Exploring the Role of Alternative Perceived Performance Measures and Needs-Congruency in the Consumer Satisfaction Process

Jochen Wirtz
Department of Marketing
Faculty of Business Administration
National University of Singapore

Anna Mattila
School of Hotel Restaurant and Recreation Management
The Pennsylvania State University

Prior research suggests that perceived performance of a product or a service is directly linked to postpurchase satisfaction. We argue that this causal relationship might be a measurement artifact and/or insufficient modeling of the satisfaction process rather than an accurate assessment of how consumers form satisfaction judgments.

To test our hypotheses, a $2 \times 2 \times 2$ (Performance $\times$ Expectations $\times$ Needs) factorial design was used with 2 types of perceived performance measures (value-laden and objective). The findings demonstrate that the observed direct link from perceived performance to overall satisfaction diminishes when more objective perceived performance indicators replace the commonly used value-laden measures. Furthermore, desire-congruency was found to contribute independently to satisfaction over and above a disconfirmation-of-expectations standard. In fact, desire-congruency was found to be a better predictor of satisfaction than disconfirmation-of-expectations. Finally, our results suggest that the direct performance–satisfaction link becomes insignificant when the modeling of the satisfaction process is improved. Taken together, these findings support the view that the frequently observed high correlations between perceived performance and satisfaction might be a reflection of the type of measures used and/or insufficient capturing of the evaluative process leading to satisfaction, rather than support for a direct causal link.

Consumer satisfaction is a central issue in marketing thought and practice (Oliver, 1997; Yi, 1990). Satisfaction is a major outcome of marketing activity and serves to link processes of decision making and consumption with postpurchase phenomena, such as attitude change, complaining behavior and word-of-mouth, repeat purchase, and brand loyalty (Bearden & Teel, 1983; Fornell, 1992; Oliver, 1980). The centrality of consumer satisfaction is reflected by its inclusion in the marketing concept, which focuses on profit generation through “determining the needs and wants of target markets and delivering the desired satisfactions” (Kotler, Ang, Leong, & Tan, 1996, p. 20). The need to translate this philosophical statement into operational guidelines has kindled extensive research on consumer satisfaction since the early 1970s.

Consumer satisfaction is generally defined as an evaluative response to the perceived outcome of a particular consumption experience (e.g., Cadotte, Woodruff, & Jenkins, 1987; Day, 1984; Westbrook & Oliver, 1981, 1991; Yi, 1990). A psychological comparison of some sort is a central component in the conceptualization of the satisfaction process. A vast majority of the current models are based on a comparison between perceived performance and a preconsumption comparison standard (Yi, 1990). The disconfirmation model is parsimonious and intuitively appealing (Iacobucci, Ostrom, Braig, & Bezjian-Avery, 1996) and it has received strong empirical support (e.g., Boulding et al., 1993; Churchill & Surprenant, 1982; Oliver, 1980; Oliver & DeSarbo, 1988; Westbrook & Oliver, 1991).
In addition to the mediated impact of perceived performance via a comparison process, a number of studies has also shown a strong direct link from perceived performance to satisfaction (Anderson & Sullivan, 1993; Oliver, 1993, 1994; Patterson, 1993). In this article, we postulate that the existence and strength of such a link may depend on several factors. First, evaluative components in commonly used perceived performance measures (e.g., “the performance was excellent” to “poor”) may capture part of the satisfaction construct, effectively measuring at least partly the same construct, and therefore showing high correlations. Second, we argue that any potential impact of perceived performance on satisfaction must be indirect. That is, any performance perception is mediated through a mental process of some sort, placing this perception into a person’s value system. The observed high correlations between perceived performance and satisfaction (often interpreted as causal) may therefore be a reflection of the applied comparison standards (often disconfirmation-of-expectations) not adequately capturing those mental processes leading to satisfaction.

The main objective of this study is to examine under what conditions there is a direct link between satisfaction and performance. To explore this issue, we distinguished between two types of performance measures: objective and subjective. As a secondary goal, we examined the additional contribution of a needs-congruency standard in explaining the relation between satisfaction and perceived performance.

**LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT**

**Perceived Performance—A Direct Causal Antecedent of Satisfaction?**

Recent empirical findings seem to confirm the conceptualization that, apart from driving the intervening disconfirmation variable, perceived performance directly influences satisfaction (e.g., Anderson & Sullivan, 1993; Oliver, 1993, 1994; Patterson, 1993). It has been empirically shown that direct causal links from perceived performance to satisfaction can significantly increase the proportion of explained variance in satisfaction, and that sometimes perceived performance was an even better predictor of satisfaction than the disconfirmation-of-expectations variable (Churchill & Surprenant, 1982; Tse & Wilton, 1988). These results may cast some doubt on the appropriateness of disconfirmation-of-expectations as a mediating variable and expectations as a comparison standard.

A measurement problem, whereby the applied perceived performance and satisfaction measures captured almost entirely or at least partially the same construct, might offer a potential explanation for the direct link between performance and satisfaction within the disconfirmation-of-expectations framework. In Churchill and Surprenant’s (1982) study, some between-construct correlations (between perceived performance and satisfaction measures) were higher than within-construct correlations (between different satisfaction measures). This finding indicated a lack of discriminant validity between measures, and therefore, three out of five satisfaction measures had to be dropped. Other researchers encountered similar problems. Tse and Wilton (1988) obtained a correlation of .81 between satisfaction and perceived performance measures, which was the second highest correlation between any two measures in their study.

Cadotte et al. (1987) indicated concern about the use of anchor words in performance measures that imply evaluation. Commonly used performance scales include anchor words such as good/bad, fast/slow, and friendly/unfriendly. For instance, widely accepted standard measures used in Churchill and Surprenant’s (1982) article used the following 7-point semantic differential scales anchored by good sound–poor sound, not very good–excellent, very inferior–very superior, and terrible–excellent. These scales show strong evaluative components, and therefore probably already measure some aspect of satisfaction (Wirtz, 1993). Spreng and Olshavsky (1992) suggested that perceived performance should be separated into two distinct constructs: (a) perceived performance that is “evaluationless,” or cognitive registering of outcomes, and (b) an evaluative component. This notion of a more objective-performance perception versus a subjective evaluation of perceived performance is advanced in this study.

Similarly, Cadotte et al. (1987) suggested that scales measuring true beliefs about an objective level of performance should be used instead of evaluations. Measures that cue the respondent to an objective level of an attribute performance can be developed. For instance, the speed of a service can be measured in time units rather than fast/slow, and menu variety can be measured by the number of items on the menu instead of few/many. In the remainder of this article, we apply objective performance measures to refer to evaluationless or more objective performance measures and subjective performance measures to refer to the more value-laden performance measures.

Independent of the performance measurement argument, the notion of a direct causal link between performance and satisfaction can be rejected on logical grounds. A performance of x has no meaning to us, unless we can position it in our value system. Performance can only have meaning in the context of evaluation, and evaluation implicitly requires some comparison standard. This logic is illustrated in the following example. A perceived travel time of 12 hr (objective perceived performance measure) for a trip from London to New York could be perceived as “good” or “bad” (subjective performance measure) depending on the frame of reference. One hundred fifty years ago it would have been an incredible performance. Twenty years ago it would have been a good “as expected” performance, and today it seems slow with jumbo jets traveling this distance in less than 7 hr and the Concorde in less than 4 hr. Therefore, for evaluating the 12-hr perfor-
Discussions on these standards are provided elsewhere (e.g., Spreng & Olshavsky, 1993; Yi, 1990), and this review focuses on needs and desires as a promising alternative (or additional) standard to the commonly used disconfirmation-of-expectations.

Westbrook and Reilly (1983) considered it a major weakness of the disconfirmation-of-expectations model that it does not provide sufficient differentiation between cognitive and evaluative notions. Expectations refer to beliefs about the expected product performance, but its antecedents remain unclear. Although most researchers tend to agree that needs and values influence expectations, Oliver (1997) went a step further. He argued that predictive standards can incorporate other standards such as desires and needs. However, what is expected may or may not correspond to what is wanted or desired. Product breakdown, improper function, and unattractive appearance produce dissatisfaction, regardless of whether they are expected (LaTour & Peat, 1979a). In practice, desires (or needs and wants) and expectations often fall together, because consumers choose purposefully to achieve their goals, but not always. For example, a consumer might be forced, due to time and location constraints, to dine at a restaurant that does not possess the desired attributes. Although the dining experience may be as expected, it is unlikely that this consumer will be fully satisfied, because he or she probably compares this dining experience with one at a more highly valued restaurant (LaTour & Peat, 1979a; Woodruff et al., 1983). Wirtz and Bateson (1995) showed in an experimental setting that a long response time (9 sec for accessing a bank’s mainframe via a modem and telephone line and mainframe processing time) of a personal computer homebanking service resulted in dissatisfaction even when the response time was actually shorter than expected. In sum, the two comparison standards—needs and expectations—may be distinct constructs, psychologically processed separately at least in some situations.

The notion that disconfirmation can only occur for aspects on which consumers hold prior expectations represents an additional shortcoming of the expectations-disconfirmation framework. Consumers can be dissatisfied with aspects of a product they were not aware of before consumption (Westbrook & Reilly, 1983).

To address these shortcomings, Westbrook and Reilly (1983) drew on Locke’s (1967) seminal analysis of job satisfaction. Locke asserted that satisfaction/dissatisfaction is a response triggered by a cognitive-evaluative process in which perceptions of an object, action, or condition are compared to one’s values (or needs and wants). The smaller the discrepancy between perceptions of the object, action, or condition and one’s values, the more favorable is the evaluation and the greater the level of satisfaction. Conversely, the greater the value-percept disparity, the less favorable the evaluation, followed by goal frustration and dissatisfaction. When values and expectations have been separated experimentally, values rather than expectations

Would Improved Modeling of the Evaluation Process Help?

Assuming that the frequently used mediating variable disconfirmation-of-expectations is not a good (enough) standard and does not sufficiently explain the variance in satisfaction, then an indirect, by disconfirmation-of-expectations unmediated impact from perceived performance would show as a direct causal link to satisfaction. This supposition implies that if one or more suitable standards replaced or was added to disconfirmation-of-expectations, the direct link from perceived performance to satisfaction should become smaller. It should disappear when all the mental processing used to evaluate perceived performance (i.e., mostly comparison processes) was captured. In other words, rather than introducing a direct causal link from perceived performance to satisfaction, more suitable comparison standards should be used in addition to (or instead of) disconfirmation-of-expectations.

Unfortunately, the most controversial variable in the disconfirmation paradigm is precisely this comparison standard. A large amount of theoretical debate and empirical research has revolved around the question on what standard(s) people use in the comparison process. This line of research has resulted in six broad classes of pre-experience standards. They are

1. Predictive expectations (e.g., Liechty & Churchill, 1979; Oliver, 1980; Woodruff, Cadotte, & Jenkins, 1983).
2. Ideal performance (e.g., Miller, 1977; Sirgy, 1984).
3. Needs and wants coined as value-percept (Swan & Trawick, 1981; Westbrook & Reilly, 1983) or desires (Spreng & Olshavsky, 1993).
4. Experience-based standards (Cadotte et al., 1987; LaTour & Peat, 1979a, 1979b; Woodruff et al., 1983).
5. Comparisons with social norms (Swan, 1983).
6. Multiple standards (e.g., Sirgy, 1984; Spreng, MacKenzie, & Olshavsky, 1996; Tse & Wilton, 1988).
determined satisfaction (Locke, 1967). In other words, performance in relation to needs and wants, rather than expectations, has been proposed to be the primary determinant of satisfaction (Westbrook & Reilly, 1983).

Spreng and Olshavsky (1993) operationalized desires as a comparison standard in a means-end framework. They defined values as consumers’ mental representations of end states, where “higher-level values leading to desires for products that will provide certain benefits, and these benefits in turn specifying the attributes and the levels of attributes desired in the product” (p. 171). They provided an example in which living a healthy life represents a high-level value. This higher order value can manifest itself in a desire to eat those foods that provide the benefit of reducing the risks of developing certain illnesses. This desired benefit may in turn be specified in terms of attributes (e.g., cholesterol content) and their performance levels (e.g., cholesterol-free). Prior research (Barbeau, 1985; Spreng & Mackoy, 1996; Spreng & Olshavsky, 1993) suggests, with empirical support, that the relation between a product’s performance and a consumer’s values is a significant determinant of satisfaction, explaining more variance in satisfaction than disconfirmation-of-expectations. Specifically, satisfaction occurs when a product meets or exceeds a consumer’s desires and dissatisfaction results when the performance falls short of the consumer’s desires.

In this article, we distinguish between expectations and needs congruency. In particular, expectations here are referred to as predictive expectations in accordance with recent interpretations by Spreng and Olshavsky (1993) and Spreng et al. (1996). One could, however, argue that expectations could be operationalized as “desired expectations.” This conceptualization of expectations would be very similar, if not identical, to needs congruency, and the disconfirmation-of-expectations model would still apply. Nevertheless, we feel that in practice the disconfirmation-of-expectations variable is predominantly based on predictive expectations (perhaps with a dose of experience-based expectations), and not on other alternative standards. Moreover, desire or needs congruency might provide a less ambiguous label. Accordingly, Spreng et al. (1996) stated the following:

The only way to gain a clear understanding of the impact of expectations on satisfaction is to avoid confounding predictive expectations (what a person believes is likely to happen in the future) with judgments that implicitly require the use of several possible standards of comparison (e.g., desires, industry norms, equity, best brand). (p. 16)

In summary, we propose that the inadequate capturing of the evaluation process might offer a plausible explanation for the observed direct link between perceived performance and satisfaction. Although needs-congruency might fail to account for alternative comparison standards such as affect or equity, in prior studies this variable has been more powerful in predicting satisfaction than disconfirmation-of-expectations as a comparison standard (Barbeau, 1985; Spreng et al., 1996; Spreng & Olshavsky, 1993). If needs congruency indeed better captures the evaluation process, then the observed direct link from perceived performance to satisfaction should be weaker when needs-congruency is used as a mediating variable instead of disconfirmation-of-expectations.

H2: A more complete model of the evaluation process, specifically the inclusion of needs-congruency and the distinction between subjective and objective measures of performance, will reduce or may even eliminate the observed direct link from perceived performance to satisfaction.

METHOD

Experimental Design

A $2 \times 2 \times 2$ (Expectations $\times$ Needs $\times$ Performance) between-subject factorial design was used. Although prior experiments on comparison standards have successfully manipulated expectations, they have failed to include other standards for comparison. This neglect introduces potential bias toward the expectations standard, given that much of the variance in satisfaction is likely to be due to this manipulation. To test H2, it was therefore considered crucial to manipulate both needs and expectations so that their respective effects on the perceived performance–satisfaction link could be compared.

The participant pool was composed of 115 MBA students from a Singaporean university. Four incomplete questionnaires had to be discarded, thus resulting in 13 to 14 participants per experimental cell. Of the sample, 69% were men and 31% were women, the age of respondents ranged from 25 to 40 years with a mean age of 29, and there were 78% part-time and 22% full-time MBA students.

Scenario Development and Manipulations

To enable the simultaneous manipulation of expectations, needs, and performance, the focal product had to be carefully chosen. There were two main criteria for the selection of a study context. First, to allow a true experimental manipulation, distinctly high and low needs had to exist within a single group of consumers. Second, the product had to possess an attribute whose performance level could be measured on an objective scale.

Experimentation with four pretest scenarios (fast food, weight-loss center, key duplication service, and courier service) indicated that a courier service scenario, using delivery time to manipulate attribute expectations, needs, and performance, was realistic and effective. A pretest of the courier service scenario ($n = 16$) indicated that all three ma-
Manipulations were significant at the .05 level and in the anticipated direction.

The courier service scenario described a situation where an employee of a medium-sized company had to make an urgent courier delivery from Singapore to Da Nang, a small city in Vietnam. Because past experience with a product can influence consumer expectations (LaTour & Peat, 1979a, 1979b; Miller, 1977; Pieters, Koelmeijer, & Roest, 1995; Rust, Inman, & Jia, 1997), a manipulation of expectations is easily contaminated by the respondents’ past consumption experiences. To overcome this problem, a small city in Vietnam was chosen as a destination in this study. At the point in time of this study, participants were unlikely to be familiar with courier services to this destination (Vietnam had just started opening its economy to international trade and foreign direct investment at the time of data collection). In fact, none of the respondents had sent a courier package to Vietnam before (although 97% had personal experience in using courier services). This lack of specific experience enabled us to cleanly manipulate the participants’ expectations in an experimental setting.

Needs were manipulated at two levels: fast (Wednesday p.m.) versus slow (Friday a.m.). The scenarios the participants read described an urgent meeting with bankers, government officials, and lawyers. This meeting would have to be postponed if the documents had not arrived by the manipulated date, leading to significant inconvenience to the parties involved. Expectations of the delivery date were manipulated via the courier company’s service guarantee (fast: Wednesday a.m. vs. slow: Friday p.m.). Using seller’s promises was considered an appropriate way to manipulate predictive expectations (Woodruff, Clemons, Schuman, Gardial, & Burns, 1991). The promise was operationalized in the form of a guarantee, which is considered a strong and explicit promise (McDougall, Levesque, & VanderPlaat, 1998). The performance of the service was manipulated as fast (Wednesday p.m.) versus slow (Thursday p.m.).

The chosen manipulation levels allowed for positive and negative disconfirmation-of-expectations. For example, performance > expectations (Wednesday p.m. vs. Friday p.m. for performance and expectations, respectively) and performance < expectations (Thursday p.m. vs. Wednesday a.m.). For positive and negative needs-congruency, for example, performance > needs (Wednesday p.m. vs. Friday a.m.) and performance < needs (Thursday p.m. vs. Wednesday p.m.). Past studies often only manipulated the direct antecedents of disconfirmation-of-expectations (i.e., expectations and perceived performance) and not the antecedents of alternative standards (e.g., Spreng & Olshavsky, 1993; Tse & Wilton, 1988; one notable exception is the study by Spreng et al., 1996, in which the antecedents of all competing standards were manipulated), a practice likely to introduce biases in favor of the disconfirmation-of-expectations model. This study avoided such biases as the antecedents of both standards were manipulated.

Measures

Needs perception was measured by a two-item, 7-point scale. The first question asked for the time at which the meeting was scheduled in Vietnam (NEED1). The second item on this scale instructed participants to estimate the latest time by which the documents should arrive in Vietnam to avoid being late for the meeting (NEED2; as a manipulation check of needs). Similarly, a two-item, 7-point scale was administered to measure expectations (as a manipulation check of expectations). The first question asked for an expected delivery time for the documents (EXPECT1). The second item probed estimation about the time by which the courier firm could assure the delivery of the documents (EXPECT2).

Two measures of perceived performance were used to test H1. The first measure included an evaluative component similar to those used in numerous past studies. Specifically, participants were asked to rate the performance of the courier service on a scale anchored by extremely slow and extremely fast (SUBJ-PERF). The second measure, being time-based, provided a more objective measure of perceived performance. The participants were asked to indicate the time at which the documents actually arrived in Vietnam on a scale ranging from 1 (Wednesday a.m.) to 7 (Saturday a.m.). This OBJ-PERF variable was recoded to 7 for Wednesday a.m. and 1 for Saturday a.m. so that a high value reflects a high performance.

Disconfirmation-of-expectations was measured with a two-item, 7-point, semantic differential scale. The first item was Oliver’s (1980) scale with worse than expected and better than expected as anchors (EXP-DISC1). The second item was adapted from Churchill and Surprenant (1982) and was anchored by the following statements: (a) Mr. Lim’s expectations about the speed of the courier service were too low: The delivery was faster than expected and (b) Mr. Lim’s expectations about the speed of the courier service were too high: The delivery was slower than expected (EXP-DISC2).

The construct of needs congruence was operationalized as a subjective evaluation of the discrepancy between the service performance and needs. Again, a two-item, 7-point scale was used. The first item was an adaptation from Westbrook and Reilly’s (1983) value-percept scale, which was anchored at the positive end by provides exactly what I need. The measure was redefined in accordance with Spreng and Olshavsky’s suggestions (1993) so that it also captured performances exceeding customers’ needs. The new 7-point semantic differential scale was anchored by provided less than he needed and provided more than he needed (NEED-CON1). The second measure was a Likert-type scale with a following statement United Express courier service far exceeded Mr. Lim’s needs (NEED-CON2).

Overall satisfaction was measured using Westbrook and Oliver’s (1981) three-item scale. The scale consisted of a 7-point rating scale ranging from delighted to terrible, a 11-point rating scale ranging from 100% (completely satisfied) to 0% (not at all satisfied; SAT2), and a behavioral ten-
dency 11-point rating scale ranging from Certain I’d do it again to No chance I’d do it again (SAT3). For hypothesis testing, these individual items were combined to yield an overall measure of satisfaction (SAT).

### Procedures

A cover story disguising the real purpose of the study was used to reduce hypothesis guessing. Each participant was randomly assigned to one of the eight experimental conditions. To avoid communication among participants, they were instructed not to look at each other’s questionnaires and not to communicate with neighboring participants for the duration of the experiment. Immediately after the cover page, a written scenario of the consumption experience was presented. This page also contained manipulations of needs and expectations. Participants were instructed to complete reading this scenario and to answer six questions (including the manipulation checks of needs and expectations) before proceeding to the next page. The scenario continued on the second page in which the performance was manipulated. Having completed the entire scenario, the participants were asked to complete the questionnaire.

### RESULTS

#### Manipulation Checks and Reliability

A series of analyses of variance (ANOVAs) were performed to assess the effectiveness of the three manipulations. The results from these one-way ANOVA analyses showed that all manipulations were successful. Significant main effects were found for the needs perception, $F(1, 109) = 529.1, p < .001$ ($M = 2.0$ for the low-needs condition and $M = 4.8$ for the high-needs condition) and expectations, $F(1, 109) = 40.5, p < .001$ ($M = 3.0$ for the low-expectations condition and $M = 4.9$ for the high-expectations condition). The main effects for the two performance measures were also significant. For subjective perceived performance, $F(1, 109) = 39.9, p < .001$ ($M = 3.3$ for the low-performance condition and $M = 5.3$ for the high-performance condition) and for objective performance, $F(1, 109) = 296.0, p < .001$ ($M = 4.1$ for the low-performance condition and $M = 6.0$ for the high-performance condition).

The reliability of all multi-item measures was adequate. The Cronbach’s alpha values were .86 for NEED, .78 for EXPECT, .89 for EXP-DISC, .81 for NEED-CON, and .95 for SAT. The correlation between the two perceived performance measures (OBJ-PERF and SUBJ-PERF) was .44, indicating that, as expected, the objective perceived performance measure and the subjective performance measure do not capture the same construct.

#### Comparing the Determinants of Objective and Subjective Performance Perceptions

Two three-way ANOVAs were conducted, examining the effects of the three manipulations on the two performance measures. The main effect of the performance manipulation on OBJ-PERF was significant, $F(1, 109) = 292.8; p < .001$ ($M = 4.2$ for the low performance and $M = 6.0$ for the high performance). No other main or interaction effect reached significance. Consequently, the objective perceived performance measure was driven exclusively by the actual performance manipulation.

The second ANOVA of the manipulations on SUBJ-PERF showed three main effects and one interaction effect. The performance manipulation showed the strongest impact, $F(1, 109) = 72.6; p < .001$ ($M = 3.3$ for the low-performance manipulation and $M = 5.4$ for the high-performance manipulation), followed by the expectations manipulation, $F(1, 109) = 60.7; p < .001$ ($M = 3.5$ for the low-expectations manipulation and $M = 4.9$ for the high-expectations manipulation), and the needs manipulation, $F(1, 109) = 16.8; p < .001$ ($M = 3.8$ for the low-needs manipulation and $M = 4.8$ for high-needs manipulation). Finally, there was a significant interaction effect between the performance and the needs manipulations, $F(1, 109) = 5.7, p = .02$. An examination of the cell means (see Figure 1) showed that in the high-performance conditions, performance was rated significantly higher when needs were also high (cell mean of 5.5) than when needs were low; cell mean of 4.0; $t(108) = 4.68; p < .001$. This may be due to an increased level of utility associated with this higher level of per-
formance in the high-needs condition. For low-performance conditions, the impact of needs on performance was small and not significant. The cell means were 2.4 in the low-needs condition and 2.9 in the high-needs condition, \( t(108) = 1.27; p > .10 \). It seems that as the level of performance fell below what was needed, needs had less impact on performance ratings.

To summarize, SUBJ-PERF was driven by all manipulations, suggesting that this measure may capture more complex processes beyond simple performance perception. Consequently, these findings imply that SUBJ-PERF did capture some aspect of an evaluation, which included needs and expectations, as well as a need by performance interaction. Because expectations and needs are often used to model evaluation processes, it is not surprising that they are reflected in our subjective performance measure.

We used an experimental design and conducted an initial analysis with the traditional ANOVA. However, given that the conceptualization posited interconnections among the dependent variables (i.e., between perceived performance, disconfirmation, and satisfaction), a structural equation modeling (SEM) approach was used. First, the measures were examined using a Confirmatory Factor Analysis (CFA) and then the hypotheses were tested with a staged SEM approach.

Measurement Model

The covariance matrices were read into LISREL 7 to examine the fit of the proposed models. Before testing the causal relations specified in the conceptual models, the measurement model was assessed using CFA. The constructs of interest were objective and subjective perceived performance, expectations, disconfirmation-of-expectations, needs congruency, and overall satisfaction. The fit of the CFA model is acceptable (\( \chi^2 = 21.6, p = .12 \)), with a goodness-of-fit index of .956 and an adjusted goodness-of-fit of .895. The squared multiple correlations ranged from .634 to .944, thus suggesting an acceptable level of reliability. Similarly, the average variance extracted (Fornell & Larcker, 1981) was adequate (.797 for disconfirmation-of-expectations, .657 for needs congruency, and .921 for satisfaction). These results suggest that most of the variance was captured by the constructs rather than due to measurement error. The CFA also showed that in every case, path coefficients from latent constructs to their corresponding indicators were significant at \( p < .001 \), thus providing evidence of convergent validity (Sujan, Weitz, & Kumar, 1994). Taken together, the results suggest that the data provide a good fit with the hypothesized measurement model.

Perceived Performance–Satisfaction Link: Test of H1

Table 1 contains the correlations of all the overall measures used for hypothesis testing. The correlations show that the objective perceived performance measure (OBJ-PERF) has a lower correlation with satisfaction (\( r = .44 \)) than the subjective measure SUBJ-PERF (\( r = .80 \)). Fisher’s \( r \)-to-\( z \) transformation was used to test the significance of the difference in correlation coefficients. The results confirm that OBJ-PERF is significantly lower correlated with satisfaction than SUBJ-PERF (\( z = 6.62, p < .001 \)), tentatively supporting H1.

Furthermore, SUBJ-PERF shows the highest correlation with SAT amongst all measures, indicating that none of the measured comparison standards (EXP-DISC and NEED-CON) completely mediated the impact from SUBJ-PERF. This observation is consistent with the interpretation of subjective performance measures already capturing parts of satisfaction.

A structural equation approach, using the basic expectations-disconfirmation framework, was used to formally test H1. Overall, satisfaction was conceptualized as a function of disconfirmation-of-expectations. Disconfirmation-of-expectations in turn was determined by perceived performance and prior expectations. A link between expectations to perceived performance was included in the model because expectations can influence performance perceptions via assimilation effects, and empirical studies have supported this link (see Oli-

### Table 1

Correlations Between Key Summary Measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Performance manipulation (MPERF)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Expectations manipulation (MEXPECT)</td>
<td>-.01</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Needs manipulation (MNEEDS)</td>
<td>-.01</td>
<td>-.01</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Objective performance (OBJ-PERF)</td>
<td>.85</td>
<td>.02</td>
<td>-.10</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Subjective performance (SUBJ-PERF)</td>
<td>.52</td>
<td>.47</td>
<td>.26</td>
<td>.44</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Expected performance (EXPECT)</td>
<td>.02</td>
<td>.67</td>
<td>.31</td>
<td>-.11</td>
<td>.51</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Needed performance (NEED)</td>
<td>-.01</td>
<td>.09</td>
<td>.91</td>
<td>-.09</td>
<td>.27</td>
<td>.40</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Disconfirmation-of-expectations (EXP-DISC)</td>
<td>.32</td>
<td>.52</td>
<td>.23</td>
<td>.23</td>
<td>.71</td>
<td>.63</td>
<td>.27</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Needs-congruency (NEED-CON)</td>
<td>.47</td>
<td>.18</td>
<td>.47</td>
<td>.41</td>
<td>.66</td>
<td>.29</td>
<td>.40</td>
<td>.48</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>10. Satisfaction (SAT)</td>
<td>.50</td>
<td>.40</td>
<td>.40</td>
<td>.44</td>
<td>.80</td>
<td>.43</td>
<td>.40</td>
<td>.60</td>
<td>.79</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note. \( 109 < n < 112 \).
ver, 1997). Finally, a direct causal link from perceived performance to satisfaction was included to directly test H1.

The analysis for testing H1 is divided into three sections. We first examine the effects of the traditional expectations-disconfirmation paradigm followed by an analogous model using the needs-congruency construct. Within each framework, both objective and subjective performance measures are fitted separately. In the final section, we present the results for a grand model that incorporates both comparison standards and both performance measures.

Model 1 in Figure 2 shows the standardized path coefficients when the subjective perceived performance measure was used. As indicated by the path coefficients, satisfaction was solely determined by SUBJ-PERF, hence disconfirmation of expectations failed to influence satisfaction evaluations. This finding is consistent with earlier studies using similar subjective perceived performance measures (e.g., Churchill & Surprenant, 1982; Tse & Wilton, 1988).

Model 2 reports the results of the same analysis when perceived performance was measured on an objective scale (OBJ-PERF). This time, satisfaction was largely determined by disconfirmation-of-expectations, whereas perceived performance had only a marginal impact. These findings suggest that disconfirmation-of-expectations mediates much of the effects of expectations and perceived performance on satisfaction. The goodness-of-fit statistics summarized in Table 2 indicate that Model 2 with an objective performance measure had a better fit than Model 1 with the subjective performance measure. Overall, the fit of Model 2 seems adequate. Furthermore, as expected, the standardized path coefficient of perceived performance on satisfaction was lower when an objective performance measure was applied rather than a subjective measure. In other words, employing a more objective measure to capture perceived performance reduced the observed direct relation between perceived performance and satisfaction, thus supporting H1.

Improved Modeling of the Comparison Process: Further Testing of H1

Prior research suggests that needs or desires congruency can be a better predictor of satisfaction than disconfirmation-of-expectations (e.g., Barbeau, 1985; Spreng & Olshavsky, 1992; Spreng et al., 1996). Consequently, as shown in Figure 3, disconfirmation-of-expectations from Figure 2 was replaced with needs congruency as a mediating variable.

The impact of needs congruency as a comparison standard was first tested on a subjective performance measure (Model 3) and then followed by an objective measure (Model 4). As indicated in Figure 3, the subjective measure showed a significant link to overall satisfaction. When the more objective performance measure was used (Model 4), the direct path between objective performance and satisfaction became insignificant. In other words, the effects of objective perceived performance on overall satisfaction were completely mediated by needs-congruency.

These results provide additional support for H1. This time the conceptualization of satisfaction used needs-congruency as a comparison standard. The link between perceived performance and satisfaction was insignificant when an objective performance measure was applied. However, the link between performance and satisfaction remained strong with a subjective performance measure, probably as this evaluative measure already captures part of satisfaction.

Improved Modeling of the Comparison Process: Test of H2

H2 predicted a decrease in the observed direct link from perceived performance to satisfaction, if the modeling of the comparison process is improved. To test H2, the path coefficients from perceived performance to satisfaction are compared between Models 1 and 3 for the subjective performance measure, and Models 2 and 4 for the objective performance measure.

The path coefficients in Model 3 indicate that perceived performance (with a subjective performance measure) had a lower direct impact on satisfaction in the needs-congruency framework.
The Grand Model

A grand model incorporating both comparison standards (expectations-disconfirmation and needs congruency) and both performance measures (subjective and objective) was developed to further advance our understanding of the evaluation process. This model enabled us to examine whether disconfirmation-of-expectation explains variance in satisfaction over and above needs-congruency, thus contributing to the literature on alternative standards. Moreover, the grand model made it possible to investigate the relative influence of the two performance measures within a single framework. Because the results from our prior analyses suggest that the subjective performance measure might already capture part of satisfaction, subjective performance was included as an additional indicator of satisfaction in the grand model (see Figure 4).

Both disconfirmation measures showed a significant link to satisfaction. The path coefficients were .81 \( (p < .001) \) for needs-congruency and .22 \( (p < .05) \) for expectations disconfirmation. The results suggest that needs-congruency is a more important predictor of satisfaction than expectations disconfirmation. Nevertheless, the results also suggest that each of the comparison standards captures distinct aspects of the evaluation process, replicating Spreng et al.’s (1996) findings. As expected, the relation between subjective performance and satisfaction was significant with a path coefficient of .59, \( p < .001 \). Moreover, objective performance had a stronger link to performance (path coefficient = .67) than subjective performance (path coefficient = .30).

The fit of the grand model was relatively poor with \( \chi^2 = 167 \) at 59 degrees of freedom, an adjusted goodness of fit of .729, and a root mean square residual approximation of .130. Several factors might have contributed to the poor fit of the grand model. First, the research design was experimental, and hence, our sample size \( (n = 111) \) might not have been sufficient to obtain trustworthy parameter estimates with the grand model. In fact, the ratio of sample size to number of free parameters (4.7:1) was below recommended levels (see Bentler & Chou, 1987).

Summary Findings

The results of this investigation suggest that the frequently observed direct link from perceived performance to satisfaction may, at least partially, be a function of the type of measures used and of the insufficient modeling of the evaluation process. First, applying a more objective measure of per-
ceived performance greatly reduced the observed direct link between the two constructs. Second, using needs-congruency as a mediating variable resulted in this link becoming insignificant. In sum, the observed link between perceived performance and satisfaction was reduced from a standardized path coefficient of .815 ($p < .001$) in Model 1 (using a subjective perceived performance measure and disconfirmation-of-expectations) to only .075 ($ns$) in Model 4 (using an objective perceived performance measure within the needs-congruency framework).

Finally, the results of this study suggest that needs-congruency should be incorporated into our conceptualization of satisfaction. Needs-congruency had a strong link to overall satisfaction, whereas disconfirmation-of-expectations played a smaller, but still significant role. This finding suggests that consumers use multiple standards in the satisfaction process.

THEORETICAL IMPLICATIONS

Role of Perceived Performance in the Satisfaction Process

Overall, the findings of this investigation suggest that the direct link between perceived performance and satisfaction may reflect a measurement issue rather than accurately mirroring the cognitive processes used by consumers in forming satisfaction evaluations. Subjective scales that already include an evaluative component have typically been used to measure perceived performance. In this investigation, however, perceived performance was measured on both a subjective and a more objective scale. The results indicate that employing a more objective measure greatly reduces the extent of the direct relation between performance and satisfaction. In fact, the link between perceived performance and overall satisfaction was either small or insignificant, regardless of the model, when an objective or value-free perceived performance measure was used. These implications are consistent with prior research. Cadotte et al. (1987) postulated that measures of performance using anchor words (e.g., very good/very bad) might imply evaluation and therefore measure part of satisfaction. Spreng and Olshavsky (1992) speculated that perceived performance should be separated into two distinct constructs: an “evaluationless” judgment and a subjective evaluation.

Desire Congruency and Multiple Standard Models

The results of this study cast some doubt over the use of disconfirmation-of-expectations as the only variable for modeling the comparison process leading to satisfaction. The results are consistent with Spreng and Olshavsky’s (1993), Spreng and Mackoy’s (1996), and Barbeau’s (1985) findings that needs or desires-congruency should be incorporated into our conceptualization of satisfaction. All four studies showed that desire congruence is a better predictor of satisfaction than disconfirmation-of-expectations across a number of products. These findings indicate that values derived from needs and wants may be superior to expectations as a comparison standard across many product categories. However, disconfirmation-of-expectations may be more influential with high involvement purchases or in situations where expectations are strong. Prior research suggests that high involvement may enhance the role of disconfirmation in the evaluation process (e.g., Oliver & Bearden, 1983). Product involvement was not measured in our study, thus leaving this alternative explanation subject to future verification. Moreover, prior expectations for products and services may vary in strength and valence. It can be argued that expectations that are strong and value-laden may be more powerful in the disconfirmation process than expectations that are weak. Consequently, the results of past studies using disconfirmation-of-expectations may be partially attributed to ignoring the moderating role of expectancy strength and valence. Furthermore, other comparison standards based on consumers’ affective reactions or equity perceptions, for example, may perform better under certain circumstances than the needs congruency standard tested in this investigation. These limitations are clearly issues for future research.

In addition to the crucial role of needs as a comparison standard, our findings suggest that multiple standards may be used in the comparison process, replicating the findings of Spreng et al. (1996). Consequently, consumers may under
certain circumstances process expectations and needs separately in the satisfaction evaluation. This reasoning may also apply to other standards (e.g., perceived equity and ideal standards) and further research on the use of multiple standards or different forms of expectations should enhance our understanding of the satisfaction process (e.g., Erevelles & Leavitt, 1992; Tse & Wilton, 1988).

As with most laboratory experiments, interpretation of the results of this study must be moderated by an understanding of how well the experimental tasks parallel natural instances in which consumers form satisfaction judgments. Future work is needed to focus on issues of external validity by verifying that the results can be extended to actual product evaluations in a more natural setting. For example, consumer expectations (e.g., of waiting time via posting “expected wait” as part of an advanced queuing systems) and actual performance (short vs. long duration) could be manipulated in a field experiment. Or, shopper intercept surveys could be used across various products and services, whereby expectations and needs could be measured prior to a consumption experience. Conversely, perceived performance and satisfaction indicators could be collected on an after-consumption basis.

In conclusion, the results of this investigation suggest that the casual relation between satisfaction and perceived performance found in prior studies might mirror a measurement artifact and/or insufficient modeling of the satisfaction process. In this study, the observed direct link from perceived performance to overall satisfaction diminished when more objective perceived performance indicators replaced the traditional value-laden measures. Moreover, our findings indicate that the performance–satisfaction link becomes insignificant when a desire-congruency as opposed to the traditional disconfirmation of expectations is used as a comparison standard in the satisfaction modeling process.

ACKNOWLEDGMENTS

Anna Mattila was visiting the Department of Marketing, National University of Singapore, during the conduct of this study. She is now with the School of Hotel Restaurant and Recreation Management, The Pennsylvania State University.

We gratefully acknowledge the research assistance of Ang Eu Gene and Lee Meng Chung, and the partial financial assistance provided by the National University of Singapore.

REFERENCES


tributions to consumer satisfaction and complaining behavior (pp. 27–32). Bloomington: Indiana University School of Business.


Accepted by Dawn Iacobucci.