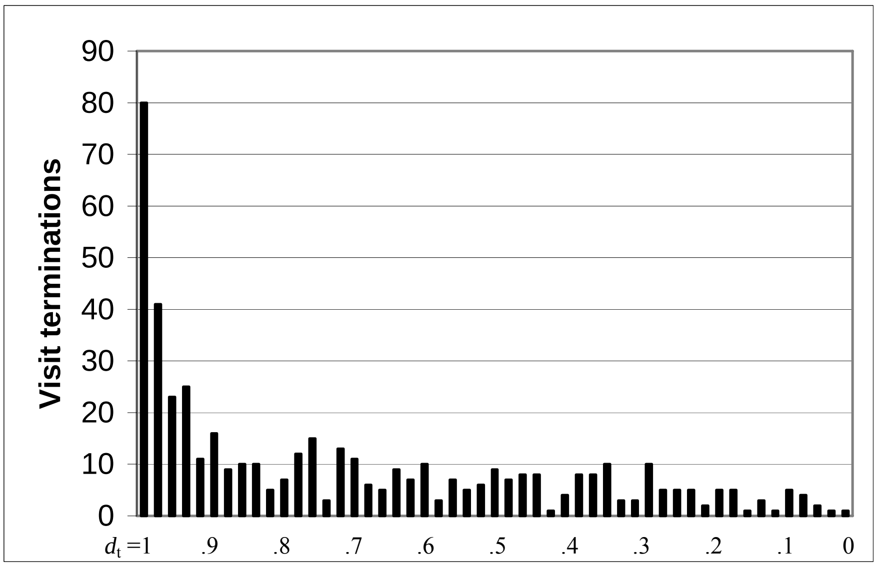
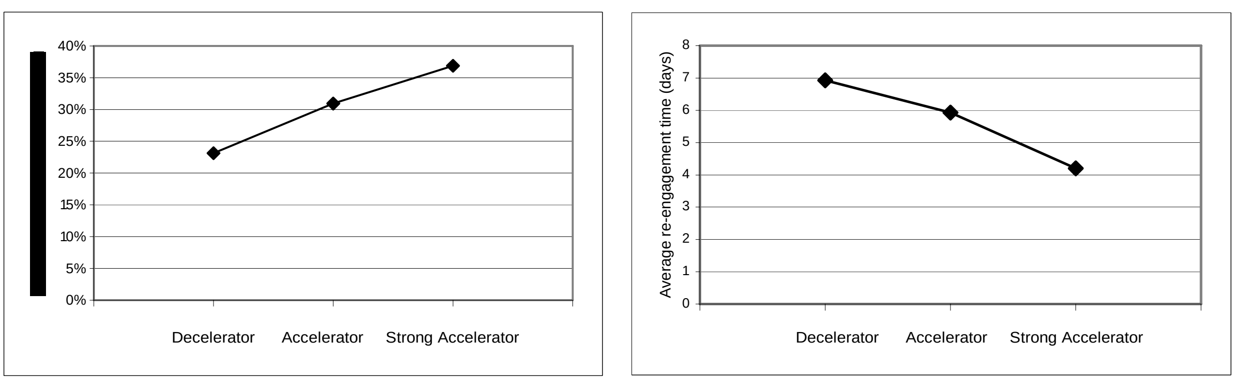


Figure 8: Effect of Card 1 Acceleration on Reengagement Probability and Timing



References

Anderson, Amos C. (1933), "Runway Time and the Goal Gradient," *Journal of Experimental Psychology*, 16, 423-428.

Atkinson, John W. (1957), "Motivational Determinants of Risk-Taking Behavior," *Psychological Review*, 64 (November), 359-372.

Bell, David R. and James M. Lattin (2000), "Looking for Loss Aversion in Scanner Panel Data: The Confounding Effect of Price Response Heterogeneity," *Marketing Science*, 19 (2), 185-200.

Breen, Richard (1996), *Regression Models: Censored, Sample-Selected or Truncated Data* (Sage University Paper Series on Quantitative Applications in the Social Sciences # 07-111). CA: Sage.

Brown, Judson S. (1948), "Gradients of Approach and Avoidance Responses and Their Relation to Level of Motivation," *Journal of Comparative and Physiological Psychology*, 41, 450-465.

Carver, Charles S. and Michael F. Scheier (1990), "Origins and Functions of Positive and Negative Affect: A Control-Process View," *Psychological Review*, 97, 19-35.

Cox, D. R. (1972), "Regression Models and Life Tables," *Journal of the Royal Statistical Society, Series B*, 34, 187-220.

Deighton, John A. (2000), "Frequency Programs in Service Industries," in Dawn Iacobucci and Teresa A. Swartz (eds.), Handbook of Services Marketing and Management, Thousand Oaks, California: Sage, 401-407.

Dowling, Grahame R. and Mark Uncles (1997), "Do Customer Loyalty Programs Really Work?" *Sloan Management Review*, 38 (4), 71-82.

Fisher, Ronald A. and Frank Yates (1957), *Statistical Tables for Biological, Agricultural and Medical Research* (5th ed.). New York: Hafner Publishing Co.

Forster, Jens, E. Tory Higgins, and Lorraine C. Idson (1998), "Approach and Avoidance Strength During Goal Attainment: Regulatory Focus and the "Goal Looms Larger" Effect," *Journal of Personality and Social Psychology*, 75 (5), 1115-1131.

Gupta, Sunil (1988), "Impact of Sales Promotions on When, What, and How Much to Buy," *Journal of Marketing Research*, 25, 342-355.

--- (1991), "Stochastic Models of Interpurchase Time with Time-Dependent Covariates," *Journal of Marketing Research*, 28, 1-15.

Hardie, Bruce, Eric J. Johnson, and Peter Fader (1993), "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice," *Marketing Science*, 12 (4), 378-394.

Heath, Chip, Richard P. Larrick, and George Wu (1999), "Goals as Reference Points," *Cognitive Psychology*, 38, 79-109.

Heilizer, Fred (1977), "A Review of Theory and Research on the Assumptions of Miller's Response Competitions Models: Response Gradients," *The Journal of General Psychology*, 97, 17-71.

Helsen, Kristiaan and David C. Schmittlein (1993), "Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models," *Marketing Science*, 12 (4), 395-414.

Herrnstein, Richard J. and Drazen Prelec (1991), "Melioration: A Theory of Distributed Choice," *Journal of Economic Perspectives*, 5 (Summer), 137-156.

Hsee, Christopher K., Frank Yu, and Joe Zhang (2003), "Medium Maximization," Journal of Consumer Research, 30 (1), 1-14.

Huber, Joel, John W. Payne, and Christopher P. Puto (1982), "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, 9 (June), 90-98.

Hull, Clark L. (1932), "The Goal Gradient Hypothesis and Maze Learning", *Psychological Review*, 39, 25-43.

--- (1934), "The Rats' Speed of Locomotion Gradient in the Approach to Food", *Journal of Comparative Psychology*, 17, 393-422.

Kamakura, W. and G. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26 (November), 379-390.

Kivetz, Ran (2000), "Preferences towards Streams of Efforts for Future Rewards: Understanding Frequency Programs," doctoral dissertation, Graduate School of Business, Stanford University.

--- (2003), "The Effects of Effort and Intrinsic Motivation on Risky Choice," *Marketing Science*, 22 (4), 477-502.

---, Oded Netzer, and V. Srinivasan (2004), “Alternative Models for Capturing the Compromise Effect,” *Journal of Marketing Research*, 41 (3), 237-257.

---, and Itamar Simonson (2003), ”The Idiosyncratic Fit Heuristic: Effort Advantage as a Determinant of Consumer Response to Loyalty Programs,” *Journal of Marketing Research*, 40 (4), 454-467.

Lal, Rajiv, and David E. Bell (2003), “The Impact of Frequent Shopper Programs in Grocery Retailing,” *Quantitative Marketing and Economics*, 1 (2), (June).

Lewin, Kurt (1951), *Field Theory in Social Science: Selected Theoretical Papers*, Dorwin Cartwright, ed. Oxford, England: Harpers.

Lewis, Michael (2004), “The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention,” *Journal of Marketing Research*, 41 (3), 281-292.

Miller, Neal E. (1944), “Experimental Studies of Conflict,” in *Personality and the Behavior Disorders* (Vol. 1), J. Hunt, ed. Oxford, England: Ronald Press, 431-465.

Payne, John R., James R Bettman, and Eric J. Johnson (1992), “Behavioral Decision Research: A Constructive Processing Perspective,” *Annual Review of Psychology*, 43, 87-131.

Prelec, Drazen and George Loewenstein (1998), “The Red and the Black: Mental Accounting of Savings and Debt,” *Marketing Science*, 17, 1.

Seetharaman, P. B. and Pradeep K. Chintagunta (2003), “The Proportional Hazard Model for Purchase Timing: A Comparison of Alternative Specifications,” *Journal of Business and Economic Statistics*, 21 (3), 368-382.

Sharp, Byron and Anne Sharp (1997), “Loyalty Programs and Their Impact on Repeat-Purchase Loyalty Patterns”, *International Journal of Research in Marketing*, 14, 473-486.

Simonson, Itamar (1990), “The Effect of Purchase Quantity and Timing on Variety-Seeking,” *Journal of Marketing Research*, 27 (2), 150-162.

--- and Amos Tversky (1992), “Choice in Context: Tradeoff Contrast and Extremeness Aversion,” *Journal of Marketing Research*, 29 (August), 281-295.

--- and Russell S. Winer (1992), “The Influence of Purchase Quantity and Display Format on Consumer Preference,” *Journal of Consumer Research*, 19 (1), 133-138.

Soman, Dilip and Mengze Shi (2003), “Virtual Progress:The Effect of Path Characteristics on Perceptions of Progress and Choice,” *Management Science*, 49 (9), 1129-1151.

Stevens, S. (1957), “On the Psychophysical Law,” *Psychological Review*, 64, 153-181.

Thaler, Richard (1980), “Toward a Positive Theory of Consumer Choice,” *Journal of Economic Behavior and Organization*, 1 (March), 39-60.

Tobin, James (1958), “Liquidity Preference as Behavior Towards Risk,” *The Review of Economic Studies*, 25, 65-86.

Trope, Yaacov and Nira Liberman (2003), “Temporal Construal,” *Psychological Review*, 110 (3), 403-421.

Van Heerde, Harald J., Peter S. H. Leeflang, and Dick R. Wittink (2000), “The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data,” *Journal of Marketing Research*, 37 (3), 383-395.

van Osselaer, Stijn M. J., Joseph W. Alba, and Puneet Manchanda (2003), “Irrelevant Information and Mediated Intertemporal Choice,” *Journal of Consumer Psychology*, 14(2).

Wertenbroch, Klaus (1998), “Consumption Self-Control by Rationing Purchase Quantities of Virtue and Vice,” *Marketing Science*, 17 (4), 317-337.

Winer, Russell S. (1986), “A Reference Price Model of Demand for Frequently-Purchased Products,” *Journal of Consumer Research*, 13 (September), 250-256.

--- (1999), “Experimentation in the 21st Century: The Importance of External Validity,” *Journal of the Academy of Marketing Sciences*, 27 (3), 349-358.

Wittink, Dick R. (2004) “Journal of Marketing Research: 2 Ps,” *Journal of Marketing Research*, 41 (1), 1-6.