

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/247274027>

# Customer-Satisfaction Measurement Performance Counts

Article in *Cornell Hospitality Quarterly* · December 1998

DOI: 10.1177/001088049803900611

---

CITATIONS

75

---

READS

2,454

2 authors, including:



[Atila Yuksel](#)

Adnan Menderes University

84 PUBLICATIONS 1,983 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



sleep quality and tourists' biorythms [View project](#)

# Cornell Hotel and Restaurant Administration Quarterly

<http://cqx.sagepub.com>

---

## **Customer-Satisfaction Measurement: Performance Counts**

Atila Yüksel and Mike Rimmington

*Cornell Hotel and Restaurant Administration Quarterly* 1998; 39; 60

DOI: 10.1177/001088049803900611

The online version of this article can be found at:

<http://cqx.sagepub.com/cgi/content/abstract/39/6/60>

---

Published by:



<http://www.sagepublications.com>

On behalf of:

The Center for Hospitality Research of Cornell University

**Additional services and information for *Cornell Hotel and Restaurant Administration Quarterly* can be found at:**

**Email Alerts:** <http://cqx.sagepub.com/cgi/alerts>

**Subscriptions:** <http://cqx.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

# Customer-Satisfaction Measurement

## Performance Counts

by Atila Yüksel and  
Mike Rimmington

Perhaps the best way to determine customer satisfaction may also be the most simple. Just ask, "How'm I doing?"

**D**etermining customer satisfaction is fundamental to effective delivery of services. Successfully being able to judge customers' satisfaction levels and to apply that knowledge potentially gives a hospitality manager an advantage over competitors via such benefits as product differentiation, increased customer retention, and positive word-of-mouth communication.<sup>1</sup> Given the vital

<sup>1</sup> G.M. Gundersen, M. Heide, and H.U. Olsson, "Hotel Guest Satisfaction Among Business Travelers," *Cornell Hotel and Restaurant Administration Quarterly*, Vol. 37, No. 2 (April 1996), pp. 72-91; L.J. Crompton and L.L. Love, "The Predictive Validity of Alternative Approaches to Evaluating Quality of a Festival," *Journal of Travel Research*, Vol. 34, No. 1 (Summer 1995), pp. 11-25; H. Oh and C.S. Parks, "Customer Satisfaction and Service Quality: A Critical Review of the Literature and Research Implications for the Hospitality Industry," *Hospitality Research Journal*, Vol. 20, No. 3 (1997), pp. 36-64.

**Atila Yüksel**, Ph.D. candidate, is on the faculty of Sheffield Hallam University and Adnan Menderes University. **Mike Rimmington** is M.B.A. program director in the School of Hotel Management at Oxford Brookes University.

© 1998, Cornell University

role of customer satisfaction, one should not be surprised that a great deal of research has been devoted to investigating the process by which customers form judgments about a service experience. As a result, noticeable progress has been achieved in the application of customer-satisfaction information within service industries, despite the fact that perhaps no other area has generated as many theoretical and methodological difficulties.

Attempting to provide a theoretical explanation of the concept, academics have largely focused on conceptual issues and underlying processes, giving less attention to the more pragmatic task of measurement.<sup>2</sup> Although several methods have been developed to provide information for managers, the literature (particularly in the area of hospitality and tourism) lacks empirical studies that compare measurement methods' relative validity and reliability.<sup>3</sup> Given the limited research in this area, we decided to conduct a comparative study scrutinizing the predictive power and reliability of existing methodologies in determining customer satisfaction.

The study described in this paper examined the relative validity of six alternative ways of assessing customer satisfaction. To this end, the first section of the paper provides a brief theoretical background of customer-satisfaction theory together with the reported operational merits and limitations of existing measurement approaches. We then discuss our research methodology, followed by the findings and implications of the study.

### Customer-satisfaction Measurements

Drawing on a number of studies, customer satisfaction can be defined as a post-consumption evaluative

judgment concerning a product or a service.<sup>4</sup> While early research considered the satisfaction construct as a form of attitude, considerable research in the 1980s revealed that satisfaction is notably more complex and that the measurement issues are numerous.<sup>5</sup> Researchers have developed various competing theories of and measurement approaches for satisfaction, most notably the expectancy-disconfirmation paradigm and its variants. Expectancy-disconfirmation, which is a derivative of adaptation-level theory, states that customers compare actual product and service performance with their prior expectations. If expectations are met or exceeded, the consumer is satisfied. If perceived performance falls short of expectations, on the other hand, dissatisfaction results.

Researchers have proposed two basic methods of investigating confirmation and disconfirmation of expectations—an inferred approach and a direct method.<sup>6</sup> The inferred approach involves computing the discrepancy between expectations of performance and evaluation of outcomes. This technique requires researchers to educe separate data sets relating to customer-service expect-

tations and perceived performance. The scores for performance are then subtracted from those of expectations to form the third variable, the confirmation-disconfirmation (or difference) score, which is used in subsequent analysis. The direct approach, by contrast, requires the use of summary-judgment scales to measure confirmation and disconfirmation (e.g., a Likert-type scale of "better than expected" to "worse than expected"). The researcher avoids the necessity of calculating difference scores, since the respondents can be asked directly the extent to which the service experience exceeded, met, or fell short of expectations. Both inferred and direct methods of expectancy-disconfirmation paradigm have been used by hospitality and tourism researchers in various hotel- and tourism-related studies.<sup>7</sup>

Notwithstanding its growing popularity, the expectancy-disconfirmation paradigm has received considerable theoretical and operational criticism—in particular, for including expectations and for using difference scores in assessing customer satisfaction.<sup>8</sup> Some have

<sup>4</sup> For example, see: G.R. Churchill and C. Surprenant, "An Investigation into Determinants of Customer Satisfaction," *Journal of Marketing Research*, Vol. 19 (1982), pp. 491–504; or R.L. Oliver, "A Cognitive Model of the Antecedents of Satisfaction Decisions," *Journal of Marketing Research*, Vol. 17 (1980), pp. 46–49.

<sup>5</sup> R.N. Cooper, B.M. Cooper, and F.D. Duhan, "Measurement Instrument Development Using Two Competing Concepts of Customer Satisfaction," *Journal of Consumer Satisfaction/Dissatisfaction and Complaining Behaviour*, Vol. 2 (1989), pp. 28–35.

<sup>6</sup> See: V. Prakash and W.J. Lounsbury, "A Reliability Problem in the Measurement of Disconfirmation of Expectations," in *Advances in Consumer Research*, Vol. 10, ed. P.R. Bagozzi and M.A. Tybout (Ann Arbor, MI: Association for Consumer Research 1983), pp. 244–249; and A. Meyer and P. Westerbarkey, "Measuring and Managing Hotel Guest Satisfaction," in *Service Quality in Hospitality Organisations*, ed. D.M. Olsen, R. Teare, and E. Gummeson (New York: Cassell, 1996), pp. 185–204.

<sup>7</sup> See, for example: A. Pizam and A. Milman, "Predicting Satisfaction among First-time Visitors to a Destination by Using the Expectancy-Disconfirmation Theory," *International Journal of Hospitality Management*, Vol. 12 (1993), pp. 197–209; J.D. Barsky, "Customer Satisfaction in Hotel Industry: Meaning and Measurement," *Hospitality Research Journal*, Vol. 16 (1992), pp. 51–73; or J.D. Barsky and R. Labagh, "A Strategy for Customer Satisfaction," *Cornell Hotel and Restaurant Administration Quarterly*, Vol. 33, No. 5 (October 1992), pp. 32–40.

<sup>8</sup> Among others, from: J.T. Brown, A.G. Churchill, and P.L. Peter, "Improving the Measurement of Service Quality," *Journal of Retailing*, Vol. 69, No. 1 (1993), pp. 127–139; F. Buttle, "SERVQUAL: Review, Critique, and Research Agenda," *European Journal of Marketing*, Vol. 30, No. 1 (1996), pp. 7–27; Churchill and Surprenant, pp. 491–504; and C. Gronroos, "Toward a Third Phase in Service Quality Research: Challenges and Future Directions," in *Advances in Service Marketing and Management: Research and Practice*, Vol. 2, ed. T.A. Swart, D.E. Bowen, and W.S. Brown (Greenwich, CT: JAI Press, 1993), pp. 49–64.

<sup>2</sup> Gundersen et al., pp. 72–91.

<sup>3</sup> See: Crompton and Love, pp. 11–25; and Oh and Parks, pp. 36–64.

argued, for instance, that if expectations are measured after or even simultaneously with the service experience, those expectations have been biased by the experience.<sup>9</sup> Any expectations thus biased might be understated if the customer has a negative experience, for instance, resulting in a smaller difference than would otherwise be the case, and a positive experience might cause an overstatement of expectations.

Considering this contamination effect, some researchers suggest that expectations should be solicited before the service experience.<sup>10</sup> Getty and Thomson, for instance, argue that “to be of value, expectations should be elicited prior to the service’s being provided; otherwise, the risk is great that expectations will be contaminated by perceptions of the actual service provided.”<sup>11</sup> However, other researchers argue that measuring expectations prior to the service experience is also problematic.<sup>12</sup> For instance, the guest may modify her expectations during the service encounter and use those modified (perhaps more realistic) expectations as the standard of comparison. In her research on tourist satisfaction, for instance, Weber points out that modification in expectations can occur as the impor-

tance attached to pre-trip expectations may change during the trip and a new set of expectations may be formed as a result of multiple experiences during the trip.<sup>13</sup> This implies that as the customer progresses from one encounter to the next, say, from greeting to seating to ordering in the restaurant, the guest may revise his expectations based on the experience of the previous encounter.<sup>14</sup> Events that are completely unanticipated prior to a service encounter, therefore, may become significant contributors to consumers’ overall dissatisfaction or satisfaction.<sup>15</sup>

Additionally, from a practical perspective, Dorfman argues that a difficulty of measuring expectations is that such measurement leads to consistently high expectation ratings.<sup>16</sup> Referring to issues concerning the measurement of service quality, for example, Babakus and Boller point out that “in general, when people are asked to indicate an expected level and an existing level they seldom rate the expected level lower than the existing level.”<sup>17</sup>

That is, respondents may feel motivated to demonstrate an I-have-high-expectations social norm and also indicate that they desire the maximum amount of quality on every attribute. The level of expectations, therefore, may exceed the experienced or existing level for no other reason than guests’ tendency to proclaim high expectations. If that is true, restaurateurs would find

it difficult to satisfy customers, as expectations will never be met or exceeded.<sup>18</sup> In this sense, Crompton and Love argue: “If these scores are almost constant, then there is little point in including them on an instrument, since they will not give responses significantly different from using the perception scores alone.”<sup>19</sup>

Given the basic assumption of the expectancy–disconfirmation paradigm—that a customer must have pre-purchase expectations to be able to experience disconfirmation of those expectations—a number of researchers contend that measuring expectations, and thus disconfirmation, may not be valid in situations where customers do not have well-formed expectations.<sup>20</sup> Lack of experience with a service or lack of familiarity with a destination may cause expectations to be tentative and uncertain.<sup>21</sup> In such situations, considering expectations as a set of firm criteria against which evaluative judgments are made is likely to be fallacious. For example, in the case of a first time trip to eastern Europe, one might assume that expectations based on travel to western Europe would be an appropriate proxy. Such prior experience, though, may have a little relevance to the actual experience owing to the discontinuity of development in eastern and western European nations.<sup>22</sup>

Wider experience, on the other hand, seems to foster more-realistic

<sup>9</sup> For example, see: Gronroos, pp. 49–64.

<sup>10</sup> J.M. Carman, “Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions,” *Journal of Retailing*, Vol. 66, Spring 1990, pp. 35–55.

<sup>11</sup> M.J. Getty and N.K. Thomson, “The Relationship Between Quality, Satisfaction, Recommending Behaviour in Lodging Decisions,” *Journal of Hospitality & Leisure Marketing*, Vol. 2, No. 3 (1994), p. 8.

<sup>12</sup> Among them: P.J. Danaher and J. Mattsson, “Customer Satisfaction During the Service Delivery Process,” *European Journal of Marketing*, Vol. 28, No. 5 (1994), pp. 5–16; Gronroos, pp. 49–64; and D. Iacobucci, A.K. Grayson, A.L. Ostrom, “The Calculus of Service Quality and Customer Satisfaction: Theoretical and Empirical Differentiation and Integration,” in *Advances in Service Marketing and Management: Research and Practice*, Vol. 3, ed. T.A. Swart, D.E. Bowen, and W.S. Brown (Greenwich, CT: JAI Press, 1994), pp. 1–67.

<sup>13</sup> K. Weber, “Assessment of Tourist Satisfaction Using the Expectancy Disconfirmation Theory, A Study of German Travel Market in Australia,” *Pacific Tourism Review*, Vol. 1 (1997), pp. 35–45.

<sup>14</sup> Danaher and Mattsson, pp. 5–16.

<sup>15</sup> Weber, pp. 35–45.

<sup>16</sup> P.W. Dorfman, “Measurement and Meaning of Recreation Satisfaction: A Case Study in Camping,” *Environment and Behaviour*, Vol. 11, No. 4 (1979), pp. 483–510.

<sup>17</sup> E. Babakus and W.G. Boller, “An Empirical Assessment of the SERVQUAL Scale,” *Journal of Business Research*, Vol. 24 (1992), pp. 253–268.

<sup>18</sup> A. Yüksel and M. Rimmington, “An Integrated Approach to Customer Service Evaluation, Advances in Hospitality and Tourism Research,” *Proceedings of the Second Conference on Graduate Education and Graduate Students Research*, Vol. 2 (1997), pp. 31–45.

<sup>19</sup> Crompton and Love, p. 15.

<sup>20</sup> For example, see: D. Halstead, D. Hartman, L.S. Schmidt, “Multisource Effects on the Satisfaction Formation Process,” *Journal of the Academy of Marketing Science*, Vol. 22, No. 2 (1994), pp. 114–129; and Iacobucci *et al.*, pp. 1–67.

<sup>21</sup> Crompton and Love, pp. 11–25.

<sup>22</sup> Halstead *et al.*, pp. 114–129.

expectations. In one study, the researchers noted that people with extensive travel experience developed more-realistic expectations regarding trips and gave greater satisfaction ratings than did people without previous experience.<sup>23</sup>

The role played by expectations in determining the level of satisfaction has also been questioned. In her research on tourist satisfaction, Hughes comments that "even though experiences did not fulfill expectations, a considerable number of tourists were relatively satisfied."<sup>24</sup> Similarly, Pearce also argues that tourists may be satisfied even though their experiences did not fulfill their expectations.<sup>25</sup> In this regard, measuring expectations may be an inefficient way of addressing customer-satisfaction issues.

### Perceived Performance Only

Given the theoretical and operational problems related to the measurement of expectations, many researchers doubt the validity of disparity theories for measuring customer satisfaction and consider that perceived performance is the best predictor.<sup>26</sup>

The limited empirical evidence indicates that the performance dimensions alone predict behavioral intentions at least as well as the complete expectancy-disconfirmation

### Definitions of terms used in this research

Term	Example
<b>Predictive expectations:</b> The service that the consumer anticipates during a visit to a particular restaurant.	<i>The waiting time for a dish in this particular restaurant will be... (short, long).</i>
<b>Perceived performance:</b> The subjective evaluation made by the consumer after a service encounter.	<i>The waiting time for the dish in this particular restaurant was... (short, long).</i>
<b>Inferred disconfirmation</b> (subtractive): The difference between what the customer expected and what the customer received, expressed numerically.	<i>The score of predictive expectations is subtracted from the score for perceived performance.</i>
<b>Direct disconfirmation</b> (subjective): The consumer's evaluation of the restaurant's product performance relative to pre-purchase expectations.	<i>The waiting time for the dish in this particular restaurant is...(much better than expected, much worse than expected).</i>
<b>Importance:</b> The weight or significance that the customer attaches to a restaurant's service attribute.	<i>The waiting time for this dish is... (extremely important, not important at all, some point in between).</i>

model.<sup>27</sup> This suggests that including the confirmation-disconfirmation calculation as an intervening variable is unnecessary. That is, when a service or product performs well, the consumer will be satisfied regardless of any confirmation-disconfirmation effect.<sup>28</sup> Moreover, Halstead *et al.* argued that when customer expectations have become well established, such as in the case of continuously provided services or when there is high familiarity with the service, the confirmation-disconfirmation process will not operate unless performance is clearly outside the range of those norms that are based on experience.<sup>29</sup> Performance bears a preeminent role in the formation of customer satisfaction because it is the main feature of the consumption experience. Perceived performance,

<sup>23</sup> R.A. Westbrook and W. Newman, "An Analysis of Shopper Dissatisfaction for Major Household Appliances," *Journal of Marketing Research*, Vol. 15, (August 1978), pp. 456-466.

<sup>24</sup> K. Hughes, "Tourist Satisfaction: A Guided Tour in North Queensland," *Australian Psychologist*, Vol. 26, No. 3 (1991), p. 168.

<sup>25</sup> P.L. Pearce, "Introduction to the Tourism Psychology," *Australian Psychologist*, Vol. 26, No. 3 (1991), pp. 145-146.

<sup>26</sup> For example, see: Babakus and Boller, pp. 253-268; Brown *et al.*, pp. 127-139; G.M. Engeset and M. Heide, "Managing Hotel Guest Satisfaction: Towards a More Focused Approach," *Tourist Review*, Vol. 2 (1997), pp. 23-33; and S. Erevelles and C. Leavitt, "A Comparison of Current Models of Consumer Satisfaction/Dissatisfaction," *Journal of Consumer Satisfaction/Dissatisfaction and Complaining Behaviour*, Vol. 5 (1992), pp. 104-114.

<sup>27</sup> Brown *et al.*, pp. 127-139; and J. Swan and E.I. Travick, "Satisfaction Explained by Desired versus Predictive Expectations," in *Changing Marketing Environment: New Theories and Applications*, ed. Kenneth L. Bernhardt (Chicago: American Marketing Association, 1981), pp. 170-173.

<sup>28</sup> Erevelles and Leavitt, pp. 104-114.

<sup>29</sup> Halstead *et al.*, pp. 114-129.



**One may not need to  
measure consumers'  
pre-event expectations  
to gauge ultimate  
customer satisfaction.**

therefore, seems to be more straightforward, convenient, and typical of the human cognitive process.<sup>30</sup>

In contrast, some researchers argue that performance ratings alone may not lead to the same practical applications as, or have the diagnostic value of, difference scores.<sup>31</sup> These researchers suggest that the use of difference scores gives managers a better understanding of whether increasing expectations or diminishing performance might be responsible for declining service quality and customer satisfaction.

### **Attribute Importance**

Despite wide acceptance of the expectancy-disconfirmation paradigm as the theoretical construct that best explains satisfaction, many researchers also acknowledge the potential contribution of attribute importance.<sup>32</sup> Without considering attribute importance, one has no indication of the relative importance that respondents attach to particular aspects of a service performance.<sup>33</sup> Barsky and others thus have advocated that attribute importance be used to weight confirmation-disconfirmation scores in assessing customer satisfaction.<sup>34</sup> This implies that customers' satisfaction levels would be related to the strength of their beliefs regarding each attribute's importance multiplied by how well the attributes meet those expectations.

The weighting has been calculated in two ways: indirect inference through a regression model and direct questioning of subjects. For an indirect measurement, researchers

have used a constant-sum scale to derive importance by asking respondents to allocate 100 points among a set of dimensions.<sup>35</sup>

In the case of the direct-measurement method, the most popular weighting method has been to multiply customers' rating of each attribute by its importance, although the underlying assumption of statistical independence of the two variables has not been clearly established.<sup>36</sup>

Statistical independence aside, the multiplication approach has other shortcomings. For example, Crompton and Love argue that this approach is not capable of distinguishing between the relative contribution of the importance and perception scores: "A score of 3 on the importance scale and 5 on the perception scale gives the same score of 15 as does a 5 on importance scale and 3 on the performance scale. Although these two combinations yield the same result, the implications in terms of satisfaction are substantially different."<sup>37</sup> To alleviate such limitations, Barsky suggested assigning a range of numbers to each combination.<sup>38</sup> However, Duke and Persia argue that "such multiplicative models often do not resemble either the original performance ratings or the original importance ratings."<sup>39</sup> Moreover, the multiplicative approaches are argued to be of little help to managers, pressed for time and without access to sophisticated software.

In this regard, scholars continue to discuss whether to include im-

<sup>30</sup> Meyer and Westerbarkey, pp. 185-204.

<sup>31</sup> A. Parasuraman, V.A. Zeithaml, and L.L. Berry, "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Future Research," *Journal of Marketing*, Vol. 58 (1994), pp. 111-124.

<sup>32</sup> Among them: Barsky, pp. 51-73.

<sup>33</sup> As noted by: Carman, pp. 35-55.

<sup>34</sup> See: Barsky and Labagh, pp. 32-40.

<sup>35</sup> A. Parasuraman, V.A. Zeithaml, and L.L. Berry, "SERVQUAL: a Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing*, Vol. 64, Spring 1988, pp. 12-40.

<sup>36</sup> Oh and Parks, pp. 36-64.

<sup>37</sup> Crompton and Love, p. 14.

<sup>38</sup> Barsky, pp. 51-73.

<sup>39</sup> C.R. Duke and A.M. Persia, "Performance-Importance Analysis of Escorted Tour Evaluations," *Journal of Travel and Tourism Marketing*, Vol. 5, No. 3 (1996), p. 208.

portance in measuring attitudes. Oh and Parks, for instance, commented: "Those who advocate inclusion mainly focus on the conceptual and realistic role of the relative importance of different attributes in human-decision processes, whereas those who dismiss inclusion tend to emphasize statistical and methodological efficiency."<sup>40</sup> Given that debate, we believe that the literature dealing with customer satisfaction needs empirical studies to assess the contribution of attribute importance to the various models' predictive power.

### Satisfaction Measures Compared

The perfect rating scale almost certainly does not exist, but some scales are more likely to produce reliable and valid findings than others.<sup>41</sup> In this study we compared six forms of customer satisfaction—performance only; performance weighted by importance; importance minus performance; direct confirmation–disconfirmation; confirmation–disconfirmation weighted by importance; and performance minus predictive expectations. Given that performance is the closest match with the human cognitive process and given its ease of application, this study hypothesized that (1) a performance-only approach is a more satisfactory framework for measuring customer satisfaction than the other five alternatives, and (2) weighting performance and direct confirmation–disconfirmation scores by importance would not make a substantial improvement on the predictive validity of these methodologies.

We assessed which methodology is most appropriate by examining the construct validity of the several

scales under investigation. Construct validity is concerned with the extent to which a particular measure relates to other measures consistent with theoretically derived hypotheses.<sup>42</sup> As it is most directly related to the question of what the measurement is in fact measuring, construct validity lies at the heart of measurement validity. For a scale to have construct validity, it must satisfy such conceptual and empirical criteria as convergent, discriminant, and nomological validity.

**Convergent validity.** Convergent validity of a measure is the extent to which the measure correlates or converges with other measures designed to measure the same concept—indicating that a particular variable value is not an accident.

**Discriminant validity.** Another theory-based measurement is discriminant analysis. For measures to have discriminant validity, the correlation between two different measures of the same variable should be higher than the correlation between the measure of that variable and the measures of any other variable.<sup>43</sup>

**Nomological validity.** Nomological validity of a measure is the extent to which the measure correlates in a theoretically predicted way with a measure of a different-but-related construct.<sup>44</sup>

To help assess convergent, discriminant, and nomological validity within each scale, as well as the relative validity of the methodologies, we used an overall customer-satisfaction measure and two behavioral-intentions measures.<sup>45</sup>

<sup>42</sup> G.E. Carmines and A.R. Zeller, *Reliability and Validity Assessment* (London: Sage Publications, 1979).

<sup>43</sup> J.J. Cronin, Jr., and S.A. Taylor, "Measuring Service Quality: A Re-examination and Extension," *Journal of Marketing*, Vol. 56 (1992), pp. 55–68.

<sup>44</sup> Brown *et al.*, pp. 127–139.

<sup>45</sup> As suggested, for example, by: Cronin and Taylor, pp. 55–68; and Crompton and Love, pp. 11–25.

### The Research Instrument

The study used four different types of questionnaires. The first questionnaire was designed to assess only respondents' performance perceptions. The second questionnaire assessed expectations and performance evaluations. The third investigated attribute importance and perceptions of performance. The final questionnaire was designed to investigate a summated judgmental scale and attribute importance.

Each of the four research instruments comprised the same set of 24 questions accompanied by four different scales, according to the customer-service measure that the questionnaire was testing. The questions were divided into the following areas: general information about the purpose of visit, ratings on 12 restaurant-service attributes, ratings on overall service evaluation, behavioral intentions (i.e., would they return and would they recommend), and a demographic profile. The 12 service-attribute items on the questionnaire were derived from previous restaurant studies and were based on seven-point semantic-differential scales. The study also used a single overall measure of customer satisfaction. Although some researchers contend that satisfaction should be measured by a combination of attributes, ease of use and empirical support for a summative overall measure of satisfaction led us to use that approach.<sup>46</sup> The study employed a seven-point "delight-terrible" scale for measuring customer satisfaction, as it has been reported to be the most reliable customer-satisfaction scale.<sup>47</sup> To ensure that the question-

<sup>46</sup> As proposed by: D. Halstead, "Expectations and Disconfirmation Beliefs as Predictors of CS, Repurchase Intentions, and Complaining Behaviour: An Empirical Study," *Journal of Consumer Satisfaction/Dissatisfaction and Complaining Behaviour*, Vol. 2 (1989), pp. 17–21.

<sup>47</sup> N.R. Maddox, "Measuring Satisfaction with Tourism," *Journal of Travel Research*, Winter 1985, pp. 2–5.

<sup>40</sup> Oh and Parks, p. 53.

<sup>41</sup> J.S. Dewlin, K.H. Dong, and M. Brown, "Selecting a Scale for Measuring Quality," *Marketing Research: A Magazine of Management and Applications*, Vol. 5, No. 3 (1993), pp. 12–17.



## Exhibit 1

### Results of reliability, multiple-regression, and correlation analyses

Model [n]	Coefficient alpha	Multiple-regression values	Correlation with...		
			Overall satisfaction	Return intentions	Recommending restaurant
Performance only [106]	.6287	.73905	.5792	.3662	.4418
Performance weighted by importance	.7713	.63844	.6861	.6144	.5842
Performance component of the performance-weighted-by-importance scale [105]	.7517	.65573	.6981	.6136	.5777
Importance minus performance	.7836	.52880	-.6157	-.5460	-.4391
Confirmation–disconfirmation [102]	.9136	.52616	.4591	.2015	.2942
Confirmation–disconfirmation weighted by importance	.9150	.51140	.4456	.2019	.2937
Performance minus predictive expectations [88]	.6162	.30693	.3868	.3170	.2345
Expectation	.7349	.60020	.3328	.3816	.4725
Performance component of the performance-minus-predictive-expectations scale	.7882	.68432	.6132	.5992	.6084

naires' design, wording, and measurement scales were appropriate, we ran a pilot test with 30 restaurant customers. The pilot test revealed that there was no major problem related to the clarity and appropriateness of the items included in the questionnaire, although we did make some minor changes.

**Sampling.** The study setting of this research was a chain restaurant. The survey was carried out with a sample of 460 restaurant customers (115 for each instrument) during a two-week period in June 1997. Forty customers refused to participate and 19 of the returned questionnaires were incomplete, resulting in an *N* of 401. Fifty-seven percent of the respondents were women, and 76 percent of all respondents had visited the restaurant at least once before.

To build the sample we handed out a questionnaire to every seventh customer on Sundays and to every fifth customer on Wednesdays and Thursdays. We distributed the questionnaires at the door as diners entered the restaurant and collected the questionnaires as those diners departed.

We rotated distribution of the questionnaires. Though it might seem preferable to obtain comparative information from the same respondents (and thus eliminate between-subject variability), asking identical questions using four different scales would have introduced a severe respondent-fatigue effect.

On the inferred-type questionnaire, we asked respondents to answer questions about their expectations before they ate and about perceptions afterward. We acknowledge Weber's concern that this approach may be subject to "hindsight" bias—particularly if the respondent waited until after dinner to complete the expectations section.<sup>48</sup>

<sup>48</sup> Weber, pp. 35–45.

We had to accept this possibility, however, and we tried to account for it in our interpretation of results (as discussed later).

In the administration of other questionnaires, the respondents were instructed both to evaluate the actual service and state the importance of the service attributes after their meal.

## Results and Discussion

After analyzing the data using SPSS software, we subtracted scores on expectations from scores on performance perceptions to yield difference scores. Similarly, we constructed weighted confirmation–disconfirmation and weighted–performance scores by computing respondents' scores on performance and item importance.<sup>49</sup>

To assess the possibility of sample bias, we followed Childress and Crompton's suggestion of conducting a series of chi-square tests on demographic and visit-related variables (i.e., age, sex, purpose, previous visit, dining alone or accompanied).<sup>50</sup> No significant differences were found between those variables, suggesting that differences found in the subsequent analysis could reasonably be attributed to reasons other than the sample differences.

We also performed a series of one-way ANOVA and t-tests to compare data from the four different instruments in terms of respondents' overall satisfaction, likeliness to return, and likeliness to recommend. We found no significant difference ( $p > 0.05$ ) among the respondents' ratings on those factors. Similar t-tests, however, revealed significant differences ( $p < 0.05$ ) in

the ratings of first-time patrons and repeat visitors on overall satisfaction, likeliness to return, and likeliness to recommend. Repeat visitors gave the restaurant significantly higher scores than first-timers on those variables. Finally, we subjected the collected data to reliability analysis to assess the internal consistency of the set of items.<sup>51</sup> The total scale reliability for all of the methodologies was moderately high, which indicates that the sample of the items performed well in capturing the measured construct (Exhibit 1).

Using Pearson product-moment correlation and multiple-regression procedures, we investigated the four methodologies' convergent, discriminant, and nomological validity (hence, their construct validity). The Pearson scores were computed for each of the satisfaction–calculation methods, using respondents' scores on the aforementioned seven-point scale (from delighted to terrible) for their responses to the question, "Overall how satisfied are you with the meal experience?" We also calculated the correlations between each model's outcome and the respondents' intention to return and their likeliness to recommend. In addition to the correlation analysis, we used multiple regression to assess the ability of each of the six scales to explain variation in customer satisfaction. For each of the 12 attributes in the regression analysis, we applied a backward stepwise procedure to eliminate the variable with the largest probability-of-F value, provided that value was larger than 0.10.

Our correlation analysis found several statistical relationships that were significant at the  $p < 0.01$  level. The result of correlation analysis of overall satisfaction with

the performance-only scale indicated a relatively high, statistically significant correlation between the two measures (.5792). As shown in Exhibit 1, the following scales revealed statistically significant high correlations: overall satisfaction and the performance component of the performance-weighted-by-importance scale (.6981), overall satisfaction and performance weighted by importance (.6861), and overall satisfaction and importance minus performance (–.6157). The analysis showed a moderately high, statistically significant correlation between overall satisfaction and confirmation–disconfirmation (.4591) and overall satisfaction and confirmation–disconfirmation weighted by importance (.4456). The performance component of the performance-minus-predictive-expectations scale also showed a high, significant correlation (.6132) with overall satisfaction, although the correlation between overall satisfaction and the composite performance-minus-predictive-expectations scale was moderate (.3868). Similarly, the results of the stepwise multiple regression indicate that the performance-only method has a higher correlation with and a higher ability to explain more of the variation in overall satisfaction than did the other methodologies. In addition, some degree of discriminant validity for the models is evident in Exhibit 1, as correlation with overall satisfaction was higher than with other correlations.

The next step was to check each construct's nomological validity by seeing how well a particular construct's measure related to measures of the other constructs to which it was theoretically related. That is, did each customer-satisfaction construct correlate in the way that the theory predicted with a measure of a different-but-related construct? An examination

<sup>49</sup> Based in part on: Barsky and Labagh, pp. 32–40.

<sup>50</sup> D.R. Childress and L.J. Crompton, "A Comparison of Alternative Direct and Discrepancy Approaches to Measuring Quality of Performance at a Festival," *Journal of Travel Research*, Fall 1997, pp. 43–57.

<sup>51</sup> A.G. Churchill, Jr., "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, Vol. 16, February 1979, pp. 64–73.

**If customers are nevertheless  
satisfied despite the fact  
that the service experience  
did not fulfill their initial  
expectations...**

of the correlation coefficients in Exhibit 1 indicates nomological validity for the performance model, since it correlated in the theoretically predicted way with intention to return and intention to recommend. The nomological validity of the performance-weighted-by-importance and importance-minus-performance scales is also supported, for the correspondence between the models and behavioral intentions was as predicted. With regard to the confirmation-disconfirmation and confirmation-disconfirmation-weighted-by-importance scales, the correlation coefficients demonstrated moderately low correspondence between those models and the two behavioral-intention variables.

An important issue bears on the nomological validity of the performance-minus-predictive-expectations scale, however. The correlation coefficient demonstrates a relatively low correlation between "return intentions (performance minus predictive expectations)" and "recommend (performance minus predictive expectations)." An examination of correlation coefficients further demonstrates that the perception component of the difference model alone performs better than the performance-minus-predictive-expectations scale, while the expectations component has relatively small correlation with measures of other theoretically related constructs. The perception component outperforms performance minus predictive expectations .59 to .31 in predicting the intention to return and .60 to .23 in predicting intention to recommend. This outcome, consistent with others' findings, suggests that although the inferred method of expectancy-disconfirmation paradigm is intuitively appealing, the calculated difference scores do not

provide additional information in predicting behavioral intentions.<sup>52</sup>

We undertook additional analysis to examine the contribution of the expectations component of the performance-minus-predictive-expectations scale in determining customer satisfaction. That is, we wanted to see whether the complete performance-minus-predictive-expectations scale provides any additional information beyond that already contained in the perceptions component of that scale. If the conceptualization of customer satisfaction as a difference score is indeed the most valid measure, the correlation between the difference scores and overall satisfaction must be greater than that of the perceptions component alone with satisfaction.<sup>53</sup>

Contrary to that expectation, actual inspection shows that the correlation between the overall performance-minus-predictive-expectations scale and overall satisfaction is weaker (.3868) than that between the perceptions component and overall satisfaction (.6132). The dominant component in the difference score is clearly the perception of performance. This result suggests the efficacy of using only performance perceptions to measure customer satisfaction.

Thus, the first of our hypotheses, concerning the higher convergent validity and predictive power of the performance-only model in predicting customer satisfaction, has been supported by correlation coefficients and the results of multiple regression. The second proposition, in which inclusion of importance scores was assumed not to make any substantial difference in the predictive power of the models, has also

<sup>52</sup> See: Brown *et al.*, pp. 127-139; and Cronin and Taylor, pp. 55-68.

<sup>53</sup> As suggested by: Babakus and Boller, pp. 253-268.

been justified, because weighting performance or disconfirmation by importance did not increase the correlation and multiple-regression values.

Although weighting according to importance appeared to offer no substantial improvement to the reliability or validity of either the performance-weighted-by-importance model or the confirmation-disconfirmation-weighted-by-importance model, plotting the scores of importance and performance on coordinate axes has the advantage of being easily interpreted by managers.<sup>54</sup> That facet could be of critical value for managers who do not use sophisticated software packages. The weighting approach could yield valuable information for managers as they strive to allocate resources according to the service aspects customers consider important.

As we already stated, people rarely indicate lower expectations for performance than what they considered actually occurred. That is, "how much was expected" is usually higher than "how much there is now."<sup>55</sup> This phenomenon was demonstrated in the performance-minus-predictive-expectations scale. The mean score of the expectations (.8639, SD 0.6) was higher than the mean perceptions score (.5919, SD 0.67).<sup>56</sup> A similar tendency (scores on expectations are consistently higher than those for performance) is found in the results of Fick and Ritchie and of Parasuraman *et al.*<sup>57</sup> This finding suggests that it would be difficult

to satisfy customers; their expectations will never be met or exceeded. To the contrary, however, the fact that we found no significant difference in overall level of customer satisfaction between the samples calls into question the applicability of this phenomenon in marketing.

Defining customer satisfaction as exceeding expectations lends itself to a peculiar effort to manipulate expectations. Some researchers, for instance, have asserted that service providers should understate their firm's capabilities—in effect, promising less than they plan to deliver. For example, Pizam and Milman suggested that "it would be more beneficial to create modest and even below-realistic expectations."<sup>58</sup> Though this is theoretically a sensible and potentially effective suggestion, we question whether one could or should actually try it.<sup>59</sup> To begin with, customers may not want to spend time and money patronizing a restaurant that conveys a weak message of customer benefits. Moreover, establishing a threshold at which expectations are raised high enough to attract customers, but low enough to allow for expectations to be exceeded is obviously difficult, if possible at all.

### Unexpected Satisfaction

More important, our findings indicate that customers may be reasonably satisfied even if the restaurant does not totally meet their expectations. A cross-tabulation of mean scores for inferred confirmation-disconfirmation (or, performance-minus-expectation-confirmation or disconfirmation) with respondents' ratings on overall satisfaction brought out this conclusion. Although mean scores for inferred confirmation-disconfirmation (or,

**...then the assumption that customers are performance dissatisfied when performance falls short of expectations is likely wrong.**

<sup>54</sup> Crompton and Love, pp. 11–25.

<sup>55</sup> See: Babakus and Boller, pp. 253–268; and Dorfman, pp. 483–510.

<sup>56</sup> Statistically significant at  $p < 0.05$ .

<sup>57</sup> G.R. Fick and B.J.R. Ritchie, "Measuring Service Quality in the Travel and Tourism Industry," *Journal of Travel Research*, Fall 1991, pp. 2–9; and Parasuraman *et al.*, pp. 12–40.

<sup>58</sup> Pizam and Milman, p. 208.

<sup>59</sup> As did: Weber, pp. 35–45.



performance-minus-expectation confirmation or disconfirmation) show that respondents' expectations were not met (performance was lower than expectations), a considerable number of respondents ( $n = 42$ ) were relatively satisfied and well over half were either slightly satisfied or pleased. Thus, customers may be reasonably satisfied despite the fact that the service experience did not fulfill their initial expectations. Indeed, this finding brings into question the assumption of the expectancy-disconfirmation paradigm that customers are perforce dissatisfied when performance falls short of expectations.

**Logical flaw.** Another potential problem relates to the logic of calculating customer satisfaction by subtracting predictive expectations from a performance score. The approach presupposes that everyone has prior expectations about the service experience. Confirmation or disconfirmation of expectations cannot occur without such advance expectations. However, prior expectations may not be established clearly enough to serve as a basis for a comparison to an experience. Indeed, as we administered the questionnaire that measured performance minus predictive expectations, a number of respondents indicated that they had no expectations for this restaurant. Certainly, one could not calculate performance minus expectations for that group. The issue of fuzzy expectations casts some doubt over the logic and appropriateness of using the performance-minus-predictive-expectations scale for assessing customer satisfaction.

A further shortcoming of the performance-minus-predictive-expectations scale relates to its administration. As we mentioned above in our discussion of procedure, measuring expectations and

perceptions simultaneously is indefensible, because the expectations are unavoidably colored by the experience. Our resulting effort to solicit expectation scores prior to service (and thereby eliminate the risk of contaminated expectations) proved to be much more difficult and prone to bias than we had hoped, particularly because of some respondents' failure to complete the expectations part before eating.

Another substantial limitation of the use of difference scores is that the difference scores may result in misleading interpretations resulting from simple arithmetic.<sup>60</sup> For example, a seven-point scale for expectations and performance yields six ways of producing a score of -1 for performance minus expectation (e.g.,  $P = 1 - E = 2$ ;  $P = 6 - E = 7$ ) and seven ways of producing a score of zero. One could argue that no matter how the score of -1 is achieved the customer found performance to be slightly less than expectations. Each case, however, indicates a slightly different combination of customer expectations and restaurant performance. The logic of the expectancy-disconfirmation paradigm predicts otherwise, however.

By the same token, the diagnostic ability of the direct confirmation-disconfirmation-only scale is also questionable. This scale cannot indicate whether the expectations being confirmed or disconfirmed were high or low. Intuitively speaking, the case of meeting or exceeding low customer expectations is qualitatively different from meeting or exceeding high expectations. Therefore, the direct confirmation-disconfirmation scale seems to be of little use for diagnostic analysis.

<sup>60</sup> As noted by: R.K. Teas, "Consumer Expectations and the Measurement of Perceived Service Quality," *Journal of Professional Services Marketing*, Vol. 8, No. 2 (1993), pp. 35-54.

## How Are You Doing?

Hospitality and tourism managers are supposed to be responsible for guest satisfaction, but they often find little guidance on how to determine whether they have delivered a satisfactory service experience. Accurate measurement of customer satisfaction is a prerequisite to developing effective management strategies. This study attempted to ascertain the relative validity and reliability of various methodologies for measuring customer satisfaction. We found what appeared to be significant differences in the validity of those methodologies.

Measuring customer satisfaction as performance only emerged as the most reliable and valid measure of satisfaction, whereas deriving a score for customer satisfaction as the difference between customers' ratings of predictive expectations and perceived performance was the least reliable. Weighting the findings according to importance of restaurant attributes made little improvement on the predictive power of the measures. We caution, however, that although weighting importance does not add to the explanatory power of the models, we are not recommending that the importance dimension be discarded. Knowing the importance ascribed to service attributes by different individuals may still be useful for managers.

As a final note, in this research we used customers' predictive expectations about the service experience as a comparison standard. Future research needs to be carried out on the use of other types of expectations such as normative or ideal expectations and minimum expectations in measuring customer satisfaction. Given that customers reported being satisfied even when the restaurant did not fulfill their expectations, an examination of minimum expectations may be particularly fruitful. **CQ**