What You *Don't* Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions

Roland T. Rust • J. Jeffrey Inman • Jianmin Jia • Anthony Zahorik

Owen Graduate School of Management, Vanderbilt University, Nashville, Tennessee 37203, roland.rust@owen.vanderbilt.edu

School of Business, University of Wisconsin, Madison, Wisconsin 53706, jinman@bus.wisc.edu

Department of Marketing, Chinese University of Hong Kong, NT, Hong Kong, jia@baf.msmail.cuhk.edu.hk

ACNielsen Burke Institute, Nashville, Tennessee 37212, zahoriktny@aol.com

Abstract

We show that some of the most common beliefs about customer-perceived quality are wrong. For example, 1) it is not necessary to exceed customer expectations to increase preference, 2) receiving an expected level of bad service does not reduce preference, 3) rational customers may rationally choose an option with lower expected quality, even if all non-quality attributes are equal, and 4) paying more attention to loyal, experienced customers can sometimes be counterproductive. These surprising findings make sense in retrospect, once customer expectations are viewed as distributions, rather than simple point expectations. That is, each customer has a probability density function that describes the relative likelihood that a particular quality outcome will be experienced. Customers form these expectation distributions based on their cumulative experience with the good or service. A customer's cumulative expectation distribution may be conceptualized as being a predictive density for the next transaction.

When combined with a diminishing returns (i.e., concave) utility function, this Bayesian theoretical framework results in predictions of: (a) how consumers will behave over time, and (b) how their perceptions and evaluations will change. In managerial terms, we conclude that customers consider not only expected quality, but also risk. This may help explain why current measures of customer satisfaction (which is highly related to expected quality) only partially predict future behavior. We find that most of the predictions of our theoretical model are borne out by empirical evidence from two experiments. Thus, we conclude that our approach provides a useful simplification of reality that successfully predicts many aspects of the dynamics of consumer response to quality.

These findings are relevant to both academics and managers. Academics in the area of customer satisfaction and service quality need to be aware that it may be insufficient to measure only the point expectation, as has always been the standard practice. Instead it may be necessary to measure the uncertainty that the customer has with respect to the level of service that will be received. Due to questionnaire length constraints, it may not be practical for managers to include uncertainty questions on customer satisfaction surveys. Nevertheless it is possible to build a proxy for uncertainty by measuring the extent of experience with the service/good, and this proxy can be used to partially control for uncertainty effects.

The findings of the study were obtained using 1) an analytical model of customer expectation updating, based on a set of assumptions that are well-supported in the academic literature, and 2) two behavioral experiments using human subjects: a cross-sectional experiment, and a longitudinal experiment. Both the analytical model and the behavioral experiments were designed to investigate the effects that *distributions* of expectations might have, and especially the effects that might deviate from the predictions that would arise from a traditional point expectation model. The behavioral experiments largely confirmed the predictions of the analytical model. As it turned out, the analytical model correctly (in most cases) predicted behavioral effects that contradict some of the best-accepted "truisms" of customer satisfaction.

It is now clear that a more sophisticated view of customer expectations is required—one that considers not only the point expectation but also the likelihood across the entire distribution of possible outcomes. This distinction is not "just academic," because it results in predictable behavior that deviates significantly from that which was traditionally expected based on simpler models.

(*Quality*; *Customer Satisfaction Measurement*; *Customer Expectations*; *Customer Retention*; *Bayesian Updating*; *Customer Lifetime Value*)

0732-2399/99/1801/0077$05.00

1. Introduction

The trade literature in quality and customer satisfaction abounds with rarely questioned platitudes. Some of the most often repeated and/or noncontroversial are:

- *"It is necessary to exceed customer expectations."*
- *"If a customer expects a bad level of quality and receives it, he/she will reduce his/her level of preference for the brand."*
- *"Given two equally-priced options, the customer will choose the one with the higher expected quality."*
- *"Management should always focus on its most loyal customers."*

We will show, using both an analytical model and behavioral experiments, that all of these truisms are flawed. Management must instead adopt a more sophisticated view of how customer expectations are updated, and how customer expectation updating relates to preference and future choice behavior.

1.1. What's Wrong with Customer Satisfaction Measures?

Managers routinely conduct customer satisfaction surveys and use that information to produce explanatory and predictive models of customer repurchase intention, word-of-mouth intention, customer repurchase behavior, and market share (Bolton and Drew 1991a, 1991b; Boulding et al. 1993; Cronin and Taylor 1992, 1994; Fornell 1992; Oliver and DeSarbo 1988; Parasuraman et al. 1985; Rust and Zahorik 1993; Teas 1993).[1] Such models may be used to evaluate the projected profit impact of service improvement programs (Rust et al. 1995). Recently, some authors have questioned the predictive ability of customer satisfaction measures (Gale 1997, Reichheld 1996). They have complained that many customers who report being "very satisfied" or perceive quality to be "excellent" nevertheless subsequently switch to a competitor. In general, while customer satisfaction measures and/or perceived quality measures are important predictors of intentions and behavior, they often explain only 30–50% of the variance.

We argue that one of the reasons for the apparently weak satisfaction-behavior link may be that customers respond according to the perceived variance in service, in addition to expected quality. In other words, customers' *certainty* about quality has an important impact. This perceived risk can be related to the variance of a Bayesian predictive distribution of quality outcomes, resulting in a more multidimensional view of customer expectations. This viewpoint results in some apparently counter-intuitive insights that have some important managerial implications.

1.2. Dynamics and the Decision Theory Perspective

Because customer relationships unfold over time, it is important to clearly understand the dynamics of how quality perceptions are formed and updated as well as how such perceptions influence customer retention over time. With few exceptions (Anderson and Sullivan 1993, Bolton and Drew 1991a, Boulding et al. 1993), the research literature has not produced dynamic models of how these changes occur. Researchers have begun to argue that a decision theory framework offers a powerful way of conceptualizing the dynamics of quality and customer retention (Anderson and Sullivan 1993, Boulding et al. 1993). Bayesian decision theory (Berger 1985, Zellner 1971) provides a way of understanding customer retention that has thus far been underutilized. This literature provides well-explored methods of describing how people incorporate new information and form new expectations over time. Thus, it can be used to generate testable predictions of behavior over time (Kahneman and Tversky 1972).

Essentially there are two major issues in the field of behavioral decision research-risky choice behavior and probability judgments. Expected utility can be determined by mean and variance if the distribution is normal or if the utility function is quadratic (Markowitz 1959, 1987; Meyer 1987). Behavioral research on risky choice has revealed discrepancies between actual choice behavior and the prescriptions of the classic von Neumann and Morgenstern (1944) utility axioms. For example, Payne (1973) argued that factors other than mean and variance (or even higher moments) are needed to capture risk perceptions and risky choice behavior. Empirical studies have challenged the traditional utility theory on several grounds (e.g.,

[1]Strictly speaking, many customer satisfaction surveys actually measure perceived quality.

Kahneman and Tversky 1979; Bell 1982, 1985; Machina 1987; Inman et al. 1997). This has resulted in a much deeper understanding of risky choice behavior, while maintaining much of the mathematical formalism of the normative approach (see Fishburn (1988) for a comprehensive review of generalizations of expected utility theory).

Behavioral research on probability judgments has produced mixed results regarding the descriptive power of Bayesian decision theory. In studies of probability judgments, prior probability (or base-rate) information is often ignored by subjects (Kahneman and Tversky 1973), but sometimes it is utilized appropriately (Gigerenzer et al. 1988). Moreover, people have a tendency to seek confirmatory evidence in aggregating various pieces of information, which may lead to the failure to change one's opinion in the face of nonsupporting evidence (Wason 1960, Doherty et al. 1979). Another relevant issue is the underestimation of posterior probability, called "conservatism" (Edwards 1968, Slovic and Lichtenstein 1971). That is, upon receipt of new information, subjects revise their posterior probability estimates in the same direction as the Bayesian model, but revisions are typically much less extreme than those calculated from the Bayesian perspective. In general, probabilistic judgments involve contingent processing and heuristics, which sometimes yield reasonable results and sometimes lead to systematic biases (Tversky and Kahneman 1974, Fischhoff and Beyth-Marom 1983). People often utilize a variety of rules for making probabilistic inference in different environments, including both statistical and nonstatistical methods (Ginossar and Trope 1987).

Although previous research has generated several empirical findings of risky choice behavior and probability judgments, there is surprisingly little empirical evidence regarding a dynamic decision model that uses Bayesian updating for revising expectation distributions upon receipt of new information. In particular, previous research on probability judgments based on Bayesian updating has used discrete probability information and/or nonbusiness contexts. In contrast, our study focuses on the descriptive validity of a Bayesian decision model using continuous probability information in the area of quality perceptions and customer retention.

1.3. Comparison with Previous Models

Beginning with SERVQUAL (Parasuraman et al. 1988) service quality measurement instruments have universally focused on a point estimate of quality. The implicit assumption has always been that a customer's "perceived quality" or "expected quality" was a single point. This viewpoint has also been shared by the recent dynamic models of customer satisfaction. Such models involve the updating of expectations and quality perceptions. For example, Boulding et al. (1993) employ a linear updating scheme by which expectations and cumulative quality perceptions are updated according to the most recent transaction quality perception. Their model is analogous to an exponential smoothing model. Likewise, Bolton and Drew (1991a,b) employ linear updating functions. Only Anderson and Sullivan (1993) suggest (but do not develop) an updating approach that involves a distribution rather than only a point estimate.

Our model is different in that it involves a fully developed Bayesian updating scheme, in conjunction with a diminishing returns utility function. We will see that the behavioral predictions arising from our model are quite different in some cases from the predictions arising from linear updating models. Our insights arise from the observation that experience with a brand makes knowledge of the brand's quality more complete, thereby reducing the customer's risk. Because the customer's utility function is concave (i.e., diminishing returns), reduced risk is always good and increases the customer's preference for the brand. This results in some seemingly counter-intuitive results that are nevertheless quite logical under closer inspection. Importantly, these results have key consequences for managerial practice.

1.4. Plan of the Paper

In the next section we develop a theoretical framework that can be used to describe the relationship between customer retention and quality perceptions over time, based on the concepts of expected utility maximization and Bayesian updating. This framework results in several propositions about how customers should behave over time. We then present two behavioral experiments to test whether or not the theoretical propositions hold up on actual behavior. Finally, we discuss the implications of our results, as well as limitations and directions for future research.

2. Model Development

In this section we present the relationship between quality and customer retention as a model in which retention is based on the distribution of customer expectations. Importantly, we allow the distribution to be updated over time according to the quality perceived in a particular transaction. While previous research (e.g., Anderson and Sullivan 1993, Boulding et al. 1993) has pointed in this direction, this is the first attempt to formulate these phenomena in a fully Bayesian framework, complete with operationalized prior and posterior distributions of quality. We begin by presenting the assumptions on which our theoretical model is based. We then present the theoretical formulation in formal detail. Finally, we provide a list of nontrivial testable propositions that arise from the model.

2.1. Assumptions

We assume a scenario in which an individual with existing expectations chooses a particular option, experiences an outcome, and then changes his/her expectations based on the outcome.

Assumption 1. *Utility as a function of perceived quality is continuous, twice differentiable, increasing and concave. This implies that customers suffer more from not having their expectations met than they benefit from an equivalent positive disconfirmation.*

This assumption is borne out by prior empirical research in customer satisfaction and service quality (Anderson and Sullivan 1993, DeSarbo et al. 1994, Inman et al. 1997, Rust et al. 1995), and is also consistent with assumptions traditionally made in decision theory (Keeney and Raiffa 1976). This particular shape of the utility function, well-documented in a variety of academic and proprietary studies, combines with the Bayesian framework to produce interesting and managerially relevant, testable conclusions.

Assumption 2. *The customer has a predictive distribution of outcomes that reflects the relative likelihood that each outcome will occur.*

This assumption says that the customer considers not only the outcome s/he expects on average (i.e., the point expectation), but also the entire distribution of possible outcomes.

Assumption 3. *In any purchase situation, the customer's preference for an option increases with its expected utility, and probability of choosing an option increases with preference for that option.*

The assumption is consistent with the economic theory of utility and the consumer viewpoint (Luce 1959, Manrai 1995, McFadden 1976) with randomness arising from the variability of the predictive distribution.

Assumption 4. *The customer has a prior distribution of average quality for the product or service.*

Through experience, a customer learns what quality can be expected on average. This prior distribution may be formed on the basis of experience with other brands, on the basis of prior experience with the brand under consideration, or both. Similar assumptions are commonly made in a variety of Bayesian applications (e.g., Little 1966).

Assumption 5. *The customer updates the prior distribution based on the perceived quality of the transaction using a Bayesian updating process.*

Such an updating mechanism has been suggested by previous research in the service quality area (Anderson and Sullivan 1993, Boulding et al. 1993, Kopalle and Lehmann 1995).[2]

Assumption 6. *Perceived quality varies randomly across transactions.*

Even in a highly controlled environment such as a manufacturing assembly line, there are numerous uncontrollable factors that can produce variability in the quality of the output (Burr 1976). It is notable that in service, which is a growing majority of every developed economy, variability is typically much greater

[2]We tested Assumption 5 in our first experiment by estimating the relative weight subjects placed on prior perceptions versus the weight placed on new information in updating the performance expectation. The weight on new information was significantly positive. Thus, consistent with Assumption 5, subjects used the information to update their perceptions of the expected performance.

than in manufacturing. While an entire field of study, statistical process control (Juran and Gryna 1980), has arisen to identify and attempt to control the sources of variation in quality, some variability inevitably remains even in a process that is "under control" (Deming 1986). Another source of variation is perception error on the part of the customer (Thurstone 1927). We recognize that prior research indicates that variables such as prior expectations may influence the quality perception itself (e.g. Boulding et al. 1998, Oliver 1997), but to keep this initial model simple enough to provide maximum insight, we suppress secondary effects here.

2.2. Mathematical Formulation

We now build a formal mathematical structure that incorporates the preceding assumptions. Let the customer have a known prior distribution $\pi(Q)$ of the average quality (Q) of the brand. Let $\pi(Q)$ have a normal distribution with mean μ and variance $\tau^2 > 0$, reflecting the customer's degree of uncertainty. Further, for a particular transaction, we will denote the customer's perceived quality as X. We assume that X is distributed normally, with mean Q and variance $\sigma^2 > 0$ representing random variation arising from both variability of quality and errors in perception. It is clear from standard Bayesian analysis (Berger 1985, p. 127) that the joint density of Q and X is:

$$h(x,Q) = (2\pi\sigma\tau)^{-1} \exp\{-1/2\ [((Q - \mu)^2/\tau^2) + ((x - Q)^2/\sigma^2)]\} \tag{1}$$

from which we can get the predictive (marginal distribution) of X:

$$p(x) = \int_{-\infty}^{\infty} h(x,Q)dQ = (2\pi\rho)^{-1/2}\ (\sigma\tau)^{-1}\ \exp\{-(\mu - x)^2 /[2(\sigma^2 + \tau^2)]\} \tag{2}$$

where $\rho = (\sigma^2 + \tau^2)\ /\ (\sigma^2\tau^2)$. If we scale the units (without loss of generality) such that $\sigma^2 + \tau^2 = 1$, then the predictive distribution becomes

$$p(x) = (2\pi)^{-1/2} \exp\{-(\mu - x)^2/2\} \tag{3}$$

which is immediately seen as being normal with mean μ and variance one.

We now address what happens when a level of quality, x_t, is observed on the next transaction. We also define the *disconfirmation* to be $\Delta_t = x_t - \mu$, following the traditional definition (Oliver 1980). (It is worth noting that the disconfirmation, although conceptualized as a difference, is not usually calculated mathematically, but rather observed, and measured, directly.) Again from standard Bayesian updating, the posterior distribution of Q, $\pi(Q \mid x_t)$, is calculated as:

$$\pi(Q \mid x_t) = h(x_t,Q)/p(x_t) = (\rho/2\pi)^{1/2} \exp\{-(\rho/2)[Q - (1/\rho) ((\mu/\tau^2) + (x_t/\sigma^2))]^2\} \tag{4}$$

which is a normal distribution with mean $\mu + \Delta\tau^2$ and variance $\rho^{-1} = \sigma^2\tau^2$. Thus, the posterior mean increases (decreases) when the disconfirmation, Δ, is positive (negative). Also the customer's uncertainty decreases, regardless of the outcome, since $\rho^{-1} < \tau^2$. The predictive density of the quality of the next transaction, x_{t+1}, given observed quality level x_t, is

$$p(x_{t+1} \mid x_t) = \int_{-\infty}^{\infty} f(x_{t+1} \mid Q)\pi(Q \mid x_t)dQ \tag{5}$$

where $f(x_{t+1} \mid Q)$ is normal with mean μ and variance σ^2. This predictive density is easily shown to be normal with mean $\mu + \Delta\tau^2$ and variance $\sigma^2 + \sigma^2\tau^2$. The predictive density's new mean is bigger (smaller) if the disconfirmation, Δ, is positive (negative), and its variance is smaller regardless (since $\sigma^2 + \sigma^2\tau^2 \leq 1$), reflecting the greater certainty created by experience.

From (4) we can see that the mean of the posterior distribution exceeds that of the prior distribution whenever the disconfirmation, Δ, is positive and is lower when Δ is negative. The same holds for the predictive distributions. In other words, the expectations (and future predictions) move in the direction of the perceived level of quality. Note also from (4) that the variance is reduced, as makes sense from having more experience. From Assumption 1, we assume a continuous, twice differentiable utility function $U(x)$, with $U' > 0$ and $U'' < 0$. From Assumption 3, we assume an expected utility V, that is equal to $\int U(x)p(x)dx$, where $p(x)$ is the predictive distribution of x. The expected utility can be expressed in terms of mean and standard deviation since the predictive distribution is normal.

For example, if we assume an exponential utility function, such as is popularly used in modeling risky choice problems, then the expected utility can be determined by a tradeoff between expected value and risk (Jia and Dyer 1996):

$$V_i = \alpha_{1i} + \alpha_{2i}\mu_i - \alpha_{3i}\sigma_i^2 \tag{6}$$

where V_i is the preference measure of brand i, μ_i is the expected performance of brand i, σ_i^2 is the variance measuring perceived risk or uncertainty of brand i's performance, and α_{1i} and α_{21}, $\alpha_{3i} > 0$ are constants. This is a positive linear transformation of the certainty equivalent form of the exponential expected utility. The ratio α_{2i}/α_{3i} measures customers' risk tolerance, reflecting a tradeoff between the mean and variance in determining preference.

Based on Assumption 3, the consumer chooses the brand with the highest expected utility. Thus, a multinomial logit model (Luce 1959, McFadden 1974) can be used to estimate the choice likelihood for each brand:

$$p_i = \exp(V_i) \Big/ \sum_j \exp(V_j). \tag{7}$$

2.3. Propositions Arising from the Model

From the model above, a number of empirically testable propositions arise that warrant further scrutiny. (Proofs of all of the propositions are available in an Appendix that is available from the authors on request.) Following are some of the propositions resulting from the model that would seem to have a reasonable chance of being empirically falsified:

Proposition 1. *If an option is chosen and a better than expected outcome is observed, then the probability of choosing that option will increase.*

Intuitively, this proposition follows from the theoretical formulation whereby the predictive distribution's mean increases (and the predictive distribution shrinks). Preference improves because quality now seems better on average, and also less risky.

Proposition 2A. *If the preferred option is chosen and an expected outcome is observed, then the probability of choosing that option will increase.*

The logic here is that if the utility function is concave, reducing the variance of the predictive distribution will increase the expected utility. Preference improves not because the option seems better, but rather because it is less risky.

Proposition 2B. *If a nonpreferred option is chosen (as is possible in our probabilistic choice model) and an expected outcome is observed, then the probability of choosing that option will increase.*

Again, the shrinking predictive distribution increases the expected utility. We separate Propositions 2A and 2B because 2B will not obviously hold even if Proposition 2A is supported. Shrinking variance makes it more certain that the nonpreferred option will be worse, which would seem to suggest a plausible non-Bayesian basis for which this proposition may not hold. As before, preference for the option improves because the option seems less risky.

Proposition 3. *A rational consumer may choose an equally-priced option for which the expected quality is worse.*

The intuition behind this proposition is risk aversion. If one option has a larger variance for its predictive distribution, the uncertainty regarding this performance may induce many consumers to choose a lower expected performance/less variance option. Due to this, a less risky option can be preferable to a more risky option with a higher mean.

Proposition 4. *A worse than expected quality outcome may still increase the probability of choice for that option.*

The intuition for Proposition 4 is that there are two countervailing effects at work. The worse than expected outcome (all other things being equal) will tend to lower the predictive distribution and the expected utility. However, at the same time the variance of the predictive distribution is being reduced which will increase expected utility, given a concave utility function. For a small enough negative disconfirmation, the positive effect of the variance reduction will outweigh the negative effect of the lowered expectation, leading to an increase in preference.

Proposition 5. *A negative disconfirmation will evoke a greater relative change in preference relative to the case of* *zero disconfirmation than will a positive disconfirmation of equal magnitude.*

As mentioned previously, there will be two effects on preference: a variance reduction effect and a disconfirmation effect. Experience always reduces variance, resulting (all other things being equal) in increased preference. Positive disconfirmation increases preference, and negative disconfirmation (all other things being equal) decreases preference. Because of the concavity of the utility function, increasing absolute disconfirmation will create a disparity between the absolute effect of a negative disconfirmation and that of a positive disconfirmation. Proposition 5 refers to the case of zero disconfirmation as the basis, which eliminates the variance reduction effect (i.e., all cases have the same reduced variance). Since there is only a disconfirmation effect at work, this results in a simple pattern of asymmetric changes in preference.

PROPOSITION 6: *Given diffuse priors, and an equal historically-observed mean and variance, a sufficiently large negative disconfirmation will cause a greater preference shift in a less experienced customer. (Also a positive disconfirmation will cause a greater preference shift in a less experienced customer.)*

This effect occurs because more experience produces less uncertainty, but at a decreasing rate. In other words, experience has diminishing returns. Therefore the value gained from experience is large at first, resulting in large changes in preference, but diminishes, eventually resulting in small changes in preference. This is consistent with other research that shows that preference shifts less for more experienced customers (Bolton 1998, Boulding et al. 1993, 1998).

While the theoretical decision model forms a mathematically rigorous way of thinking about quality perceptions and customer retention, it is still only a mathematical abstraction. To determine the usefulness of this approach, we must test whether customers' behavior over time corresponds with the model's predictions. We now present results of two experiments in which the propositions were subjected to empirical test. We then discuss the implications of these results and directions for future research.

3. Longitudinal Experiment

3.1. Overview

One hundred and sixty undergraduate students at two large universities participated in a computerized decision-making exercise in return for extra credit. The experiment was designed to measure the effect of expectation distributions on discrete choice and on choice probability. The exercise consisted of the construction of a history of experiences among three brands of camera batteries. Disconfirmation was manipulated, while probability of choice, performance expectation, and perceived variance in performance for each brand of battery were measured at several points in the experiment.

The independent variables are 1) the disconfirmation ($\Delta = x_t - \mu$) between the actual and expected performance of the chosen battery, 2) the expected performance (μ) of each battery, and 3) the perceived variance (σ^2) of each battery. Disconfirmation was manipulated, while expected performance and perceived variance were measured. Expected performance for each battery was measured by the question "I'd expect the next battery (A/B/C) to last hours," while perceived variance for each battery was captured via subjects' response to the question "About 95% of the batteries for Brand (A/B/C) last between and hours."

Five different levels of disconfirmation were used: ten hours above expected, zero disconfirmation, and one, three, and ten hours below expected. The computer exercise interactively managed the amount of disconfirmation so that each subject was exposed to two of the five different levels of disconfirmation in experiences 4 and 7. For example, if the subject was to be exposed to the three hour disconfirmation treatment, the computer provided outcome feedback for the purchase by subtracting three hours from the subject's expectation for that brand.[3]

3.2. Method

We used an unbalanced design with more subjects in the zero and small disconfirmation (negative one and three hours) conditions to increase statistical power for those disconfirmation conditions. Sample sizes are

[3]The order of exposure to the amount of disconfirmation was counterbalanced.

shown in Table 2. Further, the brand labels were counterbalanced across subjects to control for order effects. Each subject was asked to imagine that s/he had recently purchased a brand of camera which used a special type of battery. A special type of battery was used in order to mitigate effects of category familiarity (i.e., so that each subject began the procedure with a diffuse expectation). The subject was told that s/he had sampled each brand of battery three times and was accordingly shown the first series of three experiences (see Table 1, columns 1–3). The subject was asked to provide his/her choice probability for each battery and to give his/her performance expectation and perceived variance for each battery. At the beginning of the fourth experience, the subject was told that s/he had purchased one of the three brands[4] (randomized across subjects) and was exposed to the first disconfirmation condition. Choice probability was then remeasured. The outcomes for the other two brands were then given and expected performance and perceived variance in performance were measured.

Following the fourth experience, subjects completed a brief distractor task and were then given outcome feedback on two more purchases of each brand (see Table 1, columns 5 and 6). Performance expectations and perceived variance were remeasured following the sixth experience. At the beginning of the seventh experience, the subject was again told that s/he had purchased one of the three brands (different from the brand in the fourth experience) and was given outcome feedback for this brand. Choice probability was remeasured. The subject was then given outcome feedback on the other two brands to complete the seventh experience. Following a second distractor task, subjects were provided outcome feedback for three more purchases of each battery, yielding a final history of ten purchases of each battery. Finally, expected performance and perceived variance were measured and the subject was asked which battery she would choose.

[4]Subjects were told to imagine that they had purchased the brand, "perhaps because it was on sale or because you had a coupon for it."

Table 1 Longitudinal Experiment—Performance Feedback Across Brands

Brands	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	Exp 9	Exp 10
1	72	59	67	E_{13}–D_1	74	57	E_{16}–D_2	72	56	69
2	63	69	60	E_{23}–D_1	59	68	E_{26}–D_2	67	64	65
3	66	60	64	E_{33}–D_1	65	67	E_{36}–D_2	59	67	63

E_{ir} = Subject's performance expectation for Brand *i* after round *r*.
D_j = Disconfirmation level for condition *j*.

3.3. Results

Table 2 shows the means for each experimental condition. Most of our propositions regard shifts in choice probability at specific levels of disconfirmation. Proposition 1 predicts that the probability of choosing an option will increase when the observed outcome is greater than expected. This proposition is supported, as probability of choice increased by an average of 8.7 points for the chosen option when the observed outcome exceeded the expected performance ($t_{32} = 4.76$, $p < 0.01$).[5] As predicted, when consumers experienced positive disconfirmation, their probability of choosing the option increased.

Proposition 2A predicts that if an expected outcome is observed (i.e., zero disconfirmation) for the most

[5]Although most of our propositions are directional, to be conservative we use a two-tailed test in our analysis.

Table 2 Longitudinal Experiment—Means Across Conditions

		Shift in Choice Probability		
Disconfirmation Level	N	Overall Shift	Low Experience	High Experience
+10 Hours	33	8.70**	14.31	3.41
0 Hours	112	5.21**	5.93	4.44
Preferred option	40	2.88*		
Nonpreferred option	72	6.51**		
−1 Hours	70	0.63	1.44	−0.52
−3 Hours	72	−0.42	0.76	−1.41
−10 Hours	33	−12.09**	−13.81	−10.47

**Shift statistically different than zero ($p < 0.01$).
*Shift statistically different that zero ($p < 0.05$).

preferred option, the probability of choosing the option will increase. Proposition 2B makes a similar prediction, but for a nonpreferred option. Both propositions are supported. The probability of choosing the option increased when the option was the most preferred option, shifting almost three points ($t_{39} = 1.73$, $p < 0.10$), while the probability of choosing the option when it was not the most preferred option increased over six points ($t_{71} = 5.26$, $p < 0.01$).

Proposition 3 predicts that subjects will not necessarily choose the brand with the highest expected performance. We tested Proposition 3 by asking subjects to choose a brand after observing the ten experiences with each brand. Using the multinomial logit model in (6) and (7), as one would anticipate, the expected performance exerts a significant effect on choice ($\alpha_2 = 0.235$, $p < 0.01$). Importantly, consistent with Proposition 3, the perceived variance also exerts a significant impact on subjects' choice ($\alpha_3 = -0.383$, $p < 0.01$). Thus, a brand's choice probability increases as the expected performance increases and decreases as the perceived variance increases.

Proposition 4 predicts that an outcome that is worse than expected may still increase the probability of choosing the option. We test this proposition by examining the effect on choice probability of disconfirmation levels of one, three, and ten hours less than expected. This proposition is not supported. The average probability of choosing the option increased slightly (i.e., in the expected direction) in the negative one hour disconfirmation condition, but the increase is not statistically significant ($t_{69} = 0.50$, NS). In the negative three hour disconfirmation condition, the average probability of choosing the option decreased slightly, but again this decrease is not statistically significant ($t_{71} = -0.34$, NS).

Proposition 5 predicts that a negative disconfirmation will have a greater impact on preference than a positive disconfirmation of equal magnitude. We test this by contrasting the -10 and $+10$ hour disconfirmation groups. Let X_1 be the difference between $+10$ and zero disconfirmation effects and X_2 be the difference between zero and -10 disconfirmation effects. The variance of X_1 and X_2 is the sum of the variance of the $+10$ and zero and -10 and zero disconfirmation effects, respectively, and the variance of $X_1 - X_2$ is the sum of the variances of X_1 and X_2. The t-test of the difference between X_1 and X_2 is statistically significant ($t_{65} = 4.69$, $p < 0.01$), supporting Proposition 5.

Since Proposition 6 states that an equivalent disconfirmation level results in a smaller shift in preference for more experienced customers, we conducted an ANOVA with experience level (high/low), disconfirmation, and their interaction as independent variables and the shift in preference as the dependent variable.[6] As predicted by Proposition 6, the effect of experience was significant ($F_{1,310} = 3.99$, $p < 0.05$), with smaller shifts at the higher experience level. Specifically, choice probability shifted by 2.7 points when experience was relatively low (four experiences) and shifted only 0.3 points when experience was higher (seven experiences). Not surprisingly, the shift in choice probability was an increasing function of the amount of disconfirmation ($F_{4,310} = 22.09$, $p < 0.01$). Further, the interaction between experience level and disconfirmation was marginally significant ($F_{4,310} = 2.02$, $p < 0.10$).

In sum, we find support for Propositions 1, 2A, 2B, 3, 5, and 6. Subjects' probability of choosing an option increased if performance for that option met (Propositions 2A and 2B) or exceeded (Proposition 1) the expected outcome. Per Proposition 3, subjects did not necessarily choose the brand with the greatest expected performance. Rather, they balanced the brand's expected performance against its variability in performance. Further, while subjects' probability of choosing an outcome was not adversely affected by an outcome that was slightly below their prior expectations, their probability of choice did not increase (significantly) as predicted by Proposition 4. The preference shifts created by negative vs. positive disconfirmations were greater in the negative disconfirmation direction, as predicted in Proposition 5. Proposition 6 was supported as well—at a given disconfirmation level, more experienced subjects tended to update their expectation to a lower degree than less experienced subjects.

While these results are encouraging, our first experiment is flawed in two respects. First, the longitudinal

[6]An analysis using prior preference as a covariate and preference as the dependent variable produced almost identical results. For ease of exposition, we discuss the analysis with shift in preference as the dependent variable.

nature of the design may have increased the chance that subjects saw that two factors were being manipulated and inferred the appropriate responses. However, it is difficult to see how this can explain our asymmetry results. Second, we did not clearly demonstrate that, holding satisfaction/quality fixed, choice is influenced by the variance around this construct. Our second experiment addresses both of these concerns.

4. Cross-Sectional Experiment

4.1. Overview

The second experiment was a 2 (high/low amount of experience) $\times$ 2 (zero/negative disconfirmation) between-subjects design. Two hundred and twenty three undergraduates participated in the experiment in return for course credit. Three subjects misunderstood the directions and were eliminated from the analysis, leaving a sample of 220 subjects.

4.2. Independent and Dependent Measures

As in the first experiment, our independent variables are experience level (a proxy for τ^2) and disconfirmation (Δ). High experience subjects were exposed to 20 experiences, while low experience subjects were exposed to only three experiences. In the zero disconfirmation condition, subjects were shown an additional outcome that was equal to the mean of the previous experiences. Subjects in the negative disconfirmation condition were shown an outcome that was two standard deviations below the mean of their previous experiences.

Our primary dependent variable is perceived quality. We adopted the measures of perceived quality used by Boulding et al. (1993), a 100-point scale anchored by "Very unfavorable" and "Very favorable." This measure was taken immediately following the experience manipulation (μ) and again following the specific visit outcome manipulation ($\mu + \Delta\tau^2$). We also used the measures developed by Boulding et al. (1993) to assess purchase intentions and likelihood to spread word of mouth, 100-point scales anchored by "Very unlikely" and "Very likely." Like Boulding et al. (1993), we find that these two measures are highly correlated ($\alpha = 0.86$), so we combine them in our analysis.

As manipulation checks, we took several measures following the experience manipulation.[7] First, we assessed perceived variability of the service and perceived consistency of the service. The first measure asked subjects to rate how variable they had found the service to be, while the second asked them to rate how consistent the service had been. We then asked subjects to rate how sure they were of the average level of quality of Café Au Lait and whether they were an occasional customer or a regular customer. Following the specific visit manipulation, we asked subjects how closely the experience on that occasion matched their expectations to assess subjective disconfirmation. This was accomplished using a standard disconfirmation question, "Overall, how closely did your experience with Cafe Au Lait on this occasion match your expectations?," measured on a 100-point "Much Worse Than Expected" to "Much Better Than Expected" scale. We also measured subjects' perceptions of how interesting and realistic they found the study, as well as their gender, age, and how many times they had stopped at a coffee shop in the last 30 days.

4.3. Method

Subjects were shown a scenario describing a coffee shop, the Café Au Lait. They were told that the coffee shop had recently opened near campus and that they had visited it (a) a few times in the low experience condition, or (b) every day for the past four weeks in the high experience condition. They then read a verbal and graphical description of their series of experiences. In the low experience condition, subjects examined a graph of three experiences. In the high experience condition, subjects examined four graphs with one graph on each page. The first graph showed the first week, the second graph showed the first two weeks, and so on, so that by in the fourth graph the subjects could see the entire series of 20 past experiences. We were careful to construct the graphs so that the mean (83) and the standard deviation (5.5) were identical between the low and high experience conditions.

Subjects were run in groups so that we could control the amount of time (40 seconds) that each page was

[7] All of the manipulation checks suggest that the experience manipulation was successful and that service variability perceptions were equivalent across groups. These are available from the authors.

viewed. For the experience manipulation to be successful, it is critical that the high experience subjects internalize the past experiences. Thus, we ensured that the high experience subjects examined each graph for a specified time to prevent them from speeding through the survey. After subjects examined the graphs, they completed a series of measures described in the next section. Following the first set of measures, subjects were exposed to the disconfirmation condition. As before, they were timed so that each subject examined this stimulus for the same amount of time (25 seconds). They were then instructed to complete the remainder of the survey at their own pace, were thanked, and released.

4.4. Results

4.4.1. Updated Quality Perceptions. To analyze our data we use analysis of covariance with updated quality perceptions as the dependent variable and initial quality perceptions and subjective disconfirmation as covariates (Cronbach and Furby 1970, Lord 1958).[8] The fit for the model is relatively good, with an R^2 of 0.51 (see Table 3 for means). As predicted, the level of experience exerts a main effect on updated quality perceptions ($t_{214} = 2.21$, $p < 0.05$). Specifically, quality perceptions for high experience subjects updated less than those of low experience subjects (-6.5 versus -8.7 for high and low experience subjects, respectively. Importantly, effects of prior perceptions of quality are controlled for in the analysis via the covariate ($t_{214} = 8.64$, $p < 0.01$).

Not surprisingly, the effect of disconfirmation on updated quality perceptions is significant ($t_{214} = 6.66$, $p < 0.01$). Subjects who experienced no disconfirmation updated their quality perceptions less (-1.8) than subjects who experienced negative disconfirmation (-14.2). Further, subjective disconfirmation exerted a significant effect on updated quality perceptions ($t_{214} = 2.81$, $p < 0.01$), replicating recent findings in the service quality literature (e.g., Anderson and Sullivan 1993, Boulding et al. 1993, Inman et al. 1997).

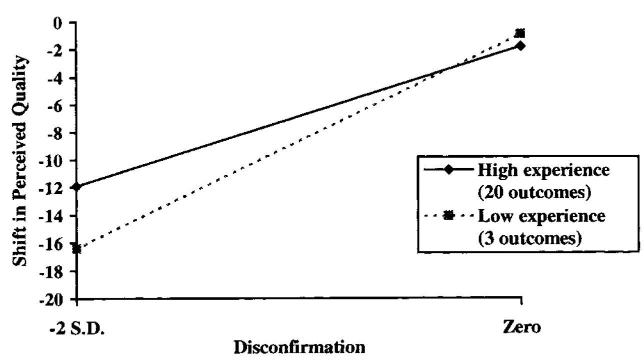
The interaction between experience and disconfirmation is significant ($t_{214} = 2.57$, $p < 0.01$). The shift in quality perceptions across the four conditions is shown in Figure 1. Both high and low experience subjects updated their quality perceptions minimally when the outcome was equal to the mean of past experiences (-1.8 for the high experience group and -0.8 for the low experience group). However, the low experience group lowered their quality perceptions significantly more than the high experience group (-16.4 versus -11.9 for the low and high experience group, respectively) when the outcome was worse than expected. Thus, consistent with prediction, experience appears to "inoculate" consumers to some extent against a single substandard outcome.

[8]Importantly, the groups were equivalent in terms of both initial quality perceptions and subjective disconfirmation.

Table 3 Cross-Sectional Experiment—Means Across Conditions (Standard Deviation in Parentheses)

Experimental Condition	Initial Quality Perceptions	Updated Quality Perceptions	Shift in Quality Perceptions	Subjective Disconfirmation	Behavioral Intentions
High Experience/ Zero Disconfirmation	81.2 (6.9)	79.4 (8.5)	−1.8	57.9 (10.5)	174.5 (22.5)
High Experience/ Negative Disconfirmation	80.2 (7.2)	68.3 (10.6)	−11.9	36.3 (12.2)	145.1 (29.3)
Low Experience/ Zero Disconfirmation	79.0 (7.6)	78.2 (7.3)	−0.9	57.1 (11.5)	154.6 (24.0)
Low Experience/ Negative Disconfirmation	78.7 (7.7)	62.3 (13.0)	−16.4	40.7 (11.5)	119.8 (31.6)

Figure 1 Cross-sectional Experiment Shifts in Perceived Quality by Condition



4.4.2 Behavioral Intentions. In examining the effects of experience level and disconfirmation on behavioral intentions, it is unclear whether or not their impact is mediated by quality perceptions. To test for potential mediation, we use the method outlined by Baron and Kenny (1986). Specifically, we examine the effect of experience (as a proxy for uncertainty) and disconfirmation on the proposed mediator, updated quality perceptions, and the dependent variable, behavioral intentions, both with and without incorporating the effect of the mediator. Perfect mediation is demonstrated if the independent variable exerts a significant effect on the mediator as well as the dependent variable but this effect becomes nonsignificant when the mediating variable is incorporated as a covariate. If the effect remains significant but the effect size significantly reduces, partial mediation is demonstrated.

We already presented evidence of the significant effects of both experience and disconfirmation on updated quality perceptions. In the second ANOVA, using behavioral intentions as the dependent variable, both experience ($t_{216} = 6.22$, $p < 0.01$) and disconfirmation ($t_{216} = 8.82$, $p < 0.01$) demonstrate significant main effects on behavioral intentions. Specifically, high experience subjects state a much greater intention to visit the service again and to tell others than do low experience subjects (160.9 vs. 137.2 for the high and low experience groups, respectively). However, when updated quality perceptions are added to the model, the effect of experience is undiminished ($t_{215} = 5.74$, $p < 0.01$), while that of disconfirmation is greatly reduced ($t_{215} = 2.74$, $p < 0.01$). Thus, these results suggest that *cumulative experience exerts a direct influence on behavioral intentions while the effect of disconfirmation is largely mediated by updated quality perceptions.*

5. Discussion and Future Research

5.1. Management Implications

Both our analytical model and our empirical results shoot holes in some seemingly reasonable quality maxims. Let us consider in particular the truisms from the Introduction:

• *"It is necessary to exceed customer expectations."*

Both our analytical model and the longitudinal experiment contradict this (although the cross-sectional experiment, which is perhaps less sensitive to this effect, does not). The longitudinal experiment showed significant positive preference shifts if customer expectations were exactly met. The reason, based on the analytical model, is that experience causes a shrinkage in the variance of the predictive distribution for the next transaction. That is, experience with a brand leads to decreased risk, and decreased risk leads to greater preference. So in fact, meeting expectations should unambiguously result in higher preference. The analytical model also suggests that even *not quite meeting* expectations might still increase preference, although this effect was not significant in the longitudinal experiment. Of course, exceeding customer expectations will still be required if the company seeks to induce customer delight (Oliver et al. 1997), but lower levels of performance may also produce positive results.

• *"If a customer expects a bad level of quality and receives it, he/she will reduce his/her level of preference for the brand."*

The analytical model and both experiments contradict this. In the cross-sectional experiment, subjects did not lower their quality perceptions of the service when their expectations were met. The longitudinal experiment showed significant *positive* preference shifts even for the nonpreferred option, indicating that even when expectations were low, meeting expectations raised preference. The analytical model explains why this preference shift can occur. Again experience shrinks the variance of the predictive distribution and reduces risk, thereby increasing preference.

• *"Given two equally-priced options, the customer will choose the one with the higher expected quality."*

Although this statement seems obviously true, it also is contradicted by both the analytical model and the longitudinal experiment. The reason, again, is risk. A higher expected quality can be outweighed by greater perceived variability. Based on the logit analysis of the data from the longitudinal experiment, perceived variance has a significant, negative impact on choice. This resulted in nearly half of our subjects choosing an option with a lower expected quality.

• *"Management should always focus on its most loyal customers."*

Given that the most loyal customers are the most experienced (which is consistent with the most typical behavioral definitions of loyalty), our research casts doubt on this seemingly self-evident maxim. The cross-sectional experiment shows that disconfirmation has the biggest preference impact on less-experienced customers, and this phenomenon is supported theoretically by the analytical model, for any nonnegative disconfirmation, and for any sufficiently large negative disconfirmation.[9] This would seem to imply that management should pay more attention to its newer (presumably less loyal) customers, because those are the customers for which quality differences will have the greatest impact. In other words, less experienced (loyal) customers are easier to lose, while more experienced (loyal) customers are harder to lose, all other things being equal.

Several other managerial implications arise from this work. First, customer satisfaction and quality measurement surveys would benefit from including an experience variable. This is because the degree to which preferences shift is dependent upon the degree of experience, with more experienced customers being more difficult to shift. All other things being equal, for maximum shift of preference, less experienced customers should be targeted. Further, in addition to measuring perceived quality, perceived variability and/or consistency in quality is important to capture as well. Both factors influence overall quality perceptions and behavioral intentions.

Our results suggest that it is insufficient for a good or service to be perceived as better (i.e., a higher expected quality) than its competitor. For example, new Brand A, even though it has a higher expected value than Brand B, may not be preferred to Brand B because of greater perceived variability resulting in its perceived value being worse. The implication to management launching a new product is that actual trial may be more powerful than supplying information through such methods as advertising. This suggests that coupons, promotions, and other ways to induce trial may be necessary in the early stages of a product launch, even if a positive brand image has been already created through advertising.

[9]Interestingly, there is an interval for which a negative disconfirmation produces a smaller shift in preference for the less experienced customer—whenever that customer's positive preference shift from variance reduction is roughly equivalent to the negative preference shift from mean reduction.

Finally, worse-than-expected quality hurts more than better-than-expected quality helps. This replicates previous findings by several researchers (e.g. Anderson and Sullivan 1993, DeSarbo et al. 1994, Rust et al. 1995). The managerial takeaway is that problems should be addressed first, and then positive opportunities. In other words, process improvement initiatives should first focus on eliminating unsatisfactory service encounters (i.e., providing consistent quality) before trying to delight customers.

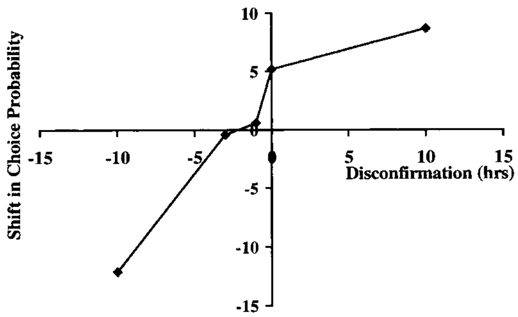
5.2. Contributions to the Quality Literature

Our results contribute to the quality literature in four respects. First, they suggest that consumers are sensitive to not only the average performance of a product, but also to its *variability* around this mean. In other words, it is not sufficient for a brand to strive to increase its overall quality—it must also strive to reduce the risk of an outcome deviating from this performance expectation. Second, consumers are more likely to rechoose a brand (i.e., retention is increased) if the brand performs as expected. Contrary to the traditional satisfaction model, meeting expectations can sometimes be interpreted by consumers as a favorable outcome, as the experience provides more information about the brand and reduces the perceived variance of the brand's future performance. Third, probability of choice is *not* necessarily adversely affected by an outcome that is slightly below expected. This is apparently because the customer becomes less concerned with downside risk if the disconfirmation is relatively minor. Finally, if negative disconfirmation is relatively substantial, quality perceptions and behavioral intentions are negatively affected, but this effect is moderated by the amount of prior experience.

Figure 2 shows a graph of the mean shifts in choice probability across the five disconfirmation conditions in the longitudinal experiment. One notes that consumers' reaction to disconfirmation is nonlinear and they appear more sensitive to negative disconfirmation than to positive disconfirmation. While this curve is similar to the effects one typically expects under Kahneman and Tversky's prospect theory (1979), our theoretical perspective and empirical results suggest that the curve may be "choice-neutral"[10] at a value less

[10]We call the curve "choice-neutral" at the point where observed quality results (on average) in unchanged probability of choice.

Figure 2 **Shift in Choice Probability in Response to Disconfirmation**



than zero disconfirmation. Specifically, the curve is choice-neutral at a negative disconfirmation of about two hours below expected, which is equivalent to approximately one half of a standard deviation from the mean. Importantly, the direction of this deviation is consistent with our prediction (Proposition 4) that a small negative disconfirmation can produce an increase in preference.

Our results suggest that the most important time for a brand to establish its quality perceptions in the minds of consumers is at the time that consumers have little prior experience with the category. In Bayesian terms, this is at the point when consumers hold a more diffuse prior expectation regarding the brand's performance. Ironically, the time where many consumers are inexperienced with the product is precisely the point in the product life cycle where the manufacturer is often struggling with maintaining consistent product quality. This may explain the Golder and Tellis (1993) findings regarding the failure of many pioneers. If the pioneer struggles with quality, a subsequent (more consistent quality) entrant can gain a perceived quality advantage with consumers. Alternatively, once the pioneer has established a relatively high prior expectation of performance and a low perceived variability (through knowledge gained by consumption experience), this imposes a barrier to later entrants into the market since the pioneer finds customers are much easier to retain. A subsequent entrant must have a significantly higher level of expected quality to counteract its higher perceived variability and risk.

5.3. Future Research and Limitations

There are many interesting ways in which this research might be extended. For example, the model might be made richer (albeit more complicated) by the inclusion of effects not currently accounted for, such as allowing for variables that affect perceived quality, or otherwise change (bias) the perception and updating process. Interesting work in this regard is already underway (Boulding et al. 1998). The existence of threshold effects is another promising area for research. In other words, updating may only take place when absolute disconfirmation exceeds a certain threshold. Otherwise the observation is simply viewed as an expected outcome. It is also possible that those thresholds may vary across customers. That might explain the fact that many of our subjects did not update their preferences.

While our results suggest that, on average, consumers update their expectations in a largely Bayesian manner, some consumers are probably more Bayesian than others. In the first experiment, choice probability for most subjects increased in the 10 hour disconfirmation condition and decreased in the -10 hour disconfirmation condition. However, in the zero, negative one hour, and negative three hour disconfirmation conditions, the modal subjective probability shift was zero. Approximately half the subjects in both the zero and negative one hour conditions reported that their choice probability would be unaffected, while a third of the subjects in the negative three hour condition gave this response. Even allowing for measurement error, some "updating heterogeneity" appears evident among consumers.

Additional research is warranted to explore the variables that account for the heterogeneity in consumers' updating processes. For instance, the Boulding et al. (1993) notion of "should" expectations could be extended to expectations of *variability* in quality. Further, individual differences in need for cognition (Cacioppo and Petty 1982) or tendency to engage in post-purchase regret (e.g., Inman and Zeelenberg 1998) might be important moderators of perception updating and changes in probability of choice in response to disconfirmation. On the other hand, it is important to explore the contexts in which consumers tend to form expectations of quality that are resistant to change.

Limitations of our study deserve mention. First, as in all research, one must be careful in overgeneralizing results. However, we tested our updating model in two quite different experimental designs and found consistent support for our thesis. Second, we tested our propositions in lab contexts and our results are based on self-reports. It is important to replicate this research in other contexts to provide greater confidence in our results. For example, our subjects observed several iterations of outcomes over a short period of time. In a field context, consumers' priors might become more diffuse (due to forgetting) as the interpurchase time increases or become less diffuse as additional information is obtained from advertising or word of mouth. Finally we should note the assumption of normal distributions. If customers' priors and likelihood distributions are not normally distributed, then a more complicated Bayesian formulation would result. This research provides the basis for subsequent work in the important area of how consumers dynamically update their quality perceptions and their preferences.[11]

[11]The authors thank Vanderbilt's center for Service Marketing for providing partial support for this project, and gratefully acknowledge the comments of Priya Raghubir, Steve Hoch, Rick Staelin, the area editor and reviewers, and seminar participants at the University of Chicago.

References

Anderson, Eugene W., Mary W. Sullivan. 1993. The antecedents and consequences of customer satisfaction for firms. *Marketing Sci.* **12** (Spring) 125–143.

Baron, Reuben M., David A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J. Personality and Soc. Psych.* **51** (Dec.) 1173–1182.

Bell, David. 1982. Regret in decision making under uncertainty. *Oper. Res.* **30** 961–981.

——. 1985. Disappointment in decision making under uncertainty. *Oper. Res.* **35** 1–27.

Berger, James O. 1985. *Statistical Decision Theory and Bayesian Analysis.* Second edition. Springer-Verlag, New York.

Bolton, Ruth, James Drew. 1991a. A longitudinal analysis of the impact of service changes on customer attitudes. *J. Marketing* **55** (January) 1–9.

——, James Drew. 1991b. A multistage model of customers' assessment of service quality and value. *J. Consumer Res.* **17** (March) 375–384.

——. 1998. A dynamic model of the duration of the customer's relationship with a continuous service provider: the role of satisfaction. *Marketing Sci.* **17** (1) 45–65.

Boulding, William, Ajay Kalra, Richard Staelin, Valarie A. Zeitham. 1993. A dynamic process model of service quality. *J. Marketing Res.* **30** (February) 7–27.

——, Ajay Kalra, Richard Staelin. 1998. The quality double whammy: The rich stay rich and the poor stay poor: Working paper. Duke University, Durham, NC.

Burr, Irving W. 1976. *Statistical Quality Control Methods,* Marcel Dekker, New York.

Cacioppo, John T., Richard E. Petty. 1982. The need for cognition. *J. Personality and Soc. Psych.* **42** (January) 116–131.

Cronbach, Lee J., Lita Furby. 1970. How we should measure 'change'—or should we? *Psych. Bull.* **74** (July) 68–80.

Cronin, J. Joseph, Steven A. Taylor. 1992. Measuring service quality: a reexamination and extension. *J. Marketing* **56** (July) 55–68.

——, Steven A. Taylor. 1994. SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service quality. *J Marketing* **58** (January) 125–131.

Deming, W. Edwards. 1986. *Out of the Crisis,* MIT Center for Advanced Engineering Study, Cambridge, MA.

DeSarbo, Wayne, Lenard Huff, Marcello Rolandelli, Jungwhan Choi. 1994. On measurement of perceived service quality: A conjoint measurement approach. Roland T. Rust and Richard L. Oliver Eds. *Service Quality: New Directions in Theory and Practice.* Sage Publications, Thousand Oaks, CA, 199–220.

Doherty, Michael E., Clifford R. Mynatt, Ryan D. Tweney, Michael D. Schiavo. 1979. Pseudodiagnosticity. *Acta Psych.* **43** 111–121.

Edwards, Ward. 1968. Conservatism in human information processing. B. Kleinmuntz, ed. *Formal Representation of Human Judgment.* Wiley, New York.

Fischhoff, Baruch, Ruth Beyth-Marom. 1983. Hypothesis evaluation from a Bayesian perspective. *Psych. Rev.* **90** 239–260.

Fishburn, Peter C. 1988. *Nonlinear Preference and Utility Theory.* John Hopkins University Press, Baltimore, MD.

Fornell, Claes. 1992. A national customer satisfaction barometer: The Swedish experience. *J. Marketing* **56** (January) 6–21.

Gale, Bradley. 1997. Satisfaction is not enough. *Marketing News* (October 27) 18.

Gigerenzer, G., W. Hell, H. Blank. 1988. Presentation and content: the use of base-rates as continuous variable. *J. Experiment. Psych. Human Perception and Performance* **14** 513–525.

Ginossar, Zvi, Yaacov Trope. 1987. Problem solving in judgment under uncertainty. *J. Personality and Soc. Psych.* **52** 464–474.

Golder, Peter N., Gerald J. Tellis. 1993. Pioneering advantage: Marketing logic or marketing legend? *J. Marketing Res.* **30** (May) 158–170.

Inman, J. Jeffrey, James S. Dyer, Jianmin Jia. 1997. A generalized utility model of disappointment and regret effects on post-choice valuation. *Marketing Sci.* **16** (2) 97–111.

——, Marcel Zeelenberg. 1998. Wow, I could've had a V8!: The role of regret in consumer choice. Working paper, University of Wisconsin, Madison, WI.

Jia, Jianmin, James S. Dyer. 1996. A standard measure of risk and risk-value models. *Management Sci.* **42** 1691–1705.

Juran, J. M., Frank M. Gryna, Jr. 1980. *Quality Planning and Analysis: From Product Development through Use,* McGraw-Hill, New York.

Kahneman, Daniel, Amos Tversky. 1972. Subjective probability: A judgment of representativeness. *Cogn. Psych.* **47** 263–291.

——, ——. 1973. On the psychology of prediction. *Psych. Rev.* **80** 237–251.

——, ——. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47** 263–290.

Keeney, Ralph L., Howard Raiffa. 1976. *Decision with Multiple Objectives: Preferences and Value Tradeoffs,* John Wiley & Sons, New York.

Kopalle, Praveen K., Donald R. Lehmann. 1995. The effects of advertised and observed quality on expectations about new product quality. *J. Marketing Res.* **32** (August) 280–290.

Little, John D. C. 1966. A model of adaptive control of promotional spending. *Oper. Res.* **14** (November–December) 1075–1098.

Lord, Frederic M. 1958. The utilization of unreliable difference scores. *J. Ed. Psych.* **49** (3) 150–152.

Luce, R. Duncan. 1959. *Individual Choice Behavior: A Theoretical Analysis,* John Wiley & Sons, New York.

Machina, Mark J. 1987. Choice under uncertainty: problems solved and unsolved. *Econom. Perspectives* **1**(1) 121–154.

Manrai, Ajay K. 1995. Mathematical models of brand choice behavior. *European J. Operational Res.* **82** 1–17.

Markowitz, H. M. 1959. *Portfolio Selection,* John Wiley & Sons, New York.

——. 1987. *Mean-Variance Analysis in Portfolio Choice and Capital Markets.* Basil Blackwell, New York.

McFadden, Daniel. 1974. Conditional logit analysis and subjective probability. Paul Zarembka, Ed. *Frontiers in Econometrics.* Academic Press, New York, 105–12

——. 1976. Quantal choice analysis: A survey. *Ann. Econom. Soc. Measurement* **5** 363–390.

Meyer, Jack. 1987. Two-moment decision models and expected utility maximization. *Amer. Econom. Rev.* **77** 421–430.

Oliver, Richard L. 1980. A cognitive model of the antecedents and consequences of satisfaction decisions. *J. Marketing Res.* **42** (November) 460–469.

——, Wayne DeSarbo. 1988. Response determinants in satisfaction judgments. *J. Consumer Res.* **14** (March) 495–507.

——. 1997. *Satisfaction*, McGraw-Hill, New York.

——, Roland T. Rust, Sajeev Varki. 1997. Customer delight: foundations, findings, and managerial insight. *J. Retailing* **73** (Fall) 311–336.

Parasuraman, A., Valarie A. Zeithaml, Leonard Berry. 1985. A conceptual model of service quality and its implications for further research. *J. Marketing* **48** (Fall) 41–50.

Parasuraman, A., Valarie A. Zeithaml, Leonard Berry. 1988. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *J. Retailing* **64** (1) 12–40.

Payne, John W. 1973. Alternative approaches to decision making under risk: moments versus risk dimensions. *Psych. Bull.* **80** (December) 435–453.

Reichheld, Frederick F. 1996. Learning from customer defections. *Harvard Bus. Rev.* (**March–April**) 56–69.

Rust, Roland T., Anthony J. Zahorik 1993. Customer satisfaction, customer retention and market share. *J. Retailing* **69** (Summer) 193–215.

——, Anthony J. Zahorik, Timothy L. Keiningham. 1995. Return on quality (ROQ): making service quality financially accountable. *J. Marketing* **59** (April) 58–70.

Slovic, Paul, Sarah Lichtenstein. 1971. Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organ. Behavior and Human Performance* **6** 649–744.

Teas, R. Kenneth. 1993. Expectations, performance evaluation, and consumer perceptions of quality. *J. Marketing* **57** (October) 18–34.

Thurstone, L. L. 1927. A law of comparative judgment. *Psych. Rev.* **34** 273–286.

Tversky, Amos, Daniel Kahneman. 1974. Judgment under uncertainty: heuristics and biases. *Science* **185** 1124–1131.

von Neumann, John, Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior,* Princeton University Press, Princeton, NJ.

Wason, P. C. 1960. On the failure to eliminate hypotheses in a conceptual task. *Quart. J. Experiment. Psych.* **12** 129–140.

Zellner, Arnold. 1971. *An Introduction to Bayesian Inference in Econometrics,* John Wiley & Sons, New York.

This paper was received January 2, 1996, and has been with the authors 26 months for 2 revisions; processed by William Boulding.