© Academy of Management Journal
2004, Vol. 47, No. 3, 433–445.

THE EFFECTS OF PHYSICAL AND SOCIAL CONTEXT ON EVALUATIONS OF CAPTIVE, INTENSIVE SERVICE RELATIONSHIPS

DONALD E. CONLON
LINN VAN DYNE
MORGAN MILNER
Michigan State University

KOK YEE NG
Nanyang Technological University

Expanding a conceptual framework, we differentiated services on the basis of their levels of captivity (the difficulty of a customer's leaving) and intensity (the number of services performed), arguing that context is especially critical to service delivery when these levels are high. Data from cruise ships generally supported our hypotheses. We report effects of physical and social context on evaluations made by passengers, industry experts, and government regulators. Implications for managers of other service settings are discussed.

Understanding evaluations of service organization performance is important because services comprise a rapidly growing sector of the economy (Gutek, 1995; Rothman, 1998). In this work, we argue it is important to acknowledge that different evaluator audiences have different perspectives and criteria for assessing performance (Cyert & March, 1963; Scott & Lane, 2000). Receiving positive evaluations of service performance from multiple perspectives is critical to service firm reputation, customer satisfaction, and long-term firm success (Bowen, Gilliland, & Folger, 1999; Mitchell, Agle, & Wood, 1997; Seiders & Berry, 1998).

In the major conceptualization of services to date, Gutek (1995) differentiated service encounters from service relationships. In this research, we build on Gutek's work and further develop the conceptualization of services by considering two additonal facets: the *intensity* of service (whether service delivery involves one primary service or multiple services) and the *captivity* of service (whether it is difficult for the party being served to leave the service setting once service delivery has started). We argue that service setting is especially salient in service relationships characterized by high captivity and high intensity (relationships that are difficult to leave and in which customers receive a wide range of services). In developing and testing this basic assertion, we focused on one cell in this framework: high-captivity and high-intensity service relationships, which we also call *captive, intensive service relationships.*

We aimed to make several contributions. First, we focused on a unique and underresearched type of service setting, suggesting that unique characteristics of these settings heighten the salience of context for organizational success (Bowen et al., 1999; Gutek, 1995). We developed and tested our hypotheses in the specific context of cruise ships and discuss applicability to other captive (difficult to exit), intensive (wide-ranging) service settings. Second, we focused on physical and social characteristics as overlooked aspects of service settings (Cappelli & Sherer 1991). Third, we focused on the judgments of three different sets of evaluators: customers, industry experts, and government regulators (cf. Mitchell et al., 1997; Seiders & Berry, 1998).

THEORY DEVELOPMENT

Gutek and her colleagues (Gutek, Bhappu, Liao-Troth, & Bennett, 1999) distinguished between two basic types of service exchanges. *Service encounters* occur when customers may interact with different service providers every time they enter a setting (interacting with different bank tellers is an example). *Service relationships* occur when customers have repeated contact with the same service provider (seeing the same doctor or hair stylist over time is an example). As indicated above, we extend

We thank George Barnekov and Jeff Moretz for coding and entering the data, and Maria Alexandre, Jason Colquitt, David Deephouse, Karin Fladmoe-Lindquist, Amy Hillman, Gerry Keim, George Madsen, Greg Miller, Gerry McNamara, William Ross, Frank Rothaermel, Bob Skarik, and the staff at *Cruise Industry News* and the Cruise Line International Association for helpful comments on this research project.

Gutek's conceptualization by further differentiating types of services on the basis of the dimensions of captivity and intensity. Table 1 summarizes this more elaborate conceptualization of services and provides specific examples for each of the eight cells.

Captive services are services whose customers cannot easily exit the service setting once service delivery has begun because loss of customer control or discretion over the execution of the service is inherent in the situation. As illustrated in Table 1, captivity can occur in service encounters or in service relationships, and it may be of short duration (as in giving blood) or of long duration (as in transoceanic air travel) duration. *Intense services* are varied services, and intensity can also apply to both service encounters and service relationships. When intensity is high, customers receive a wide range of services, such as when a contractor finishes a basement or a family member receives in-home care from service providers. Both captivity and intensity can be viewed as continuums ranging from low to high.

We focused on cruise ships as an example of a setting for service relationships characterized by high captivity and high intensity. Theoretically, we suggest that this cell in the typology creates the highest level of customer dependence on providers, thus making the service context especially salient to evaluators. In our focal context, service delivery is particular and repeated—that is, passengers interact with the same cabin steward, waiter, busboy, and cruise director—creating relationships rather than encounters. Second, the option of leaving the service setting is constrained (captivity is high), since after a ship leaves port and goes out to sea, it is difficult for a passenger to leave (exit) the service context if dissatisfied with the service. Third, the organization in our focal context provides a wide variety of services (intensity is high). For example, on cruise ships service providers deliver food, entertainment, social activities, and accommodations to passengers.

Multiple Perspectives on Service Evaluation

Ginzel, Kramer, and Sutton (1993) argued that firms must attend to their multiple publics and/or audiences in managing organizational impressions. Similarly, in service relationships, an organization's reputation is based on various audience or stakeholder evaluations. For example, Scott and Lane (2000) contrasted internal audiences (managers and employees) with external audiences (customers, suppliers, regulatory agencies, professional associations, and so forth). Different audiences selectively focus on different information in judging organizational effectiveness. This variation occurs because different audiences often have different priorities and evaluation criteria, which leads to their having different perceptions of firm reputation (Deephouse, 2000; Fombrun, 1996). Accordingly, in our study, we assessed evaluations from multiple perspectives.

The most obvious evaluator perspective is that of the *customers,* who in our case are passengers. For cruise ships and other captive, intensive service relationships, customers and their evaluations are critical to firm success. Even after a cruise, evaluations of customers are critical for two reasons. First, it takes three to five times more money to attract new customers than to retain existing cus-

TABLE 1
An Expanded Conceptual Framework Differentiating Types of Services

	Low Captivity: It is not difficult to exit the service situation even after service delivery starts.		High Captivity: It is difficult to exit the situation even after service delivery starts.	
Variable	**Low Intensity: One Primary Service**	**High Intensity: Multiple Services**	**Low Intensity: One Primary Service**	**High Intensity: Multiple Services**
Service encounter: No provider continuity	Walk-in haircut Teller transaction Fast-food lunch Discount store purchase Glasses in 60 minutes	Day spa visit Drop-in day care use Bed and breakfast stay	Giving blood MRI procedure Mammogram Taxi ride	Long plane flight Urgent care medical visit Weight loss clinic
Service relationship: Particular provider continuity	Office UPS delivery Six-month dental cleaning Weekly dry cleaning pickup Haircut with regular stylist	Annual physical Tax planning appointment Estate planning appointment Insurance update appointment	Dental cavity filling Hair coloring and styling Weekly manicure Personal training Contractor remodeling	Cruise Boarding school residency Nursing homes Wilderness retreat Extended hospital stay residency

tomers (Jones & Sasser, 1995). Second, former passengers powerfully influence ships' reputations and the decisions of others to cruise on specific ships. Favorable word of mouth and positive recommendations can increase the potential customers who consider firm services. Since customers are critical to the success and survival of service organizations, we focused on their evaluations first, considering three passenger assessments: (1) overall evaluation, and specific evaluations of (2) ship activities and (3) cabin accommodations.

The second evaluator perspective we examined is that of *industry experts.* When selecting expensive or important services, customers often use publications and Web sites to evaluate service providers (for example, *U.S. News & World Report* evaluates hospital services such as cardiology; *Morningstar* evaluates mutual funds; *Consumer Reports* evaluates insurance providers). In the cruise industry, the Berlitz and the Frommers guidebooks provide evaluations that allow customers to select cruises that fit their personal objectives. In our research, we focused on the overall evaluations of ships provided by these expert information intermediaries.

Our third evaluator perspective was that of *government regulators.* Service providers routinely interact with government regulators (for instance, the Occupational Health and Safety Administration, and local fire and health authorities). One of the most serious events that can happen on a cruise is an outbreak of illness resulting from inappropriate food preparation or sanitation. For example, an outbreak of gastrointestinal illness in 1994 aboard the Royal Caribbean Cruise Lines *MS Viking Serenade* produced diarrhea, vomiting, and fever in 583 passengers and 40 crew members (Marti, 1995). More recent outbreaks have occurred on *Disney Magic* (June 2–5, 2000, and again in November 2002) and Holland America's *MS Amsterdam* (November 2002). Failure to meet regulatory standards can cause a ship to lose its license to sail. For example, in July 2001, the U.S. Centers for Disease Control (CDC) issued a "no sail order" for the *MTS Arcadia* until the ship completed extensive clean-up efforts. In our study, we examined CDC ship inspection scores as one important government regulator perspective.

We suggest that each evaluator perspective differs and that these differences will influence evaluations of a service setting. For example, customers want to have an enjoyable cruise, and thus they should be particularly interested in ship characteristics that influence their own personal cruising experience. In contrast, government regulators must assure safe and healthy service environments, leading them to focus on a narrow set of objective criteria and precise conformity to health and safety requirements. Industry experts have two goals: evaluating ships and services relative to those of other ships, and providing detailed descriptions that allow passengers to choose cruises that best suit their goals. Cruise ships, like other service providers, must respond to the differing perspectives of diverse audiences and must manage their self-presentations so that each group has favorable views of them. Although different factors can influence evaluations, we focused on both physical and social aspects of the cruise context for our hypotheses because context has special relevance to high-captivity, high-intensity settings.

The Role of Physical and Social Context in Services

Although early management work acknowledged that the physical environment or context can influence attitudes, evaluations, and behavior (Barnard, 1938; Roethlisberger & Dickson, 1939), much of the management and organizations literature has underemphasized the role of setting and has focused more narrowly on microlevel stimuli and responses (Pfeffer, 1998; Stokols, 1995). For example, although Pfeffer identified several streams of research on the contexts of organizations and observed that aspects of context can influence a wide variety of organizational outcomes, he noted the need for additional research on context.

The proponents of recent typologies have argued for the importance of physical and social elements in service behavior settings (e.g., Fischer, Gainer, & Bristor, 1997). Physical aspects reflect the tangible, immutable aspects of work contexts, while social aspects incorporate characteristics of individuals in the context of a specific setting. It was consistent with Pfeffer's (1998) observation that we could find little research addressing the influence of physical and social aspects of context on evaluator perspectives. For exceptions, see the research on physical setting effects on customer and employee satisfaction; Bitner (1990), for instance, studied customer satisfaction in hotels and restaurants, and Oldham and Rotchford (1983) examined the issue of employee privacy. In addition, most of this existing research has taken one or a few perspectives (e.g., Quinn & Rohrbaugh, 1983).

In service contexts, *physical aspects* of context are important factors that provide tangible cues regarding the quality and level of service. This view is consistent with Pfeffer's (1982) argument that both work requirements and symbolic needs related to status determine physical settings. One

important physical aspect of cruise ships in particular and of service contexts more generally is the age of the space—in our setting, the date when a ship was built. In captive, intensive service relationships, we expected newness of the physical space to be salient to all of our evaluators. For the passengers, we predicted that newer ships would have positive effects on evaluations. Newer ships offer the latest innovations and have more amenities than older ships (Dickinson & Vladimir, 1997). Also, newer ships offer greater proportions of cabins with balconies, a popular amenity among passengers.

Turning to our other evaluator perspectives, we expected that evaluations of industry experts and government regulators would be more positive for newer ships. The role of industry experts is to evaluate ships by taking the perspective of potential passengers. They assess physical aspects of a cruise context in terms of their likely impact on passenger satisfaction. Thus, we expected a positive relationship: the newer the ship, the more positive the industry expert evaluations. Similarly, we expected government regulator ratings of health and sanitation to be higher for newer ships, because older ships can be more difficult to clean and maintain. Challenges include structural problems with potable water systems and lack of permanent hot and cold food setups for buffets and locations (such as poolside) where food is served outside the dining rooms (Marti, 1995). Thus, we predicted

Hypothesis 1a. Ship newness will be positively related to passengers' overall evaluations of a cruise.

Hypothesis 1b. Ship newness will be positively related to passengers' evaluations of cruise activities.

Hypothesis 1c. Ship newness will be positively related to passengers' evaluations of cabin accommodations.

Hypothesis 2. Ship newness will be positively related to industry experts' evaluations of a cruise.

Hypothesis 3. Ship newness will be positively related to government regulators' evaluations of a cruise.

Size of the physical space is another characteristic of the physical context that we expected would be critically important in captive, intensive service relationships. While at sea on a cruise, passengers are confined to the ship and spend significant amounts of time within the physical setting. Larger ships provide more opportunities to explore and discover new areas. More space can reduce monotony and increase satisfaction. Accordingly, we reasoned that the physical characteristic of ship size would positively influence passengers' overall evaluations as well as evaluations of cruise activities. In contrast, we did not expect overall size of the ship to influence passengers' judgments of their individual cabins. This is because the size of individual cabins is more a function of cruise cost than of ship size.

Shifting to industry experts, we predicted that ship size would be positively related to expert evaluations. As do many service relationships, a cruise provides both tangible and intangible benefits. Within the cruise industry, ship size symbolically conveys status and intangible benefits. Pfeffer (1982, 1998) noted that the sheer size of a physical setting conveys symbolic information. In addition, larger ships have more services that experts can consider in their assessments. Historically, large ships have garnered publicity in part for their sheer size (for instance, the *Bismark,* the *Titanic*). In sum, we expected industry experts to view this aspect of ship context positively and to rate larger ships more favorably.

In contrast to predictions for our first two groups of evaluators (positive effects of size on ratings), our prediction for government inspectors rating health and sanitation was that size would have a negative effect. In the cruise industry, evaluations by government agencies are critical for certification and licensing. Larger ships carry more passengers and crew, features that necessitate storing and processing larger amounts of food and waste. Overall, size thus increases the possibility of problems with general sanitation. Accordingly,

Hypothesis 4a. Ship size will be positively related to passengers' overall evaluations of a cruise.

Hypothesis 4b. Ship size will be positively related to passengers' evaluations of cruise activities.

Hypothesis 5. Ship size will be positively related to industry experts' evaluations of a cruise.

Hypothesis 6. Ship size will be negatively related to government regulators' evaluations of a cruise.

We also expected *social aspects* of context to have important implications for evaluations of captive, intensive service relationships by all of our evaluators. We considered two aspects of social context. One important social characteristic is spa-

tial density (the number of people in a given space or, in our case, the number of passengers relative to the size of the ship). Prior research has demonstrated that spatial density and crowding have robust effects on both animals and people, including increased aggression, insufficient privacy, unwanted social interactions, and psychological distress (Baron, 1993; Hall, 1966). According to social interference theory (Oldham, Cummings, & Zhou, 1995), unwanted or unexpected social interaction interferes with goal attainment and can cause frustration and aggression. From the passenger perspective, a crowded environment on a ship can make it difficult to find a poolside chair, play blackjack, obtain prompt and courteous food service, relax on the cabin balcony, or simply enjoy peaceful contemplation of the water and weather. Generally, higher densities on a ship should reduce passenger perceptions of personal control and generate feelings that vacation goals are being thwarted. Thus, we propose that high spatial density will have a negative effect on passenger evaluations (overall evaluations and evaluations of activities and cabin accommodations).

In keeping with the above logic, we also expected a negative relationship between spatial density and expert evaluations. Ship design and construction are key aspects of cruise quality. When ships are designed to hold more passengers in smaller spaces (spatial density is high), social aspects of the setting, such as crowding, noise, privacy, and the quality of personal interactions are affected. Since social density is a tangible indicator of intangible aspects of the cruise experience, we predicted a negative relationship between social density (crowding) and industry expert evaluations. We also expected a negative relationship between spatial density and government regulator assessments because problems with sanitation, cleanliness, and food handling are more likely in crowded or congested contexts.

Hypothesis 7a. Spatial density will be negatively related to passengers' overall evaluations of a cruise.

Hypothesis 7b. Spatial density will be negatively related to passengers' evaluations of cruise activities.

Hypothesis 7a. Spatial density will be negatively related to passengers' evaluations of cabin accommodations.

Hypothesis 8. Spatial density will be negatively related to industry experts' evaluations of a cruise.

Hypothesis 9. Spatial density will be negatively related to government regulators' evaluations of a cruise.

For our second social aspect of context, we focused on crew-to-passenger ratio, expecting a positive relationship between the number of service providers (crew) and the quality of service. In the cruise industry, the size of a ship's crew (the "complement") relative to the number of passengers is a key measure of service quality (e.g., Dickinson & Vladimir, 1997; Ward, 1999). A high crew-to-passenger ratio allows attentive, personalized service interactions and should enhance overall passenger evaluations but may not affect evaluations of activities and cabins. A high crew-to-passenger ratio should also enhance industry expert evaluations because it is a tangible cue, signaling high levels of high-quality service. Finally, a high crew-to-passenger ratio allows a ship's crew to comply with safety, sanitation, and health procedures and should enhance government regulator evaluations.

Hypothesis 10. Crew-to-passenger ratio will be positively related to passenger overall evaluations of a cruise.

Hypothesis 11. Crew-to-passenger ratio will be positively related to industry expert evaluations of a cruise.

Hypothesis 12. Crew-to-passenger ratio will be positively related to government regulator evaluations of a cruise.

METHODS

Sample and Procedures

We used data from multiple sources, including (1) passenger information from an Internet database, (2) published materials on physical and social characteristics of ship context, (3) published materials from industry experts, and (4) published materials from government inspectors. Although the use of Web-based data is relatively new, it is increasingly important. Stanton (1998) demonstrated general comparability between electronic data collection and more traditional methods with electronic data that showed variance, a factor structure, factor "loadings," and correlations that were similar to paper-and-pencil data using the same items and instruments.

Our passenger sample started with 1,676 people who described their cruise experiences on the Cruise Opinion Web site (http://www.cruiseopinion.com). According to the *New York Times,* this site is "the most comprehensive place for travelers'

feedback on specific cruises" (Barton, 1998). The site provides cruise information to the public by encouraging passengers to rate their own cruise experiences. To participate, passengers log on to the Web site and respond to specific questions rated on a scale from 0 to 100 (90–100 = "excellent," 80–89 = "good," 70–79 = "fair," and 70 = "poor"). Passengers are also asked to describe their cruise experience in unstructured narratives (typically 1–3 pages) that are posted on the Web site. The site began collecting evaluations in August 1996; we obtained information for the 1,676 individuals who had provided reviews by October 1998. The cruise opinion questionnaire can be viewed at http://www.cruiseopinion.com/submit.htm.

Measures

Passenger evaluations. The cruise opinion questionnaire asked respondents to rate 41 specific aspects of their cruise and to provide an overall rating of the cruise. We used this last item as our *overall cruise evaluation* measure. We analyzed the remaining items using factor analysis with varimax rotation. After eliminating items with problematic cross-loadings, six factors emerged that had eigenvalues greater than 1.0 and explained 68 percent of the variance. Two factors coincided with concepts reflected by our hypotheses. We created a *cruise activities evaluation scale* from seven items measuring passenger reactions to the casino, casino staff, deck service, and the entertainment in lounges, and assessments of how good the cruise was for families and for honeymooners (α = .87). We also created a *cabin evaluation scale* from two items measuring comfort and amenities in passenger cabins (α = .87). The other four factors (evaluations of food, embarking/disembarking, the beauty salon, and wheelchair accessibility and exercise facilities) were unrelated to any hypotheses and were excluded from our analyses.

Industry expert evaluations. We used three sources that ranked ships by quality of the cruising experience for passengers. The first two were Web sites, http://www.fieldingtravel.com/cf/shipstar.htm and www.berlitz.com, that award cruise ships from one through six stars. Our third source was the *Berlitz 1999 Complete Guide to Cruising and Cruise Ships* (Ward, 1999), which awards ships 0–2,000 points (ships in our sample had scores ranging from 1,080 to 1,805). The three expert evaluations of ships were highly correlated (r = .65), so we standardized and averaged scores for *industry expert evaluations* (α = .88).

Government regulator evaluations. The Vessel Sanitation Program (VSP), overseen by the CDC, conducts random inspections of cruise ships that call at U.S. ports in order to provide healthful and safe contexts for passengers and crew. Inspection results are reported to travel agents, newspapers, and the cruise industry on "green sheets," which are also posted on the CDC Web site. Ship inspections cover 11 areas (disease reporting, potable water, swimming pools/spas, food safety, food, medical log review, equipment, toilet and hand-washing facilities, toxic substances, facilities, and environmental health). Ships scoring 86 or below are rated unsatisfactory and are inspected more frequently (100 points is a perfect score). According to the CDC Web site, low inspection scores increase the likelihood of intestinal illness on ships. For our measure of *government regulator evaluations,* we obtained each ship's health and sanitation inspection score from http://www.cdc.gov/nceh/vsp/vsp.htm.

Physical aspects of the ship context. We obtained information on physical characteristics of the ships from several Web sites (Fielding: www.fieldingtravel.com, Get Cruising: www.getcruising.com; and Conde Nast Traveler: www.travel.epicurious.com). For *ship newness,* we used the year a ship was built. Ship length and number of cabins were highly correlated, so we standardized and averaged scores for *ship size* (α = .80).

Social aspects of the ship context. The cruise industry uses the passenger space ratio to indicate *spatial density* on a ship (passenger capacity divided by ship gross tonnage); we reversed these scores so that higher values reflected more crowding. For the *crew-to-passenger ratio,* we divided the number of crew members by the number of passengers that could cruise on the ship (higher values suggest more crew members per passenger).

Passenger demographics and controls. Cruise Opinion database respondents provided their *age* and *number of prior cruises.* We coded *gender* for all names that could reliably be classified as male or female (1 = "male," 2 = "female"). For respondents who provided only initials or gender-neutral names (such as "Chris"), we read written narratives for information that might indicate gender. This allowed us to code gender for 1,569 of our 1,676 respondents (93.6%). We also recorded information on the *length of cruise* (1 = "3–4 days," 2 = "7 days," 3 = "10 days," 4 = "12 days," 5 = "14 days," and 6 = "16 days," the longest cruise in our sample). For the handful of passengers who did not specify the length of their cruise, we determined cruise length by cross-referencing their cruise itinerary (destinations) with their specific cruise ship.

Finally, we obtained information from cruise Web sites on the average *cost per day* for cruising on each ship.

Analyses

For passenger evaluations for which we proposed cross-level effects, we used hierarchical linear modeling (HLM). HLM is more appropriate for examining cross-level relationships than is ordinary least squares regression analysis (OLS) because it explicitly models both individual- and group-level residuals and acknowledges that individuals in a group may be more similar to each other than to those in other groups (Bryk & Raudenbush, 1992). In contrast, OLS can violate statistical assumptions of independent observations (for instance, by disaggregating higher-level scores to lower-level units) or discard potentially meaningful lower-level variance (by aggregating lower-level scores to higher-level units; cf. Hofmann [1997]). After satisfying HLM prerequisites (systematic between-groups variance in dependent variables and significant variance in level 1 intercepts), we specified the HLM model for the three outcomes (overall evaluation, cruise activities evaluation, and cabin evaluation). The level 1 equation was *overall evaluation* = $\beta_{0j} + \beta_{1j}(age) + \beta_{2j}(gender) + \beta_{3j}(prior\ cruise) + \varepsilon_{ij}$. The level 2 equations were $\beta_{0j} = \gamma_{00} + \gamma_{01}(length\ of\ cruise) + \gamma_{02}(cost) + \gamma_{03}(ship\ age) + \gamma_{04}(ship\ size) + \gamma_{05}(spatial\ density) + \gamma_{06}(crew\text{-}to\text{-}passenger) + \mu_{0j}$; $\beta_{1j} = \gamma_{10} + \mu_{1j}$; $\beta_{2j} = \gamma_{20} + \mu_{2j}$; and $\beta_{3j} = \gamma_{30} + \mu_{3j}$. In the level 1 equation, we controlled for age, gender, and prior cruise experience. β_{0j} is the mean level of passenger overall evaluation for ship *j,* and ε_{ij} is the residual within-ship variance. The *t*-tests for the estimated parameters γ_{10}, γ_{20}, and γ_{30} in the level 2 equations indicate whether control variables have significant relationships with the dependent variable. To test whether hypothesized ship-level characteristics influenced passenger evaluations, we modeled the ship's mean level of passenger evaluation (β_{0j}) as a function of length of cruise (γ_{01}), cost per day (γ_{02}), age of ship (γ_{03}), size of ship (γ_{04}), spatial density (γ_{05}), and crew-to-passenger ratio (γ_{06}) and calculated *t*-statistics for these estimated parameters to test cross-level hypotheses. We computed the percentages of variance explained by ship-level predictors using the strategy outlined by Hofmann (1997).

For industry expert and government regulator evaluations (where we proposed that ship characteristics would influence ship-level evaluations),

TABLE 2
Means, Standard Deviations, and Intercorrelations among the Measures[a]

Variable	Mean	s.d.	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	41.62	11.90													
2. Gender[b]	1.45	0.50	−.15**												
3. Number of prior cruises	4.52	6.12	.21**	−.09**											
4. Cost per day	241.43	84.17	.25**	−.03	.21**										
5. Cruise length	1.95	0.58	.18**	−.04	.18**	.40**									
6. Ship newness[c]	1,989.50	9.90	−.04	.02	−.05	.26**	.07**								
7. Ship size	0.00	0.97	−.22**	.04	−.14**	−.25**	.00	.38**							
8. Spatial density	35.87	5.98	−.14**	.03	−.08**	−.67**	−.31**	−.34**	−.11**						
9. Crew-to-passenger ratio	0.44	0.06	.15**	−.03	.14**	.49**	.17**	−.10**	−.47**	−.46**					
10. Overall cruise evaluation	90.15	10.82	−.07**	.04	−.01	.11**	.05	.20**	.13**	−.14**	−.02				
11. Activities evaluation	87.79	7.91	−.15**	.07**	−.08**	−.01	−.06*	.22**	.25**	−.09**	−.13**	.62**			
12. Cabin evaluation	90.26	8.66	−.01	.06*	−.03	.23**	.07**	.31**	.10**	−.28**	.09**	.60**	.48**		
13. Industry expert evaluation	−0.01	0.90	.08**	−.00	.09**	.61**	.32**	.61**	.26**	−.64**	.22**	.21**	.16**	.34**	
14. Government regulator evaluation	92.27	4.26	.00	.01	.02	.12**	.06*	.36**	.14**	−.13**	−.06*	.16**	.13**	.15**	.34**

[a] Sample sizes (*n*'s) for each variable range from 1,507 to 1,676.
[b] Coding: 1, "male"; 2, "female".
[c] Higher values indicate newer ships.
* $p < .05$
** $p < .01$

TABLE 3
Results of Hierarchical Linear Modeling Analysis Predicting Passenger Evaluations[a]

Model	Overall Cruise Evaluation	Cruise Activities Evaluation	Cabin Evaluation
Level 1: Individual variables			
Age	−0.08**	−0.07**	−0.02
Gender	0.68	0.59	0.87*
Number of prior cruises	0.02	−0.02	−0.07
R^2	.01	.01	.01
Level 2: Ship variables			
Length of cruise	−0.18	−1.07	0.06
Cost of cruise	0.00	0.01	0.01
Ship newness	0.17**	0.11**	0.23**
Ship size	0.77	1.16**	0.10
Spatial density	−0.24*	−0.18*	−0.25**
Crew-to-passenger ratio	−9.94	−5.51	4.78
R^2	.59	.57	.73

[a]Unstandardized beta coefficients are shown. At the ship level, $n = 63$. At the individual level, $n = 1{,}526$.
* $p < .05$
** $p < .01$

we used hierarchical OLS regression. We used one-tailed tests of significance, given the directional nature of our predictions.

RESULTS

Table 2 reports means, standard deviations, and correlations for the cruise opinion responses. Comparison of our respondents to the 1998 *Cruise Industry Sourcebook* "Profile of the U.S. Cruise Industry" (http://www.cruising.org/index0.htm) showed our sample was identical in average age (42 years) to cruise industry "hot prospects" and that our proportion of first-time cruisers closely matched industry norms (31.5% versus 31.4%). We had slightly more male passengers (51% versus 46%) than industry averages (perhaps owing to our use of the Internet as a source of data). The positive passenger overall evaluations in our sample ($\bar{x} = 90$) are consistent with industry data showing that the cruise experience "consistently exceeds expectations on a wide range of important vacation attributes" for most customers, according to the "Cruise Industry Overview" in the *Cruise Industry Sourcebook.* George Madsen, publisher of *Garth's Profiles of Ships,* a semiannual travel agent publication (gmadsen@profilesofships.com), confirmed this statement. Government regulator ratings are also high ($\bar{x} = 92.27$, s.d. = 4.26).

Missing data reduced the number of passengers and ships in our final sample. For instance, government regulators did not evaluate 21 ships because the CDC only inspects cruise ships that embark and disembark passengers at U.S. ports and that include international ports of call. To summarize, the HLM analyses included 1,526 passengers (over 90 percent of the original sample) who had cruised on 63 ships (mean number of passengers per ship = 24.22; range = 9–72). The OLS regressions (which included only ship-level data) included 68 ships evaluated by both industry experts and government regulators. Tables 3 and 4 respectively present the HLM and OLS results.

Effects of Physical and Social Aspects of Ship Context on Evaluations

Our first set of hypotheses (1a–3) predicts ship age will influence evaluation of services for three different perspectives. The results of the HLM analyses (see Table 3) supported Hypotheses 1a–1c, showing that newer ships received more positive overall evaluations and more positive evaluations of cruise activities and cabin accommodations from passengers. Results also supported Hypotheses 2 and 3, showing that newer ships were rated more positively by industry experts and government regulators (see Table 4).

Our second set of hypotheses (4a–6) focuses on ship size. HLM analyses provided partial support for Hypotheses 4a and 4b, showing that size was positively related to passenger evaluations of cruise activities (4b) but was unrelated to overall evaluations (4a). For the other evaluator perspectives, we noted that ship size had a significant effect on industry expert evaluations (Hypothesis 5) but no effect on government regulator evaluations (Hypothesis 6).

Our third series of hypotheses (7a–9) focuses on spatial density. Hypotheses 7a–7c predicts

TABLE 4
Results of Hierarchical Regression Analysis Predicting Industry Expert and Government Regulator Evaluations

	Industry Expert Evaluation			Government Regulator Evaluation		
Block	**β**	***t***	**ΔR^2**	**β**	***t***	**ΔR^2**
1. Controls			.38***			.03
Length of cruise	0.05	0.38		0.11	0.67	
Cost per day	0.58	4.54***		0.08	0.47	
2. Physical and social context			.35***			.24**
Ship newness	0.35	4.12***		0.40	2.82**	
Ship size	0.22	2.44*		0.02	0.15	
Spatial density	−0.29	−2.48*		0.02	0.09	
Crew-to-passenger ratio	−0.11	−1.05		−0.21	−1.25	
Overall R^2			.73			.26
Overall *F (df)*		27.83*** (6, 61)			3.65** (6, 61)	

* $p < .05$
** $p < .01$
*** $p < .001$

that spatial density will be negatively related to passenger overall evaluations, cruise activities, and cabin accommodations. The results, shown in Table 3, supported all three of these predictions. Higher spatial density led to lower overall evaluations and lower evaluations of cruise activities and passenger cabins. For the other evaluator perspectives, the results in Table 4 supported Hypothesis 8 (higher spatial density led to lower industry expert evaluations) but did not support Hypothesis 9 (high spatial density did not lead to lower government regulator evaluations). Our fourth and final series of hypotheses (Hypotheses 10–12) focuses on crew-to-passenger ratio. Tables 3 and 4 show that none of these predictions were supported. This finding contrasts with the widespread effects of our other characteristic of social context, spatial density.

DISCUSSION

The central goals of our study were to (1) expand the theoretical conceptualization of services, (2) focus on a unique, understudied context (captive, intensive service settings) highlighted by our expanded categorization of services, and (3) assess whether physical and social aspects of a service context differentially affect evaluators with different perspectives.

In terms of our first goal, the conceptual framework in Table 1 makes a theoretical contribution by extending Gutek's (1995) differentiation of services. We identified the intensity of a service (whether service delivery involves one primary service or multiple services) and the captivity of a service (whether it is difficult to leave the service setting once service delivery has started) as key aspects of service delivery that provide a finer-grained conceptualization. Although our data did not allow us to compare cells in the framework, we hope the taxonomy stimulates future research on other cells. In addition, future work could further elaborate on forms of captivity and distinguish physical captivity (prevalent on cruise ships) from psychological captivity (service recipients' feelings of financial or social pressure to continue a service relationship). In terms of our second goal, focusing on high-intensity, high-captivity services provides needed research on a unique and understudied context. Although our results have special relevance to cruise ships, they also have implications for other high-intensity, high-captivity settings (as we discuss below). As for our third goal, our analytic results provide insights into the effects of context on three different evaluator perspectives: passengers, government regulators, and industry experts. Tables 3 and 4 reveal that although results support seven of the nine hypotheses for physical context, they support only four of eight predictions for social aspects (all involving spatial density, not the crew-to-passenger ratio). In sum, both physical and social elements of context are important for understanding evaluator perspectives.

Evaluators' Different Perspectives

An important contribution of our study is the finding that the effects of aspects of physical and

social context differ by type of evaluator. Our results support the assertion that government regulators focus narrowly on health and sanitation and use objective standards such as ship newness (an objective physical feature) in making their assessments. In contrast to the government regulators, industry experts adopt a broader focus and compare ships by assuming the role of typical passengers when making their assessments. We speculate that this broader focus explains the greater number of physical and social factors that influence their ratings. More expensive cruises, newer ships, bigger ships, and more spacious ships (lower spatial density) improve expert evaluations. Hardly a surprising pattern, yet it exists only for industry experts and raises questions about the "expertise" that such experts actually provide. It is consistent with the diverse goals of passengers that the results for their evaluations are the most complex. Both the physical aspects (ship newness) and social aspects (spatial density) of context matter, and each influences numerous passenger judgments.

Practical Implications and Future Research

Our results have interesting implications for service firms as they try to manage the judgments of different constituencies. To best manage government regulator evaluations, our results suggest, organizations should put primary emphasis on physical aspects of service contexts. Thus, investing in new physical facilities (new ships) is critical for health and sanitation because it is easier to maintain sanitation standards in newer facilities, and newer space allows for updated food storage, cleanliness, and sanitation systems. In addition, it is possible that renovations involving sanitation systems may also improve government inspector scores (for instance, on safety or sanitation) and may be especially relevant in service situations where customers are held captive in such a way that poor evaluations trigger strong, negative public reactions (such as fear of illness). To follow up on our framework, presented in Table 1, we recommend future research that contrasts the effects of physical context on regulatory evaluations in settings that vary in intensity (such as a taxi ride versus a plane flight or a dental cleaning versus an annual physical) or captivity (a walk-in haircut versus giving blood or an annual physical versus an extended hospital stay). (Other exemplars can be seen in Table 1.)

In contrast to the practical implications for regulators, which emphasize physical context, the practical implications for industry expert ratings are that physical *and* social aspects of context are important. We suggest that industry experts are attuned to the relative spaciousness of ships (more than government regulators) because they spend more time on the ships with passengers than do the inspectors, who do their work in port. Overall, results for industry experts show strong support for our predictions because the model explained over 70 percent of the variance in their evaluations. Turning again to Table 1 suggests that it is not surprising that size, newness, and spatial density are important in high-captivity settings where confinement and the difficulty of leaving make context salient. We recommend that future research compare the effects of the physical characteristics of space in different cells of the taxonomy. For example, although "big is better" on cruise ships, and largeness often has positive symbolic value (such as the status benefits of a larger limousine), in restaurants or retail spaces (where captivity is lower), largeness may detract from the exclusivity, intimacy, and romantic appeal of the setting. It also would be interesting to see if crowding (social density) becomes more critical as the intensity and captivity of services increase.

Our last set of practical implications focuses on managing customer evaluations. As they did with the industry experts, the results suggest both physical and social aspects of context are important to this audience. Large, new spaces with low spatial density enhance both the overall and specific evaluations of customers. Although this finding may seem straightforward, building new facilities and maintaining large facilities is expensive and has implications for firm profitability. In addition, it is not always easy to manage spatial density because service providers often cannot control customer access to service settings. Density is specific to both space and time. Thus, what is spacious during typical or average conditions may become crowded during peak periods (for instance, golf and tennis facilities may be in high demand in the morning, deck chairs in the afternoon, and entertainment and slot machines in the evening) or at specific times of the year (such as the winter holidays or school vacations). We recommend a somewhat paradoxical approach to managing space—especially at peak times. Currently, retailers put out *more* merchandise and displays (reducing space for customers) when customer demand is high. Perhaps fewer displays or the use of more spatially efficient storage systems could reduce spatial density while enhancing customer enjoyment and purchases. Thus, it is important for service settings to be new and busy, but not crowded.

Again, we refer to our framework for examples of where our findings might be relevant and to suggest

possible contrasts for future research. When customers purchase multiple services such as complete vacation packages or extended health spa stays, the physical and social aspects of the high-intensity context may be more salient than these aspects are when they purchase only one primary service (for instance, a haircut or single meal). When customers enter service settings with high captivity (such as a nursing home, boarding school, or week-long wilderness retreat), the quality of accommodations (newness, size, and social density) may be more salient than it will be in low-captivity settings (an annual physical or an estate planning appointment). These are interesting questions for future research.

Our last specific comment relates to the crew-to-passenger ratio. We were surprised to find no positive results for this ratio, since cruise industry experts view service ratios as critical to success (e.g., Ward, 1999). Perhaps large numbers of service providers become a nuisance factor in captive, intensive service relationship contexts. For example, if customers are constantly pressured to make additional purchases (drinks, food, clothing, equipment, souvenirs, side trips), additional service providers may be an annoying and ongoing source of interruptions.

In sum, our results show convergence and divergence in evaluator perspectives. The most consistent finding is the positive role of newness. This observation makes sense since the age of physical facilities is an important indicator of the tangible and intangible benefits offered by service firms. We suggest that future research contrast absolute age with date of renovation. Perhaps renovation is an effective and less expensive alternative. Perhaps refurbishment enhances some evaluator judgments more than others. For instance, customers and industry experts could view remodeled restaurant facilities in a historic building favorably, but regulators could still provide low ratings on health and sanitation as a consequence of the constraints of older space. Comparison across raters also raises an important issue not addressed in existing research that focuses on one evaluator perspective only. Although our multiple source data allowed us to contrast findings for three important constituencies, the practical implications of these findings are complex. This is because it is difficult to satisfy multiple publics simultaneously. For example, although high demand signals positive worth to industry experts (providing tangible evidence of value), it also can detract from customer satisfaction (leading to competition for slot machines on rainy days and for deck chairs when it is sunny) and government regulator evaluations (promoting crowding and deterioration of sanitation).

Strengths and Limitations

This study used an unusual setting (captive, intensive service relationships) to examine effects of context on evaluations of service relationships from three different perspectives. Our research design has several strengths, including multiple sources of data (customer responses, printed travel resources, government documents, and various Web resources). The use of multiple sources and evaluator perspectives ensured that no common method variance linked physical and social aspects of context with evaluations. Thus, our design provided a stringent test of the hypotheses.

At the same time, future research should examine the generalizability of our findings to other service settings. Although we focused specifically on cruise ships as one example of a captive, intensive relationship context, Table 1 shows other service contexts that have similar characteristics, including in-home day care, daily elder care, nursing homes, and extended hospital stays. In addition, many educational, personal development, and recreational experiences also have high captivity and intensity, such as when students receive room and board from university housing, when corporate executives hold off-site retreats, or when people participate in outfitted camping trips or white water rafting expeditions (Arnould & Price, 1993). Thus, we encourage others to examine whether relationships demonstrated in our cruise ship setting are replicated in other captive, intensive service relationship settings.

Our use of data from Web sites also could have influenced results. To participate in the survey on the Cruise Opinion site, passengers had to access the Internet. In the Internet's early years, its users were better educated than the general population. Recent Pew Research Center reports, however, have shown that new users are less affluent and include more women than men (Gearan, 1999), so earlier differences in education may have diminished. Yet it would be useful to examine physical and social aspects of context in more traditional service settings, such as retail sales or personal service. Another potential response bias is the possibility that the Web site respondents had strong opinions (either positive or negative); perhaps the "silent majority" of cruise consumers is underrepresented. Alternatively, novice or experienced cruisers may be overrepresented (vis-à-vis typical cruise passengers). Although our follow-up data addressed these concerns, we cannot completely rule out these limitations.

Finally, we noted the limitations of archival data and recommend future research that uses other measures of physical or social context. Physical characteristics could include tempera-

ture, noise, and lighting, or customer perceptions of the size and age of a setting. We intentionally measured ship age, an objective feature of the context, but observer ratings of physical context would also provide valuable information. For social characteristics, our measures of spatial density and the crew-to-passenger ratio are imperfect because they rely on the assumption that ships sail at full capacity (the industry standard is two passengers per cabin). Although industry data from the *Cruise Industry Sourcebook* and conversations with other industry experts confirmed that most ships routinely sailed at close to capacity during the time of our study, some ships might have sailed at less than capacity.

To summarize, although our research has limitations, its results show differing effects of physical and social context on evaluations of captive, intensive service relationships from three different perspectives (customers, industry experts, and government regulators). We recommend that future research extend this initial examination of captive, intensive service relationships by continuing to assess physical and social aspects of context, using a multiple evaluator perspective. We also recommend that future research contrast effects for different types of services in the expanded conceptual framework presented in Table 1.

REFERENCES

Arnould, E. J., & Price, L. L. 1993. River magic: Extraordinary experience and the extended service encounter. ***Journal of Consumer Research,*** 20: 24–45.

Barnard, C. I. 1938. ***The functions of the executive.*** Cambridge, MA: Harvard University Press.

Baron, R. A. 1993. Affect and organizational behavior: When and why feeling good (or bad) matters. In J. K. Murnighan (Ed.), ***Social psychology in organizations: Advances in theory and practice:*** 63–88. Englewood Cliffs, NJ: Prentice-Hall.

Barton, J. 1998. Finding a berth or joining a debate online. ***New York Times on the Web:*** http://www.nytimes.com/library/cyber/week/020198cruise.html, February 1.

Bitner, M. 1990. Evaluating service encounters: The effect of physical surroundings and employee responses. ***Journal of Marketing,*** 54: 69–82.

Bowen, D. E., Gilliland, S. W., & Folger, R. 1999. HRM and service fairness: How being fair with employees spills over to customers. ***Organizational Dynamics,*** 27(3): 7–23.

Bryk, A. S., & Raudenbush, S. W. 1992. ***Hierarchical linear models: Applications and data analysis methods.*** Newbury Park, CA: Sage.

Cappelli, P., & Sherer, P. D. 1991. The missing role of context in OB: The need for a meso-level approach. In L. L. Cummings & B. M. Staw (Eds.), ***Research in organizational behavior,*** vol. 13: 55–110. Greenwich, CT: JAI Press.

Cyert, R. M., & March, J. G. 1963. ***A behavioral theory of the firm.*** Englewood Cliffs, NJ: Prentice-Hall.

Deephouse, D. L. 2000. Media reputation as a strategic resource: An integration of mass communication and resource-based theories. ***Journal of Management,*** 26: 1091–1112.

Dickinson, R. H., & Vladimir, A. N. 1997. ***Selling the sea: An inside look at the cruise industry.*** New York: Wiley.

Fischer, E., Gainer, B., & Bristor, J. 1997. The sex of the service provider: Does it influence perceptions of service quality? ***Journal of Retailing,*** 73: 361–382.

Fombrun, C. 1996. ***Reputation: Realizing value from the corporate image.*** Boston: Harvard Business School Press.

Gearan, A. 1999. Who's on the Internet now? Joan Average. ***Lansing State Journal,*** January 15: 5B.

Ginzel, L. E., Kramer, R. M., & Sutton, R. I. 1993. Organizational impression management as a reciprocal influence process: The neglected role of the organizational audience. In L. L. Cummings & B. M. Staw (Eds)., ***Research in organizational behavior,*** vol. 15: 227–266. Greenwich, CT: JAI Press.

Gutek, B. A. 1995. ***The dynamics of service.*** San Francisco: Jossey-Bass.

Gutek, B. A., Bhappu, A., Liao-Troth, M., & Bennett, C. 1999. Distinguishing between service relationships and service encounters. ***Journal of Applied Psychology,*** 84: 218–233.

Hall, E. T. 1966. ***The hidden dimension.*** New York: Doubleday.

Hofmann, D. A. 1997. An overview of the logic and rationale of hierarchical linear models. ***Journal of Management,*** 23: 723–744.

Jones, T. O., & Sasser, W. E. 1995. Why satisfied customers defect. ***Harvard Business Review,*** 73(6): 88–99.

Marti, B. E. 1995. The cruise ship vessel sanitation program. ***Journal of Travel Research,*** 33: 29–38.

Mitchell, R. K., Agle, B. R., & Wood, D. J. 1997. Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. ***Academy of Management Review,*** 22: 853–886.

Oldham, G. R., Cummings, A., & Zhou, J. 1995. The spatial configuration of organizations: A review of the literature and some new research directions. In G. R. Ferris (Ed.), ***Research in personnel and human resource management,*** vol. 13: 1–37. Greenwich, CT: JAI Press.

Oldham, G. R., & Rotchford, N. L. 1983. Relationships between office characteristics and employee reactions: A study of the physical environment. ***Administrative Science Quarterly,*** 28: 542–557.

Pfeffer, J. 1982. ***Organizations and organization theory.*** Boston: Pitman.

Pfeffer, J. 1998. Understanding organizations: Concepts and controversies. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), ***Handbook of social psychology*** (4th ed.): 733–777. New York: McGraw-Hill.

Quinn, R., & Rohrbaugh, J. 1983. A spatial model of effectiveness criteria. ***Management Science,*** 29: 363–377

Roethlisberger, F. J., & Dickson, W. J. 1939. ***Management and the worker.*** Cambridge, MA: Harvard University Press.

Rothman, R. A. 1998. ***Working: Sociological perspectives.*** Englewood Cliffs, NJ: Prentice-Hall.

Scott, S. G., & Lane, V. R. 2000. A stakeholder approach to organizational identity. ***Academy of Management Review,*** 25: 43–62.

Seiders, K., & Berry, L. L. 1998. Service fairness: What it is and why it matters. ***Academy of Management Executive,*** 12(2): 8–20.

Stanton, J. M. 1998. An empirical assessment of data collection using the internet. ***Personnel Psychology,*** 51: 709–724.

Stokols, D. 1995. The paradox of environmental psychology. ***American Psychologist,*** 50: 821– 837.

Ward, D. 1999. ***Berlitz 1999 complete guide to cruising and cruise ships.*** Princeton, NJ: Berlitz Publishing.

Donald E. Conlon *(conlon@msu.edu)* is a professor of management in the Eli Broad Graduate School of Management at Michigan State University. He received his Ph.D. from the Organizational Behavior Group at the University of Illinois at Urbana-Champaign. His research focuses on organizational justice, conflict and negotiation, and the role of project completion in organizational decision making.

Linn Van Dyne *(vandyne@msu.edu)* is an associate professor of management in the Eli Broad Graduate School of Management at Michigan State University. She received her Ph.D. from the University of Minnesota in strategic management and organizations. Her research focuses on proactive behaviors (e.g., helping, voice, and minority influence) and the effects of work context, roles, and groups on employee behavior.

Morgan R. Milner *(milnermo@msu.edu)* is a doctoral candidate in the Department of Management at Michigan State University. His research interests include the study of psychological contracts and impression management.

Kok Yee Ng *(akyng@ntu.edu.sg)* is an assistant professor of management in the Division of Strategy, Management & Organization at the Nanyang Technological University. She received her Ph.D. from Michigan State University, with a major in organizational behavior. Her research interests are in cross-cultural I/O psychology and international organizational behavior, focusing in areas such as cultural intelligence, leadership, extrarole behavior, and teams.