The Quality of Guest Comment Cards

An Empirical Study of U.S. Lodging Chains

Kenneth R. Bartkus  
Utah State University

Roy D. Howell  
Texas Tech University

Stacey Barlow Hills  
Utah State University

Jeanette Blackham  
Sinclair Oil Corporation

This study examines the quality of guest comment cards used by major U.S. lodging chains. To accomplish this objective, guidelines for comment card design were developed through a review of the relevant literature. The guidelines focus on eight issues: (1) return methods, (2) introductory statements, (3) contact information, (4) number of questions, (5) space for open comments, (6) number of response categories for closed-ended questions, (7) balanced versus unbalanced response categories for closed-ended questions, and (8) question wording. Using a sample of 63 lodging chains, the most common deviations from guidelines include a lack of secure return methods, the use of positively biased response categories, and insufficient writing space for open comments. To improve the quality of comment card feedback, these and other limitations should be corrected. Managerial implications and directions for future research are included.

Keywords: comment cards; lodging; guidelines; guest feedback; quality

What does it mean if a finding is significant, or that the ultimate statistical analytical techniques have been applied, if the data collection instrument generated invalid data at the outset?

Jacoby 1978, p. 90

There is little doubt that customer feedback is an important component of the service management equation in the travel and tourism industry. The significance of feedback has been reported in such diverse areas as theme parks (Chen 2003), airlines (Jonas 2002), cruise lines (Jainchill 2008; Travel Weekly 2005), lodging (Sampson 1996), business tours (Home 2005), business to business relationships in tourism (Weissmann 2003), tourist complaint behaviors (Gursoy, Ekiz, and Chi 2007; Schoefer and Ennew 2004; Zins 2002), and rural tourism (Opperman 1995). One industry expert even referred to customer feedback as the “breakfast of champions” (Gage, reported in Travel Weekly: The Choice of Travel Professionals 2006, p. 18).

Although companies typically use more than one customer feedback mechanism (e.g., a combination of formal satisfaction surveys, focus groups, Web sites, personal interviews, and/or toll-free numbers), one of the most ubiquitous is the comment card. These relatively short pencil-and-paper questionnaires allow customers to complete and return them at their convenience. In this sense, they represent the traditional comment card (which is distinguished from the recent introduction of electronic comment cards and other short surveys that are distributed through a company’s Web site or through an e-mail message; Litvan and Kar 2001; Tierney 2000). Traditional comment cards also differ from more formal feedback mechanisms such as customer satisfaction surveys in that they are passively solicited (Sampson 1996).

The popularity of the comment card method can be attributed to its ability to provide regular, timely feedback at, or near, the time of service. Prasad (2003, p. 17), for example, maintains that comment cards represent a “tactical information tool for immediate problem solving and for monitoring service delivery quality.” In this
sense, comment cards help identify critical incidents (both good and bad) and serve to enhance the quality of service management (Scriabina and Fomichov 2005).

Unfortunately, some research has suggested that the quality of comment cards is often flawed. Gilbert and Horsnell (1998) examined U.K. hotels and found that the amount of writing space varied considerably from a single line to a full page. Wisner and Corney (1997) surveyed buffet restaurants and found that 18% of the cards contained 6 inches or less of writing space and one card had a total of only 2 inches. They also found that only 41% of the cards provided a secure return mechanism (e.g., locked collection box or postage return). Steintrager (2001) conducted an informal survey of fast-food restaurants and found that many of the cards contained insufficient space to write even a brief comment.

Are guest comment cards as flawed as the evidence seems to suggest? Perhaps, but the aforementioned research is largely exploratory and limited. First, two of the studies relied on small convenience samples in the restaurant sector. Specifically, both Wisner and Corney (1997) and Steintrager (2001) each sampled just 22 restaurants. As such, their samples are probably too small to be meaningful, and, perhaps more important, they are not related specifically to the travel and tourism sector.

Second, while Gilbert and Horsnell (1998) utilized a cross-sectional sample of lodging establishments, they based their evaluation on guidelines drawn from formal customer satisfaction surveys. However, because response rates are low (Sampson 1996; Trice and Layman 1984) and because comment cards are subject to self-selection bias (Barksy 2001), they probably should not be used as a substitute for formal guest satisfaction surveys. Hence, the Gilbert and Horsnell study did not take into account the different objectives of guest comment cards, and, as a result, the guidelines may not have been suitable.

Since the validity of any customer feedback mechanism is contingent on whether or not the data collection instrument is properly designed, a better understanding of the quality of guest comment cards is needed. At a minimum, this involves two issues: (1) the development of guidelines that are appropriate for an evaluation of guest comment cards and (2) the use of a suitable sample.

With regard to issue 1, prior research has focused on such issues as general guidelines for formal questionnaire design (e.g., Baker 2003; Webb 2000), guidelines for specific distribution formats, such as Web-based questionnaire design (e.g., Deutskens et al. 2004; Healey, Macpherson, and Kuijten 2005), and recommendations for developing questionnaires that measure a specific outcome, such as satisfaction with a cruise line (Testa, Williams, and Pietrzak 1998). Additional research has also examined how questionnaire methods influence response outcomes (Cole 2005; Crompton and Tian-Cole 2001; Lankford and Buzton 1995). We could not, however, identify a single study that outlined a set of guidelines designed specifically for comment cards. Given the popularity of comment cards, this represents an important gap in the literature. We propose to address this issue by developing comment card guidelines adapted from a review of the established literature on questionnaire design and administration.

Issue 2 is addressed though a sample of comment cards obtained from the lodging industry. Lodging was selected for the analysis because it represents a major segment of the economy and has been examined in a multitude of travel research studies (e.g., Al-Sabbahy, Ekinci, and Riley 2004; Chen and Schwartz 2008; McIntosh and Sigs 2005; Oh, Fiore, and Jeoung 2007; Phillips and Louvieris 2005; Poria 2006; Sheehan, Ritchie, and Hudson 2007; Siguaw, Enz, and Namasiyayam 2000; Thrane 2005; Yucelt and Marcella 1996). We begin with a review of the scholarly literature as it pertains to the usefulness and design quality of comment cards. A set of formal guidelines for assessing comment card quality is then developed. After presenting the method of inquiry and analysis, we present the results and discuss their implications for service management in travel and tourism.

**Literature Review**

Scholarly research involving comment cards has focused on two primary issues: (1) usefulness and (2) design quality. We address each of these issues below.

**The Usefulness of Comment Cards**

The usefulness of comment card data has been widely examined. Much of the focus has been on the usefulness of the comments themselves. Manickas and Shea (1997), for example, conducted a content analysis of customer complaints at one property and found that resolution of service failures resulted in less satisfaction from guests than resolution of equipment failures. Shea and Roberts (1998) examined “guest room logs” (comment books) and found that virtually all of the remarks were positive, contradicting claims from others that soliciting comments asks people to be only negative (cf. Millete, as reported in Cebrzynski 2003, p. 66).

Pullman, McGuire, and Cleveland (2005, p. 342) examined comment cards from one hotel to demonstrate how qualitative comments from guests can be converted to quantitative data while at the same time “maintaining
the respondents’ meaning so that researchers are clear on the respondents’ intent.” In this way, they argued, management does not have to guess what the respondents actually mean when they rate their satisfaction 7 out of 10.

Banker, Potter, and Srinivasan (2005) examined the relationship between guest satisfaction (as measured by responses on comments cards) and subsequent changes in revenue and operating profit for a single hotel chain. Perhaps not surprising, they found that improvements in satisfaction were followed shortly by increases in revenue and profit.

Aside from an analysis of comments, other researchers have examined some limitations of the comment card method of feedback. Sampson (1996), for example, reported evidence of self-selection bias, and Trice and Layman (1984) reported relatively low response rates. Each of these issues has been subsequently debated in the literature. Barksy (2001) and Perlik (2002), for example, argued that self-selection bias and low response rates mean comment card data are not valid because they cannot be generalized to the population of guests. Alternatively, Prasad (2003, p. 17) maintained that this would only be true if the purpose of comments cards was to generalize. He contends that comment cards only “serve as a tactical information tool for immediate problem solving and for monitoring service quality and delivery.” In this sense, it has been proposed that comment cards serve a useful purpose for guests who may wish to avoid complaining in person (Meyer, as reported in Ruggless 1994).

In sum, the literature has provided two major perspectives on comment card usefulness. On one hand, it has shown that comments can provide useful and important information that can be used to improve service quality and performance. On the other hand, there continues to be some disagreement regarding the role of comment cards, whether they should be interpreted as formal survey instruments or merely as a timely feedback mechanism for immediate problem solving. Overall, the literature suggests that although comment cards can provide useful information, they should not be used as a substitute for formal satisfaction surveys. Instead, they should be regarded as an opportunity (along with phone calls, personal contact, e-mails, etc.) to receive timely feedback from customers at or near the time of experience.

The Quality of Comment Card Design

Although empirical evidence regarding the design quality of comments cards remains limited, some important insights have been provided from a few preliminary studies. One of these relates to the use of return mechanisms. Wisner and Corney (1997), for example, found that while 55% of the cards in their survey provided for a secure return option (e.g., mail or “drop box”), 23% provided no instructions at all and another 22% asked the guest to return the card to a staff member (i.e., cashier or front desk). Steintrager (2001) also examined return mechanisms and found that “some” of the cards in her sample required customers to pay for return postage.

A second insight relates to the amount of space allocated for writing comments. In the Wisner and Corney (1997) study, 18% of the cards had 6 or fewer inches of writing space. In the Steintrager (2001) study, she reported a lack of sufficient writing space on many of the cards, although no further information on the meaning of “sufficient” was provided.

A third insight has involved the issue of question wording. Wisner and Corney (1997) found 23% of the cards in their sample contained at least one double question. One question asked, “Provide your assessment of the courtesy and welcome extended by the host persons and cashier on arrival and while being seated” (p. 114, emphasis added). Gilbert and Horsnell (1998) also found evidence of double questions in their study of 45 hotels in the United Kingdom. They reported an average of 1.2 double questions per card.

A fourth insight has been the length of comment cards, as measured by the number of questions. Gilbert and Horsnell (1998) found that approximately 73% had less than the recommended minimum of 40. It is important to note, however, that this standard was adopted from guidelines established for formal satisfaction surveys and, as a result, is not directly applicable to comment cards.

In sum, while prior research has provided a useful foundation and some needed insights toward improving our understanding of the quality of comment cards, it appears that the scope of the analyses has been limited. The following section addresses this issue by proposing a formal set of guidelines for the analysis of comment card quality.

Development of Guidelines

A review of the literature reveals eight issues that need to be addressed in the design of a comment card: (1) return method and statement of confidentiality, (2) introductory statement, (3) contact information, (4) number of questions, (5) space for open comments, (6) balanced response categories, (7) number of response categories, and (8) question wording.
Guideline 1: Comment Cards Should Have a Secure Return Method and a Statement of Confidentiality

Industry observers have argued that the integrity of comment cards is enhanced when the return method is secure (Barksy 2001; Cawley 1998; Prasad 2003). The importance of a secure response mechanism is grounded in the notion that it helps ensure that the information will not be tampered with by any individuals other than those for whom it was intended.

To meet this guideline, a number of options are available: (1) a postage-paid return to management and/or (2) a locked drop box in a convenient and readily accessible location (e.g., the hotel lobby). In addition, Lewis and Morris (1987) argue that cards should not be returned to the front desk because they might be discarded. The suggested alternative is to hand the card directly to the general manager (or guest relations officer).

Although a major advantage of mail is that it provides for added security, it extends the time lag between the delivery of the comment and any subsequent action. To help overcome this dilemma, it would make sense for comment cards to be designed so that guests have the option of returning the card to the “general manager,” returning it to a locked drop box, and/or placing it in the mail. With regard to mail returns, the literature is clear that the card should be addressed and include postage (Yammarino, Skinner, and Childers 1991).

Related to the issue of a secure response mechanism is the concept of confidentiality. Scholars in the area of research methods have noted that it is common practice to include a statement of confidentiality, especially when the identity of the respondent is known or can be easily identified (Hair, Bush, Ortinau 2003).

Adapting this principle to comment cards, a statement of confidentiality assures respondents that any other identifying information will not be used in conjunction with the responses. Since guests are often asked for identifying information, the use of a confidentiality statement helps ensure that guests will be more likely to be candid with their comments.

Guideline 2: Comment Cards Should Have an Appropriate Introductory Statement

Survey protocol suggests that questionnaires should include a short introductory statement explaining their importance. Hair, Bush, Ortinau (2003, p. 468), for example, note, “One or two sentences must be included in any cover letters to describe the general nature or topic of the survey and emphasize its importance.” Although an introductory statement should probably not exceed two sentences, this represents the ideal rather than absolute standard. A less rigid, but perhaps more reasonable, guideline would be an introductory statement of less than four sentences.

When writing an introductory statement, Parasuraman (1986) maintains that they should be concise, be objective, and not lead the respondent to answer in a particular way. One way to measure this is through concepts adapted from the impression management literature. In particular, ingratiation (flattering statements about the guest), self-promotion (boasting about the lodging establishment’s expertise), and exemplification (statements about the lodging establishment’s dedication to superior service) are examples of tactics that can serve to bias guest impressions.

Jones and Pittman (1982) note these tactics can influence perceptions of likeability, competence, and dedication, respectively. As such, they have the potential to influence perceptions of performance and satisfaction. In this regard, ingratiation would be evidenced by the presence of flattering statements about the guest. For example, a statement such as “You are the most important person to us” would be regarded as ingratiation since it extends beyond a mere thank you and conveys a stronger impression. We believe that merely thanking a guest for staying at the hotel or for completing the questionnaire is simply standard protocol and not a form of impression management.

Self-promotion is defined as a statement that attests to the expertise of the lodging organization. Hypothetical claims such as “Our award-winning staff is here to help with your stay” would be regarded as self-promotion. Finally, exemplification is defined as any statement that conveys the image that the hotel is dedicated to providing quality service. A statement such as “We work hard so that you may enjoy your stay” would be an example of exemplification.

While any or all of these tactics might be appropriate in promotional brochures, we question their use in a survey instrument where they can bias impressions. Given that their use would not be accepted with any other form of surveying, their presence is viewed as compromising the quality of comment cards and, as such, should be avoided.

Guideline 3: The Card Should Provide an Opportunity for Guests to Provide Contact Information

The issue of contact information is critical when the purpose of the feedback is to assist in service recovery or employee evaluation. With regard to service recovery,
contact information allows the organization an opportunity to personally respond to the guest’s ideas, concerns, suggestions, or other comments. It also allows the organization to validate the authenticity of the card, which is particularly important when the feedback is used for performance evaluations. In the absence of such validation, the potential for employees (or others) to submit false data is heightened.

While there is no known guideline for what is appropriate, it is reasonable to expect that a comment card should include space for the guest’s name and room number. Additional information such as date of stay and any other contact information (phone number, physical address, e-mail address, etc.) could also be useful, particularly if the guest asks for a response. In such cases, it may make sense to ask for an e-mail address since it can be the most efficient means of responding.

The standard for formal surveys is to ask for contact information at the end of the questionnaire (Alreck and Settle 1995; Parasuraman 1986). It seems reasonable to expect the same of comment cards (Prasad 2003).

**Guideline 4: The Number of Questions Should Be Moderate**

A review of the literature on survey design reveals no formal or established guidelines specifying the appropriate range of questions that should be included on a comment card. However, some scholars have noted that shorter questionnaires are generally preferred to longer ones (Greer, Chuchinprakarn, Seshadri 2000). Sampson (1996) accentuates this position by arguing that since comment cards represent a passive data collection technique that relies almost completely on consumer initiative, a survey of much length will adversely affect the response rate. In this regard, Deutskens et al. (2004) found in their study of Web-based surveys that short surveys generate a higher response rate.

In support of these positions, Prasad (2003) has argued that a range of 5 to 10 closed-ended questions is appropriate along with one general open-ended question. While 5 to 10 questions may be insufficient in some situations, the range appears reasonable for most cases. Alternatively, a range of 50 to 60, as suggested by Gilbert and Horsnell (1998), seems excessive.

Finally, the use of a single open-ended question seems reasonable in most circumstances, although a combination of questions could also be supported. For example, there could be questions addressing general comments, suggestions for improvements, problems that were encountered, and things the guest found particularly enjoyable.

In addition, a hotel property might be interested in hearing ideas for improvement or what additional services guests might like to see added. Thus, a reasonable case can be made for more than one open-ended question, although this should be done with consideration of space limitations. Given that comment cards, by their nature, are not multipaged documents, it seems reasonable to conclude that by increasing the number of open-ended questions, the amount of available space for each question will decrease. Hence, the number of open questions should be moderate.

**Guideline 5: There Should Be Sufficient Space for Open-ended Comments**

Although specific guidelines for what constitutes “sufficient” have yet to be provided, Steintrager (2001) has argued that comments cards often have too little space, and Davies (2006) has argued that the space should be “sufficiently large.” Neither of these recommendations, however, provides any additional guidance.

One proposal would be to set the minimum space at no less than half the side of a standard-sized post card and, ideally, include space equal to or greater than one side. Figure 1 provides an illustration with six lines at 3.5 inches each for a total space of 21 inches. Although arbitrary, this guideline has some intuitive appeal as a minimum standard and acknowledges that additional space is preferred. That is, information quality is more likely to be hampered by too little, rather than too much, space. Webb (2000), for example, has noted that when more lines are offered to respondents, they tend to write more. The question that has yet to be adequately addressed in the literature is how much space is optimal.

A final issue concerns the appropriate location for open-ended questions. Gilbert and Horsnell (1998) suggest
that the questions should be placed at, or near, the end of
the comment card. Lehman (1985, p. 131) also suggests
that open-ended questions should be restricted to the end
of questionnaires to help “pick up any ideas the respond-
ent holds or provide the respondent with a chance to voice
his/her opinion on a subject of interest.” Others have sug-
gested they be located at both the front and back end of the
card. Johnson and Sieveking (1974, p. 776), for example,
found that when open-ended questions are placed before
and after multiple-choice questions, there were significantly
more responses than either placement alone. Conversely,
when only one location was used, the one positioned before
the multiple-choice questions was found to elicit more
“discrete ideas and response categories.”

Given that the presumed purpose of guest comment
cards is to elicit open-ended responses, it is reasonable to
suggest that these questions be positioned near the begin-
ing of the comment card (i.e., after the introduction). A
second open-ended section can also be included near the
end to address other issues that the respondent feels were
not covered.

**Guideline 6: Response Categories for Closed-
ended Questions Should Be Balanced**

To maintain objectivity, scholars in the area of
research methods have long maintained that response
categories for closed-ended questions should be bal-
anced so that there are an equal number of positive and
negative categories (Hair, Bush, and Ortinau 2003;
Lehman 1985; Zikmund 1989). This is true, of course,
unless the responses are expected to be distributed at one
end of the scale (Parasuraman 1986; Zikmund 1989). For
example, if one were to sample only brand-loyal custom-
ners, a positive balance for the response categories would
make sense since it is unlikely these people would have
a negative opinion.

A review of the hotel literature reveals, however, that
some industry observers believe that comment cards
encourage guests to be only negative. For example,
Millete (as reported in Cebrzynski 2003, p. 66) contends
that by using comment cards “you’re only asking people
to be negative.” Alternatively, Goughnour argues that
responses are mostly bimodal: “You only get the people
who really like you, and those who had a really bad
experience” (as reported in Nation’s Restaurant Review
2002, p. 4). In light of these opinions, we are led to con-
clude that response categories should be balanced.

An example of balanced response categories would be,

<table>
<thead>
<tr>
<th>Very good</th>
<th>Good</th>
<th>Average</th>
<th>Poor</th>
<th>Very poor</th>
</tr>
</thead>
</table>

The options are balanced because there are two positive
and two negative options. Note that the use of *average* is
simply a neutral option and not a determining factor in
whether the response categories are balanced. This is
because the deciding issue is whether or not there are an
equal number of positive and negative categories on
either side of the neutral option.

An example of negatively balanced response would be,

<table>
<thead>
<tr>
<th>Good</th>
<th>Average</th>
<th>Poor</th>
<th>Very poor</th>
<th>Awful</th>
</tr>
</thead>
</table>

The response categories are unbalanced because three
are in the negative direction and only one is in the posi-
tive direction. To balance the response categories, “very
good” and “excellent” would need to be added to the left
side, resulting in a 7-point response option.

An example of a scale that is unbalanced in the posi-
tive direction is,

<table>
<thead>
<tr>
<th>Excellent</th>
<th>Very good</th>
<th>Good</th>
<th>Average</th>
<th>Poor</th>
</tr>
</thead>
</table>

The options are not balanced because three are worded
in a positive direction and only one is worded in a negative
direction. To balance the response categories, “very poor
and “awful” would need to be added on the right side.

Finally, the use of both “average” and “fair” within
the same set of response options would not be appropri-
since they both imply a relatively neutral position.

**Guideline 7: There Should Be an Adequate
Number of Response Categories**

The conventional wisdom among measurement theo-
rists is that there should be an adequate number of
response categories. Too few categories (e.g., two or
three) may not be enough to capture enough meaningful
variance, and too may increase variance without a
corresponding increase in precision (Friedman and
Amoo 1999).

While there is no consensus on what constitutes an
adequate number, Lehman and Hulbert (1972) have rec-
ommended a range of between 5 and 7, while Cox (1980)
has recommended a range of 5 to 9. Friedman and
Friedman (1986) found that in some situations an 11-point
response range may produce more valid results, and, as
such, they recommend a range of 5 to 11 points. To reflect
these recommendations, we believe that fewer than 5
points is insufficient and more than 11 is excessive.

Critical to this guideline is whether or not there should
be a neutral category. Scholars in research methods have
shown that using a neutral category can significantly
increase the number of neutral responses (Kalton, Roberts, and Holt 1980; Presser and Schumann 1980). Others have argued that eliminating the neutral position might force a respondent to commit to a particular direction when he or she is actually neutral about the issue (Tull and Hawkins 1993). To address this issue, we adopt the recommendation of Sudman and Bradburn (1982), who argue that a neutral position should be included unless a persuasive argument can be made to the contrary. Since we cannot present such an argument, we argue that a neutral category is appropriate.

When using a neutral category, it has been suggested that a graded position between the two directions be used (Hernández, González-Romá, Drasgow 2004). For example, when a guest is asked to rate the quality of services, a middle point might be “average” or “fair” rather than “neither better nor worse.”

Guideline 8: Question Wording Should Conform to Generally Accepted Principles

Wording issues in question design have been studied for over 50 years (e.g., Baker 2003; Brennan 1997; Dillman 1978; Hall and Roggenbuck 2002; Hunt, Sparkman, and Wilcox 1982; Johnson, Bristow, and Schneider 2004; Payne 1951; Stynes and White 2006). While there is some variance in the way these authors describe wording errors, we find that Hunt, Sparkman, and Wilcox (1982) have provided reasonable recommendations based on Payne’s (1951) seminal treatise The Art of Asking Questions.

First, questions should not be double-barreled. Double-barreled questions are those that include two issues within a single question. For example, the question “How do you rate our restaurant and food?” is double-barreled because it asks the respondent to provide a single evaluation for two categories. Second, questions should not use ambiguous terms; that is, they should not use terms that have multiple meanings. For example, “How do you rate the room amenities?” could be considered ambiguous because an amenity can mean quite different things to different people.

Third, questions should use appropriate vocabulary. Appropriate terms are those that have the potential to be well understood by all respondents. As such, colloquial phrases or industry jargon should be avoided. For example, asking respondents to evaluate the staff’s “can-do” attitude is probably not appropriate (which can also be considered “self-promotion”), nor is a question asking if the guest was pleased with the “Athenian décor” of the room (which may also be considered ambiguous to some respondents).

Fourth, questions should be neither loaded nor leading. Both loaded and leading questions can serve to cue respondents to answer in a particular direction and, as a result, contribute to response bias. An example of a loaded questions would be, “Given our room rates are the lowest of any of our competitors, how would you rate the value of your stay?” It is loaded because it premises the question with a favorable statement. An example of a leading question would be, “Don’t you agree that our beds are comfortable?” It is leading because it encourages the respondent to answer in a positive direction. The literature has been clear in its prescription that leading and loaded questioning should be avoided.

With regard to the subject of open-ended questions, research has largely focused on the issue of wording neutrality. Sampson (1996, 616), for example, argued that complaint solicitations might encourage customers to “try to think up things to complain about when they otherwise would register not complaint.” Along a similar line, Brennan (1997) notes that since the tone of the question can adversely influence the direction of the response, a neutral or combined (i.e., balanced) format should be used.

Examples of neutral formats would be “Please comment” or “We welcome your comments.” An example of a combined cue would be something such as “Please comment on any experience you had with us that was particularly pleasing or displeasing.” In doing so, the researcher is cuing the respondent to comment from either direction (i.e., positive or negative). Questions that asked for only negative or positive responses are more likely to result in biased feedback and should be avoided.

We next set out to examine how well comment cards conform to these proposed guidelines.

Method

Selection of Sample

The sample for this study consisted of major U.S. lodging chains. Using a list provided by the Cornell School of Hotel Administration Web site (http://www.hoteleschool.cornell.edu), we were able to identify a total of 81 lodging chains. Given the manageable size of the population, we decided to conduct a census rather than a sample. We contacted all 81 and found that 75 (or approximately 93%) used comment cards. An examination of the cards revealed redundancies for 12 cases; that is, they used the identical card. These occurred with establishments that were part of a family of chains. After
accounting for this redundancy, we were left with 63 different comment cards. After the cards were collected, they were evaluated by the researchers using the guidelines established for this study.

Evaluation Procedures

Evaluations for the majority of the guidelines were straightforward and not prone to subjective interpretation. For example, whether or not a comment card has a secure return mechanism was evident from explicit directions given on the card. Nonetheless, when evaluating guidelines with objective information, two or more of the researchers evaluated the cards in tandem so as to minimize the potential for information loss or unintentional bias.

In cases where the evaluation involved a more subjective analysis, such as the case with wording errors and impression management, two of the evaluators were given instructions on how to evaluate these guidelines. They then proceeded to code the data on an individual basis. The use of this multiple-coder procedure is particularly important when evaluations require a potentially more subjective analysis, even in cases where established instructions are provided. The purpose is to determine if the same results would be obtained if different, but equivalently trained, evaluators assessed the same data.

With regard to the guidelines for wording errors and impression management, the evaluators were given coding instructions included in the guidelines. The two evaluators then coded the cards individually and compared results. Where discrepancies occurred, the evaluators discussed the issue (sometimes involving a third trained evaluator) to resolve the difference. Altogether, the interrater consistency was greater than 95%.

Analysis and Results

Guideline 1: Return Method and Statement of Confidentiality

Table 1 presents the results for the analysis of return methods. As the table indicates, the majority of cards (52.4%) do not incorporate a secure response mechanism as proposed in the guidelines. The most common return method is simply to leave the card at the front desk (31.7%). Only 28.6% of the cards contained an appropriate secure option (defined as including either the option of a locked drop box or a mail option, or both). Cards that offered a mail option, but did not include postage, were considered only marginally acceptable since a lack of postage would hamper a response.

Guideline 2: Introductory Statements

The majority of cards (60.3%) contained fewer than four sentences, as recommended in the guidelines, and 76.2% had five or fewer. Somewhat surprising, there was a long tail in the distribution with one card containing a total of 14 sentences. Approximately 14% of the cards contained no statement whatsoever other than an identification such as “comment card.”

With regard to acknowledgements, the results indicate that a majority of cards thanked the guest for staying at the hotel (i.e., 50.8%) and 39.7% thanked the guest for completing the comment card. Only one card (1.6% of the sample) provided an assurance of confidentiality, which read, “This information will be used solely for the purpose of ensuring the authenticity of this survey.”

The extent to which impression management tactics were used is presented in Table 2. Of the cards, 28 (approximately 44%) contained no evidence of impression management. The most common example of impression management was the use of exemplification, which was identified on 29 of the cards (46%). Representative
examples include “We have only one job, to provide you with the best” and “We are committed to quality service.”

Ingratiation was found on eight cards (or approximately 13%). Statements such as “You are the most important person” and “We are honored to serve you” were considered ingratiation. Self-promotion was identified on only one card: “We welcome you with genuine hospitality and offer our exclusive ‘yes we can’ service attitude.”

While the intent of these statements may have been genuine and benign, the effect may be to unnecessarily influence impressions and create bias in subsequent guest responses. As such, they are not appropriate on survey instruments.

**Guideline 3: Contact Information**

Three issues were addressed: (1) missing contact information, (2) location of contact information, and (3) whether or not there was an offer to respond. The results are presented in Table 3 and indicate that 41.3% of the cards contained all of the relevant contact information. The most common omission from the remaining cards was not requesting an e-mail address. With regard to location, 82.5% placed the contact information at the end of the comment card, as recommended in the guidelines. Only 15.9% of the cards offered to respond to the guests’ comments.

**Guideline 4: Number of Questions**

Approximately 24% of the cards contained no closed-ended questions at all, reflecting what can be described as a “pure” comment card. At the other extreme, 35% of the cards exceeded the recommended upper limit of 10 closed-ended questions. One card contained more than 30 questions (i.e., 56 questions).

The analysis also revealed that 47.6% of the comment cards were consistent with the recommended guideline of one open-ended question. This means that a slight majority of the cards contained multiple open-ended items with nearly 13.0% including three or more. In the most extreme case, a card contained open-ended space after each closed-ended question (13 in all). The majority of open-ended questions were located at the end of the card (approximately 83.0%), which is in agreement with standard survey protocol but differs from the recommended guidelines for comment cards.

**Guideline 5: Space for Open-ended Items**

The task of measuring the space was complicated by the fact that many cards contained multiple open-ended questions. In addition, some cards did not include a general comment section and instead focused on specific issues (e.g., the quality of the room). In other cases, the open-ended item most relevant to a general comment reflected things that the establishment could do to *improve* the quality of service. In one case, the only open-ended question was in reference to an exceptional employee.
To address these issues, we decided to adopt a standard of reasonableness. In the most basic case, an item that simply read “comment” or something similar was obviously included in the analysis. However, in one case, we found that an item addressing “problems” contained much more space than the item reflecting “comments” (the available spaces were 43 and 4 inches, respectively). In this case, we included the more lengthy space, rationalizing that if a guest wished to express an opinion, he or she might naturally use the area with the most writing space. As such, we used the longest available space when multiple open-ended questions were presented. In the vast majority of cases, however, the amount of writing space between alternative open-ended questions was marginal, meaning that the selection of the longer space did not introduce significant bias in the analysis.

Overall, slightly less than 40% of the cards met the recommended minimum standard of 21 inches of writing space. Almost 16% of the cards provided less than 11 inches of space, suggesting a significant deviation from the recommended guideline. Alternatively, 19% of the cards provided space that significantly exceeded the minimum guideline of 21 inches, with one card providing more than 45 inches of space.

**Guideline 6: Balanced versus Unbalanced Response Categories for Closed-ended Questions**

Table 4 presents the results of the analysis of balanced versus unbalanced response options. Only 29.2% of the response options are balanced. The remaining 70.8% are unbalanced, with all but one skewed in a positive direction.

It is important to note that two cards used a 5-point scale with the following response categories: excellent, good, average, fair, poor. The response categories were rated as unbalanced because “fair” and “average” are considered middle (or neutral) categories, meaning that there were two positive categories (i.e., excellent and good) and only one negative category (i.e., poor).

The use of semantic differential scaling can alleviate the difficulty of having to label each individual category (and the potential for unbalanced categories) but can still create bias if the anchor points do not reflect relevant opposite positions. For example, of the five examples, one was unbalanced because it used “excellent” at one end and “needs improvement” at the other. To be considered balanced, the anchor “needs improvement” should be replaced with “poor.”

**Table 4**

<table>
<thead>
<tr>
<th>Balanced and Unbalanced Response Categories</th>
<th># Cards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unbalanced, 3-point responses</td>
<td></td>
</tr>
<tr>
<td>Excellent–good–poor</td>
<td>1</td>
</tr>
<tr>
<td>Excellent–good–fair</td>
<td>2</td>
</tr>
<tr>
<td>Completely satisfied–satisfied–not satisfied</td>
<td>1</td>
</tr>
<tr>
<td>Exceeded expectations–met expectations–did not meet expectations</td>
<td>4</td>
</tr>
<tr>
<td>Very satisfied–satisfied–not satisfied</td>
<td>3</td>
</tr>
<tr>
<td>Better than expected–as expected–below expectations</td>
<td>1</td>
</tr>
<tr>
<td>Excellent–satisfactory–unsatisfactory</td>
<td>1</td>
</tr>
<tr>
<td>Unbalanced, 4-point responses</td>
<td></td>
</tr>
<tr>
<td>Excellent–good–fair–poor</td>
<td>9</td>
</tr>
<tr>
<td>Poor–fair–good–excellent</td>
<td>2</td>
</tr>
<tr>
<td>Unbalanced, 5-point responses</td>
<td></td>
</tr>
<tr>
<td>Excellent–very good–fair–poor</td>
<td>3</td>
</tr>
<tr>
<td>Excellent–good–average–fair–poor</td>
<td>2</td>
</tr>
<tr>
<td>Unbalanced, 6-point responses</td>
<td></td>
</tr>
<tr>
<td>Superior–very good–satisfactory–fair</td>
<td>1</td>
</tr>
<tr>
<td>Average–outstanding (6-point anchor)</td>
<td>1</td>
</tr>
<tr>
<td>Balanced (4- and 5-point scales)</td>
<td></td>
</tr>
<tr>
<td>Excellent–above average–needs improvements–poor</td>
<td>1</td>
</tr>
<tr>
<td>Very satisfied–somewhat satisfied–neither satisfied not dissatisfied–somewhat dissatisfied–very dissatisfied</td>
<td>2</td>
</tr>
<tr>
<td>Extremely satisfied–satisfied–neither satisfied nor dissatisfied–dissatisfied–extremely dissatisfied</td>
<td>1</td>
</tr>
<tr>
<td>Excellent–above average–average–needs improvement–poor</td>
<td>3</td>
</tr>
<tr>
<td>Semantic differential (balanced, except as noted)</td>
<td></td>
</tr>
<tr>
<td>Unsatisfactory to outstanding (5-point anchor)</td>
<td>2</td>
</tr>
<tr>
<td>Extremely satisfied to extremely unsatisfied (10-point anchor)</td>
<td>1</td>
</tr>
<tr>
<td>Low to high (7-point anchor)</td>
<td>1</td>
</tr>
<tr>
<td>Poor to excellent (7-point with neutral midpoint)</td>
<td>1</td>
</tr>
<tr>
<td>Excellent to poor (5-point anchor)</td>
<td>1</td>
</tr>
<tr>
<td>Excellent to needs improvement (4-point, unbalanced)</td>
<td>1</td>
</tr>
<tr>
<td>Graphic (balanced, except as noted)</td>
<td></td>
</tr>
<tr>
<td>Smile–neutral–frown</td>
<td>1</td>
</tr>
<tr>
<td>Smile–neutral–frown–big frown (unbalanced, negative)</td>
<td>1</td>
</tr>
</tbody>
</table>

**Guideline 7: Number of Response Points for Closed-ended Questions**

Table 5 presents the results of the analysis for the number of response points. As the table indicates, 52.1% of the cards used less than the recommended minimum of five response points. More than 25% of the cards utilized only three response points.
Guideline 8: Question Wording for Closed- and Open-ended Questions

Table 6 provides representative examples of wording errors for closed-ended questions. The results indicate that the most common errors involve the use of double-barreled questions. These were identified in 33 of the 48 cards that used closed-ended questions (68.8%). Ambiguous terms were found in 12 of the cards (25.0%). Inappropriate terms were identified in only 5 of the 48 cards. Leading questions were identified on only one comment card.

The analysis of wording for open-ended items revealed considerable variance. A representative sample of items is presented in Table 7. As the table reveals, 17 of the cards (approximately 27% of total) used a combination format as the primary open question. In some cases, the question was relatively lengthy. In others, it consisted simply of a series of cues (e.g., comments, suggestions, complaints).

Neutral wording was used on 26 of the cards (approximately 41%). These reflect the use of cues as simple as “comments” or phrases such as “your comments are appreciated.”

Negative cues were identified on a total of 13 cards (20.6%), and positive cues were identified on a total of nine cards (14.3%). In all, 11 cards (17.5%) included an additional question on how to improve the experience and/or what additional services or amenities the guest would like to see added, and 21 cards (33.3%) contained a question asking the guest to identify an exceptional employee. In one case, that was the only open-ended question on the card.

Discussion

This study is the first to systematically examine the quality of guest comment cards using guidelines expressly...
developed for comment card design. While the overall results indicate a variety of quality concerns, we believe that most, if not all, of these can be easily addressed.

First, comment cards should include a secure return mechanism. It was somewhat surprising to find more than half of the cards lacked at least one of the preferred options (i.e., a locked drop box or postage-paid mail). A locked drop box provides a combination of security and the ability to respond in a more timely manner. Because of this, we believe it is the preferred option.

Second, while most of the introductory statements were concise, many contained wording that can potentially bias guest impressions. Exemplifying statements such as “We are committed to quality service” and “Making your stay a complete success is our goal” were the most common followed by ingratiating statements such as “You are the most important person” and “We are honored to serve you.” These are expected in promotional brochures but violate the objectivity standard for survey research (Parasuraman 1986; Sobel 1984). As such, they should be avoided.

About half of the cards expressly thanked guests for staying at the property, but fewer than half of the cards thanked them for completing the comment card. Both of these are standard protocol in survey research, and there is little justification for their omission on comment cards. Perhaps of greater concern is the fact that only one card assured confidentiality. This is contrary to the results of Sobel (1984), who found that 45% of the introductory statements in his sample of formal survey instruments promised confidentiality. It may be, however, that because so many cards utilized nonsecure return mechanisms (e.g., front desk), it would have been difficult to include a promise of confidentiality. Nonetheless, there were many cards that did include a secure return mechanism, and only one made this promise.

A third issue of concern is the lack of an e-mail request in the contact information section. Given the increased importance that customers are placing on Internet access (Dunn and Gonzalez 2007), not asking for an e-mail address represents a missed opportunity. That is especially true if management actually intends to respond to the guest’s comments. Unfortunately, only 16% of the cards included an offer to respond. Again, this represents a missed opportunity that can be easily addressed.

Fourth, there appears to be an overuse of closed-ended questions in many of the cards. While the guideline suggests less than 11 questions, 35% of the cards that did use closed-ended questions exceeded this limit, with one card containing more than 50. The use of lengthy questionnaires serves to only reduce the probability that guests will complete the card and return it to management. We believe that cards should strive to limit closed-ended questions to the conservative side of the guideline (i.e., fewer than 5).

A fifth concern is the relative lack of writing space for open comments. The fact that the majority of cards had less than minimum of 21 inches of space and that almost 16% had less than 11 inches raises a question of whether they should even be called comment cards. We argue that the primary objective of comment cards is to solicit written comments, and, as such, a lack of adequate space represents a significant limitation and one for which there is little justification.

A sixth issue concerns question wording. Consistent with prior research in other areas (i.e., Wisner and Corney 1997), we found the use of double questioning to be relatively common. It may be that space restrictions contribute to the use of double questions. Nonetheless, double questions violate wording standards and should be avoided. To address this issue, we argue that questions should reflect a general nature, such as, “How would you rate your stay?” Subsequent open-ended questions can then offer the guest an opportunity to explain their rating.

Seventh, the validity of closed-ended questions was further compromised by the use of unbalanced response options that were positively biased. This was unexpected given that some industry observers have noted that comment cards serve to provide guests only with an opportunity to complain (i.e., Millete, as reported in Cebrzynski 2003). Validity is also compromised by the fact that almost 30% of the cards had fewer than four response points for each question.

We would like to conclude by placing the results of this study in perspective. First, the analysis was restricted to major lodging chains operating within the United States. It may be that comment cards from smaller chains and independents will have different characteristics. Future research would want to examine this issue. An investigation in the global arena could provide additional insights with regard to language and culture. It is also important to note that the use of comment cards extends beyond the lodging. An analysis that incorporates comment cards from other areas of travel and tourism might provide additional insights.

Other issues related to comment card design should also be considered in future research. For example, are some paper weights and colors more appropriate than others? Also, is there an optimal size for the physical card? Finally, should there be guidelines for style and size of letter font?

Future research might also want to extend this research to an evaluation of electronic comment cards. It would be
interesting to evaluate not only the quality of electronic cards but also how they differ from traditional cards. A comparison between the two formats within the same organization might also provide important insights.

We would also like to note that while the foundational support for many of the guidelines is relatively strong (e.g., avoid double questions), other guidelines relied on more subjective interpretations (e.g., appropriate amount of writing space). As such, the guidelines should be interpreted as they are intended, as a reference to improve the quality of comment cards. To the extent that a designer has a logical reason to deviate from the guidelines, this could be considered acceptable. In the end, however, the guidelines provide a basis for accountability, which we consider a worthwhile objective.

This study examined the quality of comment card design and not the quality of responses or how they are used by management. Although these issues have already been addressed by others, we believe more work is needed. We would argue that future research is needed to investigate how deviations from the proposed guidelines influence responses. For many of the guidelines, the answers have already been established (e.g., the use of doubled-barreled questions, the use of unbalanced response categories).

For others, additional research is needed. For example, does the use of ingratiating, self-promotion, and/or exemplification tactics bias comment card results? Similarly, how much space is optimal on a comment card? Does the lack of a secure (and/or confidential) return mechanism inhibit response rates? How does open-ended wording on questions influence responses?

Finally, and perhaps most important, how does management view comment card feedback, and how does management respond to it? Some preliminary research has already been conducted, and the results are not all that encouraging. Specifically, Steintrager (2001) wanted to determine whether or not the restaurants in her study took the information on comment cards seriously. She completed and returned each card in her sample with the comment, “It took longer to get what I ordered than I would have liked. I don’t usually complain, but I thought you should know” and included a name, a return address, and a phone number. After 18 days, the author had received one response. Wisner and Corney (1997) conducted a similar experiment where they wrote “Please call me for suggestions” on the back of 22 comment cards. They received only three calls. Each of these studies suggests that responding to feedback from comment cards may not be a high priority. As such, additional inquiry into the cause of this inattention would be helpful.

References


Gursoy, D., E. Ekiz, and C. Chi (2007). “Impacts of Organizational Responses on Complainants’ Justice Perceptions and...


Kenneth R. Bartkus is a professor of marketing and the executive director of The Research Group in the Jon M. Huntsman School of Business at Utah State University.

Roy D. Howell is the J. B. Hoskins professor of marketing in the Rawls College of Business at Texas Tech University.

Stacey Barlow Hills is a clinical assistant professor of marketing and codirector of the Huntsman Scholars Program in the Jon M. Huntsman School of Business at Utah State University.

Jeanette Blackham is a marketing brand integration coordinator at Sinclair Oil Corporation in Salt Lake City.