The Influence of Information Valence and Managerial Control on Consumers’ Inferences Across Service Providers

by Anja Reimer and Valerie Folkes

Services often have multiple components, each of which are performed by different units of service providers. Previous research in marketing offers little insight into how consumers make inferences across such categories of service providers. Our research examines the effects of information about one service unit (e.g., flight attendants) on quality perceptions of another service unit (e.g., baggage handlers) when both are offered by the same firm. Results suggest that consumers’ inferences are influenced by the valence of information about a service unit, with positive information having more impact on inferences about another unit and about the firm’s employees in general than negative information. The majority of study participants explained their inferences by drawing on beliefs about the firm’s managerial control over employees, which increased the likelihood of participants drawing similar inferences about service quality over different service units.

1. Generalizations from One Service Unit’s Quality to That of Another

Services often have multiple components, each of which are performed by a different type of service provider. For example, pilots, flight-attendants, and baggage handlers all work for an airline but provide different services. When service companies offer a variety of services performed by different employee units, consumers may use information about one group of service providers to draw inferences about the quality offered by another group of service providers. For example, consumers might experience excellent service from an airline’s flight attendants that sets expectations for the quality of the airline’s baggage delivery. It is important to understand to what extent consumers generalize from a service group and what mechanisms influence their inferences.

Previous research in services marketing assumes generalization, because the experience with one service provider influences the customer’s overall service quality judgment. That overall judgment then has the potential to set expectations for service delivered by other employees (e.g., Boulding et al. 1993). However, some factors appear to limit generalizations. For example, consumers generalize more from positive information about an individual service provider than from negative information (Folkes/Patrick 2003). If information valence has asymmetric effects, then service firms need to make more efforts to manage some types of information than others. Whether or not the same positivity effect will be found when consumers generalize from a whole group of service providers instead of from just an individual is unclear.

Another factor that seems to limit generalizations is suggested by the literature on brand extensions (e.g., Aaker/Keller 1990; Boush/Loken 1991; Joiner/Loken 1998; Park/Milberg/Lawson 1991; Völckner/Sattler 2006). This research largely examines when consumers generalize across products with identical brands rather than services delivered by the same firm. Findings show that consumers are more likely to generalize from one group of products to another group of products when both groups are perceived as similar. These results suggest that segregation of service providers into dissimilar specialty units will decrease the likelihood of a consumer’s generalizing from one service component to another, because each service unit would be perceived as dissimilar from the other. Yet, services have special properties...
that may give rise to unique but impactful beliefs about quality that may lead to inferences across even dissimilar services.

One of those special properties involves beliefs about managers’ ability to control employees’ behaviors. Since services are often closely linked to the people that deliver them, consumer beliefs about relationships among people may influence their inferences about service quality. Even when service providers offer very different services, consumers may assume that these separate units deliver similar quality because managers control employees’ behaviors. That control provides a basis for perceiving a firm as a unified, conceptually coherent body that incorporates even dissimilar services.

Our research investigates beliefs about managerial control over service units as well as information valence as two factors that may influence the extent to which consumers use information about the quality delivered by one group of service providers to make inferences about the quality to expect from another group of service providers that is perceived as different from the other. Based on previous research on consumers’ inductive inferences, we develop a series of hypotheses that are tested in an empirical study. Though we base some of our hypotheses about services on research about objects, the extent to which that research is applicable to services is unclear. Theorists have identified numerous ways in which the processing of object categories differs from the processing of human or social categories (Medin/Lynch/Solomon 2000).

2. Consumers’ Category-Based Inductive Inferences

Our investigation into consumers’ quality inferences draws on research on category-based induction. Most cognitive psychology research on category-based induction examines the use of information about one category to make inferences about another target category that also belongs to the same natural class (Markman/Ross 2003; Proffitt/Medin/Coley 2000; Yamauchi/Markman 2000). For example, when a person is told that robins have an ulnar artery, the person is likely to assume that ostriches also have an ulnar artery simply because both belong to the same natural class of birds (Proffitt/Medin/Coley 2000).

These distinctions about categories for natural classes (e.g., birds) have been adopted in the marketing literature in the context of brand extensions. The brand-extension literature suggests that a family brand (e.g., Sony, Ford, Nestle) can be thought of as a superordinate class encompassing categories of products (e.g., Sony televisions, Sony clock radios) (Boush/Loken 1991). Empirical evidence of consumers’ category-based inductions across different products belonging to the same brand (a superordinate class) was provided by Joiner/Loken (1998).

The authors showed that attributes associated with one product category are generalized to or projected on to another product category that shares the same brand name. For example, when consumers know that Sony receivers are dependable, they infer that Sony clock radios are dependable.

Many brand extension studies suggest that generalizations of the parent product’s attributes to an extension are stronger for similar products (often referred to as near extensions) than for dissimilar products (far extensions) (e.g., Aaker/Keller 1990; Boush/Loken 1991; Völckner/Sattler 2006). When the parent and the product share similar features (e.g., have the same physical appearance, serve the same function), evaluations of the extension are more positive than when dissimilar (Park/Milberg/Lawson 1991). A door-to-door survey of Germans’ perceptions of 66 brand extensions offered in Germany found that perceived “fit” or similarity was one of the most important factors in predicting an extension’s success (Völckner/Sattler 2006). Brand extension similarity effects are consistent with the basis for inductive inferences suggested by categorization theory. Induction occurs by identifying features that two units share. If they share dominant features (are perceived as similar on relevant attributes), then they are assumed to share additional features (Yamauchi 2005).

Services offered by a firm can also be conceptualized as categories belonging to a superordinate class (the firm). For example, an airline’s baggage handlers offer a distinctive service to passengers from an airline’s flight attendants, but both belong to the same firm. If the firm is a superordinate class for separate services, then consumers should use information about one of the firm’s services to infer that another of the firm’s services has the same quality merely because both are offered by the same firm. An attribute that suggests high quality should lead to more favorable inferences about the target unit than an attribute that suggests low quality. Consistent with this notion, corporate image influences consumers’ preferences for service brand extensions, though similar services are preferred over dissimilar services (Ruyter/Wetzels 2000).

The service quality literature supports the notion that consumers generalize from information about one service unit to another service unit, even though the units are dissimilar (e.g., Boulding et al. 1993). Service quality assessment is multidimensional, with five major dimensions: reliability, assurance, responsiveness, empathy and tangibles (Parasuraman/Zeithaml/Berry 1985, 1988). These abstract dimensions permit generalizations despite the heterogeneity inherent when different employees deliver service. The overall service quality judgment represents an averaging across the different dimensions (Boulding et al. 1993). Hence, the overall quality assessment about one unit of service providers should influence inductive inferences about the firm’s service quality as a whole, because both judgments are made at a
global, abstract level (Zeithaml 1988). The assessment of the firm’s service quality should influence expectancies about overall service quality for another unit within the firm even though that second unit is perceived as dissimilar from the first unit.

**H1:** Information about the high quality delivered by one unit of service providers leads to favorable inferences about the quality of another, dissimilar unit of service providers, whereas information about the low quality delivered by one unit of service providers leads to less favorable inferences about the quality of another dissimilar unit.

Whereas information about one category’s quality should influence inferences about others, the valence of that information may affect inferences differently. In contrast to the negativity effect often found for product evaluations (e.g., Ahluwalia 2002; Herr/Kardes/Kim 1991), previous research in services marketing has found a positivity bias in consumers’ generalizations from information about an individual to other employees within the same firm (Folkes/Patrick 2003). Information about an excellently performing individual is more likely to be generalized to other employees than is information about a poorly performing individual. Folkes/Patrick (2003) reasoned that a positivity bias occurs for generalizations from an individual because the poorly performing individual can be regarded as an anomaly.

It is unclear whether a positivity bias would also be found when generalizing not just from an individual but from a group (e.g., not just from one flight attendant to all flight attendants, but from information about flight attendants to inferences about employees in the firm as a whole and even to baggage handlers). A positivity bias might occur since people generally evaluate services positively (Fornell et al. 1996; Johnson/Anderson/Fornell 1995). Expectations of good service might suggest that one unit’s weakness might not be that diagnostic of poor service from other units or from the firm’s employees as a whole. The notion of widespread service incompetence seems incompatible with cultural beliefs that people select jobs that match their aptitudes and remain in their jobs because they perform competently. If so, then consumers should show a positivity bias when drawing inferences about a different service category and about the firm’s employees as a whole. Positive information about one service category should lead to similarly positive evaluations about the initial service category and about a different, target service category, as well as the firm as a whole, more than negative information leads to similarly negative evaluations.

**H2:** Inferences about the firm as a whole and about another service unit show a positivity bias. Information about the high quality delivered by one unit of service providers leads to more similarly positive inferences about the firm’s employees more than information about the low quality delivered by one unit of service providers leads to similarly negative inferences about the firm’s employees. Information about the high quality delivered by one unit of service providers leads to more similarly positive inferences about a different unit more than information about the low quality of the unit lead to similarly negative inferences.

An important finding of category-based induction research is that even though people do make inductive inferences across categories, inferences about another category are not as strong as inferences about the class as a whole (Osherson et al. 1990; Shafir/Smith/Osherson 1990; Sloman 1993). For example, given information about robins’ arteries, people make stronger inferences about birds in general than about ostriches’ arteries. Osherson et al. (1990) labeled the counternormative finding of stronger inferences about a natural kind (e.g., birds) than about the target category (e.g., ostriches) the inclusion fallacy.

The inclusion fallacy occurs not just for natural kinds but also when brand names serve as the superordinate class. Joiner/Loken (1998) found that consumers used information about one branded product (e.g., Sony television dependability) to make inductive inferences about Sony products’ dependability, but then did not make as strong inductive inferences about another category of Sony products (e.g., Sony cameras). The inclusion fallacy suggests that consumers fail to generalize from class to the target category to the extent that they should. A category should share the same attributes as the class. Evidence for the inclusion fallacy comes from finding differences between inferences about the class and the target category.

Since the inclusion fallacy has been found for natural kinds as well as for brands, we expect that it will also occur when making inductive inferences about services. Within the general class of a firm’s employees are categories of employees that deliver different types of services. We should find a greater difference between quality perceptions of the target category and the initial category than between the class and the initial category. For example, when consumers receive information about an airline’s flight attendants, their inferences about flight attendants’ service quality should be more similar to their inferences about the quality of the firm’s service in general than to consumers’ inferences about the quality of the airline’s baggage handlers (if the inclusion fallacy applies).

**H3:** Quality inferences about the firm as a whole differ from quality inferences about the target service category, with perceptions of the quality of the firm more similar to quality perceptions of the initial service category than are perceptions of the quality of the target.

Moreover, previous research suggests that increasing the salience of the inclusion fallacy’s illogic may accentuate the effect. Joiner/Loken (1998) reasoned that the mem-
bership of a product category in the greater class of branded products should be salient if consumers are forced to make judgments about all the branded products prior to making the judgment about a category, resulting in similarity between class and category. Yet, their empirical evidence suggested it had the reverse effect. For example, all Sony product dependability was related to Sony bicycle dependability more weakly when consumers first rated all Sony products and then Sony bicycles (i.e., the order of items was class first followed by the specific subcategory) as compared to when they rated Sony bicycles without rating all Sony products (i.e., specific products only with no class ratings). Joiner/Loken (1998) indicated that the most likely explanation for this unpredicted finding was a contrast effect. Making a general judgment about the brand may have accentuated the target product’s atypicality for the brand (e.g., made Sony bicycles seem even more atypical of Sony products in general). We therefore expect to find a greater difference between quality inferences about the class and the target category when respondents are asked to first make inferences about the class and subsequently about the target category.

**H4: Quality inferences about the class differ from quality inferences about the target category more when consumers first make inferences about the class than when they first make inferences about the target category.**

Whereas H1-4 are guided by findings about product or service inferences, our research also explored a potential influence on inductive inferences not investigated previously. We examined how consumers’ beliefs about the control that managers exert over service units can be perceived as a cause of similar quality across units and so influence inductive inferences. Several categorization theorists have suggested that belonging to the same superordinate group is only one factor that explains inductive inferences. They argue that category-based induction frequently relies on “causal knowledge that imposes structure on our beliefs about categories and supports projections that enable people to ‘go beyond the information given’” (Rehder/Hastie 2001, p. 323). By drawing on their knowledge, people reason about how and why a property might appear across categories (Boriarczyk/Alba 1994). When making inferences about the quality to expect from a group of service providers, consumers may use knowledge they have about mechanisms that influence service quality. Thus, their inferences should be based on beliefs about how and why quality might be similar across different service units.

Consumers may expect that by selecting, training and motivating employees to achieve high quality, a service firm exerts managerial control. The control that firms have over all their units should facilitate generalizing from the high quality of one unit to another. If a firm is able to ensure high-quality service delivery from one service unit, it should also be able to ensure similar quality from another unit. Similarly, if a firm fails to ensure high-quality service delivered by one unit, it seems likely that similar low quality will be delivered from the firm’s other units (Hartline/Ferrell 1996). Thus, when consumers believe that similar quality across different service units occurs because of managerial control mechanisms, they should be more likely to use information about one unit to infer similar quality delivered by another unit.

**H5: Information about the quality delivered by one unit of service providers leads to more similar inferences about the quality of another unit when drawing on theories of managerial control compared to when consumers do not perceive categories to relate to each other in terms of managerial control.**

### 3. Empirical Study

Our research investigated consumers’ generalizations across different units within a firm when the valence of information varied. We investigated category-based induction effects by comparing the outcome of consumers’ inferences (i.e., quality ratings), but also investigated the processes leading to those inferences. As with previous research examining how people arrive at inductive inferences (e.g., Proffitt/Medin/Coley 2000), we asked individual respondents to explain their judgments. Although we anticipated that respondents would report that their judgments reflected beliefs about firm’s control over both units, our study examined theories in a context in which managerial control would not be salient. Those who theorize about the firm’s control over both units should make category-based inductive inferences more than those who do not.

### 3.1. Method

As with most previous studies examining consumers’ category-based inductive inferences (e.g., Joiner/Loken 1998; Proffitt/Medin/Coley 2000), we used an experimental design that allowed us to manipulate various factors. For the purpose of our research, a 2x2x2x3 design was used. The three between-subjects factors were quality information (high quality vs. low quality), initial service category typicality (highly typical service vs. moderately typical service) and order of inferences (firm inference made last, after the target category inference vs. target category made last, after the firm inference). The within-subjects factor was the entity being rated (the initial category vs. the target category vs. the firm’s employees). Respondents were 239 undergraduates who completed a questionnaire as part of a course requirement. Although a student sample is not representative of adult service users, most undergraduate students are regular or at least infrequent airline customers and thus can be considered an appropriate group of respondents for the purpose of this study.

The questionnaire gave information about the quality of one unit within an airline and asked respondents about...
the quality of another of the airline’s unit. Airlines are a
good context to examine consumers’ inferences because those employees perform their jobs independent of obvi-
sous supervision. Further, airlines have service units that
operate in separate spheres (e.g., flight attendants and
baggage handlers), reducing the likelihood of respond-
ents identifying a single managerial control mechanism.

3.2. Pretest

A questionnaire was given to 40 students to examine per-
ceptions of airline occupations, and in particular, whether some service units might be more typical than
others. Inclusion fallacies found in Joiner/Loken’s work
(1998) were attenuated depending on the category’s typi-
cality. Consumers generalized less when the target cate-
gory was atypical for the brand, which is also true of nat-
ural classes (Osherson et al. 1990). Similar to Loken/
Joiner/Peck (2002) and Roedder/Loken/Joiner (1998),
we identified the most typical exemplars of airline occu-
pations by asking respondents “when you think about an
airline, which occupation comes to mind first?” After
writing down the first to come to mind, respondents then
wrote down the second and third occupations to come to
mind. Flight attendants were the first occupation to come
to mind for 19 respondents (47.5%), and were mentioned
either first, second or third by all 40 respondents. No
respondents mentioned baggage handlers either first or
second, but 6 respondents (15%) mentioned them third.
Respondents were then asked “what occupation do you
feel is the very best example of what an airline stands
for?” Flight attendants were mentioned by 23 respond-
ents (57.5%), leading to its selection as the very typical
unit to be manipulated in the main study. No respondents
mentioned baggage handlers as the best example. How-
ever, the ratings that followed suggested that it was not
perceived as atypical (which was the focus of compari-
sion in Joiner/Loken 1998) but as moderately typical.

Pretest respondents were asked to respond to questions
about particular occupations on 9-point rating scales.
First, we asked about the extent to which the occupation
was a good example of the kind of people who provide
passengers with services at an airline. Similar to Loken/
Ward’s (1990) three-item scale to measuring typicality,
we used example goodness (1 = poor example, 9 = good
example), typicality (1 = not at all typical, 9 = very typi-
cal) and representativeness (1 = not at all representative,
9 = very representative) (for detailed information on scale development and validity see Loken/Ward 1990).
Scale reliability was satisfying (α = .86). Flight atten-
dants were considered a better example than were bag-
gage handlers, with baggage handlers rated around the
midpoint (M = 8.67 versus 5.68, respectively), t (39) =
8.457, p < .001. We included both the very typical and
the moderately typical services in our study. Expected
quality in general was high for both occupations, but
flight attendants were rated more highly (M = 7.03 vs.
6.00), t (39) = 3.459, p = .001. As expected, familiarity
and expertise with airlines was high (M = 7.98 and 8.03,
respectively, with 1 = not at all familiar, 9 = very familiar
and 1 = none, 9 = a lot).

3.3. Main Study Procedures and Measures

Participants in the main study were told that a travel
magazine had rated the flight attendants/baggage han-
dlers that work for a particular airline at the top/bottom
of those in their profession. The service providers were
described as more/less competent, efficient and knowl-
edgeable compared to those that work for other airlines.
Competency, efficiency and knowledgeability about how
to successfully perform one type of service function were
selected because these attributes would not be demon-
strated in an identical manner in the other type of service
function. For example, the skills in loading baggage effi-
ciently are different from those involved in loading pas-
sengers on to a plane efficiently. Moreover, the SER-
QUAL dimension relevant to one unit seems different
than for the other unit (e.g., competency in baggage han-
dl suggests “assurance” for that service whereas com-
petency for a flight attendant suggests “empathy” and
“responsiveness”) (Boulding et al. 1993; Parasuraman/

After reading the scenario, participants responded to 9-
point ratings scales. The initial items served as manipula-
tion checks for the information given about the service
unit. Participants were asked to indicate the quality of
service they would expect to receive from the unit
described as either at the top or the bottom on three
scales (1 = poor quality vs. 9 = excellent quality, 1 = very
bad quality vs. 9 = very good quality, 1 = very low qual-
ity vs. 9 = very high quality) (α = .99). Then they were
asked to make inferences about another unit. Specifi-
cally, they were asked to indicate on the same three scales the quality of service they would expect to receive
from flight attendants who work for the same airline if
they had been told about baggage handlers’ service or
asked to infer baggage handlers’ quality if they had been
told about flight attendants’ service (α = .99). Then an
open-ended question asked “why did you rate the quality
of service the way you did?” for the inferred occupation.
Respondents also indicated the quality of “service you
would expect to receive from all the employees who
work for the same airline as the flight attendants/bag-
gage handlers” on the same three items (α = .99).

Additional items provide support for the notion that nei-
er unit controlled the other and that the two units were
not perceived as similar. Participants responded to a
rating scale asking, “how much control do flight atten-
dants have over the quality of the airline’s baggage han-
dlers” (1 = none, 9 = a lot). A second control item asked
about the baggage handlers’ control over the quality of
the airline’s flight attendants. Neither unit was perceived
to have control over the other (M = 2.04 for flight atten-
dants’ control over baggage handlers, M = 1.53 for bag-
gage handlers’ control over flight attendants). Respond-
ents rated the similarity between baggage handlers and
flight attendants on a 9-point-rating scale (1 = none, 9 = a lot). The mean similarity was below the midpoint (M = 3.20). To assess service familiarity, three items were used. The first asked “how many airline flights have you taken in the last year” (M = 4.85). Then respondents rated their experience with airline services (1 = none, 9 = a lot), (M = 6.68) and their knowledge about what airline employees do (1 = none, 9 = a lot), (M = 5.26).

Two judges coded the open-ended responses into one of six categories: (1) the firm has control over employees (“the same firm hires and trains the baggage handlers and the flight attendants”), (2) the units are dissimilar (“their jobs are very different because flight attendants interact with people on the plane but baggage handlers just lift the luggage”), (3) the units are independent (e.g., “the flight attendants and baggage handlers don’t interact and have nothing to do with each other”), (4) characteristics of the occupation (“flight attendants are always polite and helpful”), (5) miscellaneous, and (6) unclear or no explanation.

3.4. Results

The results were analyzed by a 2x2x2x3 ANOVA, followed by comparisons across pairs of means. The ANOVA showed three significant main effects: a service quality main effect, a typicality main effect, and an entity being rated main effect (F (1, 231) = 607.48, p < .001, F (1, 231) = 4.25, p < .05, F (2, 462) = 24.60, p < .001). Those main effects are qualified by significant two-way interactions: a quality by entity being rated interaction, and a service category typicality by entity being rated interaction, F (2, 462) = 229.63, p < .001, and F (1, 462) = 21.00, p < .001.

Manipulation check. As intended, an ANOVA of just the ratings for the initial category showed that quality was higher for the units described as being top of their profession (M = 8.39) than for the units described as being at the bottom of their professions (M = 2.36), F (1, 238) = 2,205.48, p < .001. The ANOVA of the initial category ratings showed no other effects.

H1. The results support the notion that respondents make category-based inductive inferences. Inferences about the target category were more positive when respondents received positive information about the other unit than when they received negative information about the other unit (M = 6.92 vs. 4.75). An ANOVA of just the target category quality ratings revealed a main effect for quality (F (1, 238) = 115.04, p < .001), with information about a high quality service leading to more favorable inferences than information about a low-quality service.

H2. A positivity bias in inferences from the initial category is suggested by the quality by entity being rated interaction. That interaction is consistent with the more extreme ratings of the initial category than of the other entities (see Table 1). More extreme ratings for the initial category are expected because information was given about their quality. However, the means also suggest a positivity bias in that information about the high quality of the initial product led to more similarly positive inferences about the firm’s employees in general than information about low quality led to more similarly negative inferences about the firm’s employees (M = 8.36 vs. 7.37 for high quality; M = 2.36 vs. 4.73 for low quality). That discrepancy is supported by comparing the absolute difference between each participants’ initial category rating and the firm ratings. The mean difference is smaller for the high quality conditions than for the low quality conditions (M = 1.13 vs. 2.42, t (237) = 6.46, p < .001).

The same positivity bias can be seen when comparing the initial category to the target category. Evaluations of the initial category are more similar to the evaluations of the target category in the high quality condition than in the low quality condition (M = 8.36 vs. 6.92 in the high quality condition; M = 2.36 vs. 4.75 in the low quality condition). The absolute difference between each subject’s ratings for the initial category and the target category is

<table>
<thead>
<tr>
<th>Condition</th>
<th>Initial Category</th>
<th>Firm Category</th>
<th>Target Category</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Quality</td>
<td>8.36 &lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.37&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.92&lt;sup&gt;c&lt;/sup&gt;</td>
<td>123</td>
</tr>
<tr>
<td>Very Typical Initial – High Quality</td>
<td>8.52&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.39&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.38&lt;sup&gt;c&lt;/sup&gt;</td>
<td>64</td>
</tr>
<tr>
<td>Moderately Typical Initial – High Quality</td>
<td>8.26&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.36&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.51&lt;sup&gt;b&lt;/sup&gt;</td>
<td>59</td>
</tr>
<tr>
<td>Low Quality</td>
<td>2.36&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.73&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.75&lt;sup&gt;b&lt;/sup&gt;</td>
<td>116</td>
</tr>
<tr>
<td>Very Typical Initial – Low Quality</td>
<td>2.43&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.59&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.31&lt;sup&gt;b&lt;/sup&gt;</td>
<td>59</td>
</tr>
<tr>
<td>Moderately Typical Initial – Low Quality</td>
<td>2.30&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.87&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.21&lt;sup&gt;c&lt;/sup&gt;</td>
<td>57</td>
</tr>
<tr>
<td>Very Typical Initial</td>
<td>5.57&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.05&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.38&lt;sup&gt;a&lt;/sup&gt;</td>
<td>123</td>
</tr>
<tr>
<td>Moderately Typical Initial</td>
<td>5.33&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.13&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.38&lt;sup&gt;c&lt;/sup&gt;</td>
<td>116</td>
</tr>
</tbody>
</table>

<sup>a,b</sup> Means across the rows for a given dependent variable with different superscripts are different at p < .05

<sup>1</sup> Initial category was flight attendants. Target was baggage handlers.

<sup>2</sup> Initial category was baggage handlers. Target was flight attendants.
smaller for the high quality conditions than for the low quality conditions (M = 1.59 vs. 2.42, t(237) = 3.92, p < .001).

H3. There appears to be limited evidence for the inclusion fallacy. The inclusion fallacy suggests that consumers fail to generalize from their inferences about the firm as a whole to the target category. Evidence for the inclusion fallacy would require within-subjects differences between two pairs of means – the mean for the firm and the mean for the target. Table 1 shows that ratings of the firm and of the target differ for within-subjects comparisons for the high quality condition (M = 7.37 vs. 6.92 for the firm and the target, respectively), consistent with the inclusion fallacy. However, ratings are not significantly different for the low quality conditions (M = 4.73 vs. 4.75 for the firm and the target, respectively). We examined the means for the two initial category conditions to gain additional insight into why the pattern for the high quality condition differed from the low quality condition, as well as considering the interpretation of the other significant two-way interaction.

The significant two-way interaction between category typicality and entity being rated appears to be due to more positive inferences about flight attendants (the very typical category) than of baggage handlers (the moderately typical category). Table 1 shows means for this two-way interaction when quality information is held constant. Means are higher when flight attendants are the target category (M = 6.38) than when baggage handlers are the target (M = 5.38), t(237) = 3.97, p < .001. Hence, ratings of the target category appear to be influenced by both information about the target’s quality and more generalized expectations about that category’s quality.

Although the three-way interaction of quality, category typicality and entity was not significant, Table 1 shows the means for each condition to support this interpretation of the significant two-way interaction. The same pattern emerges even when quality is not held constant (i.e., for the separate high and low quality conditions). The means are more positive for the flight attendants when they are the target (M = 7.5 and 5.21 for flight attendant quality) than for baggage handlers when they are the target (M = 6.38 and 4.31 for baggage handler quality). Those more favorable ratings for flight attendants than for baggage handlers are consistent with our pretest results. In sum, the limited evidence for the inclusion fallacy seems to be more parsimoniously explained by more negative evaluations of one service category.

H4. The ANOVA showed no order effects, indicating that whether respondents made inferences about the firm first or the target category first had no effect on quality ratings.

H5. Evidence for H5 is found by examining the open-ended responses. The single most common explanation for ratings (41% of respondents, n = 98) related to the firm’s control over employees (e.g., “the same people hire all types of employees,” “the training is the same for all employees”). These respondents explained their ratings by providing a theory for why the two categories might be similar in service quality based on beliefs about the way firms control employees’ behaviors.

The other explanations for respondents’ ratings reflect different processes. 20% of respondents (n = 48) made direct comparisons between the initial category and the target, without mentioning all employees or the firm, and concluded that the two categories were different. These respondents explained their ratings by referring to the dissimilarity between the jobs (12.1% of respondents, n = 29) (e.g., “flight attendants and baggage handlers do different things”), or by referring to the independence of the units (7.9% of respondents, n = 19) (e.g., “flight attendants and baggage handlers have nothing to do with each other”). Those respondents seemed to believe that category-based induction was unwarranted. Indeed, inferences about the target category were the same for these 48 respondents, regardless of whether the initial category was high or low quality (M = 5.63 for high quality vs. 5.55 for low quality).

In contrast to those who directly compared across the two categories, some respondents focused on the target category only. The second most common explanation for ratings focused on beliefs about the specific occupations (20.9%, n = 50) (e.g., “flight attendants have to be pleasant in their job”; “baggage handlers are generally rude”). These respondents appear to use norms drawn from experience to predict quality. Inferences about the target category also did not differ (M = 6.38 for high quality vs. 5.51 for low quality). Others (5%, n = 12) gave miscellaneous explanations that fell into none of the previous categories. Some respondents (12.9%, n = 31) reported that they guessed or they merely described their ratings (e.g., “I rated the flight attendants low”).

To compare perceived quality ratings of those who theorized about control versus those who did not, we conducted a 2x2x3 ANOVA, with quality information as one between-subjects factor and managerial control versus no managerial control as the second between-subjects factor. (We excluded the 31 respondents who reported that they guessed or that merely described their ratings from this analysis.) There was a significant quality main effect (F (1, 204) = 720.74, p < .001), a significant two-way quality by control interaction (F (1, 204) = 59.97, p < .001), a two-way quality by entity being evaluated interaction (F (1, 204) = 191.10, p < .001), and a three-way quality by control by entity being evaluated interaction (F (1, 204) = 31.02, p < .001). The managerial control interaction suggests that those who explained their ratings by referring to managerial control rated the target’s perceived quality more similar to the initial category than the other respondents (see Table 2). The absolute difference scores between the perceived quality ratings of the target and the rating of the initial category
were significantly smaller for respondents that mentioned managerial control than for those who did not (M = 1.05 vs. 2.71, t (206) = -8.13, p < .001).

4. Discussion

Our study examined inductive inferences from information about the quality of one unit of a firm to the firm’s employees as a whole and to another, independently functioning unit. The overall results show that respondents drew category-based inductive inferences from information about the quality of one service unit to the other service unit, even though the two units were not perceived as similar. Similarity has emerged as an important influence on inductive inferences from parent products to brand extensions (e.g., Aaker/Keller 1990; Boush/Loken 1991; Park/Milberg/Lawson 1991). However, belonging to the same firm appears sufficient for quality information to influence inferences about the dissimilar services we examined.

Although our study participants did draw inferences, those inferences were asymmetric depending on the valence of the information. We observed a positivity bias, with positive information having a greater impact than negative information on inferences about the firm and on inferences about the dissimilar service category. Positive information about a category’s high quality leads to more similar inferences between that category and the firm and the firm’s employees in general, as contrasted with the greater discrepancies between inferences about the category and the firm in response to negative information. The positivity effect was also observed in inferences about the dissimilar service unit. A positivity bias has been found for services previously in generalizations from the individual to the service group (Folkes/Patrick 2003). Previous experiments have not tested for a positivity bias from the service group to another, dissimilar unit. Hence, this finding represents a novel contribution for our study. A positivity bias for services contrasts with the preponderance of laboratory experiments showing a negativity bias in inferences drawn from product information under the same type of conditions used in our study (Ahlawatia 2002).

Whereas the positivity bias suggests consistencies within the service literature about consumers’ inferences, the absence of support for the inclusion fallacy suggests that consumers’ inferences about products may differ from those regarding services. Our experiment found little—if any—evidence of the inclusion fallacy. That inclusion fallacy would have been revealed by participants’ drawing quality inferences about the firm more than quality inferences about the target category. When we found means for the firm that did differ from those for the target category, those differences could be explained more parsimoniously by more positive quality perceptions of one category than of the other category. Hence, support for the inclusion fallacy is questionable in this service context.

Perhaps the inclusion fallacy is more likely to be found with products than with services because product attributes are often concrete whereas service quality is evaluated more abstractly. Certainly, the lack of evidence for the inclusion fallacy is consistent with the SERVQUAL model. Service customers should be able to generalize from information about one service unit to another service unit because the overall service quality judgment for the firm represents a global summary of multiple abstract service dimensions (e.g., empathy, assurance, Boulding et al. 1993). Consistent with the SERVQUAL model, information about one service provider influenced our study participants’ overall service quality judgment. That overall judgment appeared to set expectations for a dissimilar service’s quality that took into account expectations for that occupation (i.e., expectations that flight attendants deliver higher quality service than baggage handlers).

A novel finding from our research that may be unique to the service context is that inferences across different service providers seem to derive at least in part from beliefs about managerial control. Respondents who explained their quality ratings by referring to managerial control perceived the target to be more similar in quality to the initial service category than did the other respondents (see Table 2). This finding is important since the majority of our respondents referred to managerial control as a
basis for assuming similar quality from different service providers within a firm. They seem to believe that managers’ control over service quality causes unified service standards. Because the service categories presented to participants function independently in separate locations, it is not surprising that many study participants explained their ratings in other ways than by referring to overarching managerial control. Managerial control probably was not particularly salient. Those that did explain their inferences based on managerial control may have used it as a basis for integrating the dissimilar service categories into a conceptually coherent class.

To our knowledge, ours is the first experiment that investigated valence biases in category-based inferences for dissimilar service units. Our results suggest caution in assuming that studies about category-based inference in regard to natural kinds and tangible products also apply to services. Positivity biases in services suggest that negative information about a service category may not be so damaging to firms as one might expect from experiments investigating inferences about products.

Further, service-specific causal beliefs about managerial control have a heretofore unrecognized impact on consumers’ inferences across different groups of service providers. Hence, service companies need to understand how their customers perceive their various groups of service providers and may need to manage beliefs about managerial control, particularly when their service employees function independently and are perceived as dissimilar. Consumers’ beliefs about managerial control mechanisms can be steered by enhancing perceptions of a common management demanding the same high standards. The focus of managers’ efforts should be on guaranteeing excellent service quality from the service category that consumers have the most information about, because excellent service from that unit is influential in inferring the quality of the firm’s services (as long as it is perceived as typical of the firm’s other services).

Our experiment suffers from some limitations that should be recognized. One of the shortcomings of our experiment is the use of a student sample. Even though undergraduates are frequent airline customers and our pretest showed that our respondents were familiar with airlines, the replication of our study with a different sample may provide stronger evidence for the generalizability of our results. Similarly, more confidence in generalizability can be gained from examining inductive inferences in other service industries besides the one examined here, in other occupations, and in more naturalistic situations.

References


The Influence of Information Valence and Managerial Control on Consumers’ Inferences Across Service Providers


