

Gratitude and Gratuity:
A Meta-Analysis of Research on the Service-Tipping Relationship

Michael Lynn
and
Michael McCall*

Journal of Socio-Economics 29 (2000) 203-214

* Michael Lynn is an associate professor at the School of Hotel Administration, Cornell University, Ithaca, NY 14853-6902. Michael McCall is an associate professor at the School of Business, Ithaca College, Ithaca, NY 14850. This research was supported by a grant to the first author from the Center for Hospitality Research at the School of Hotel Administration at Cornell University. The second author was supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (R29-AA10387-01). The authors would like to thank Orn Bodvarsson for making his raw data available to us and for providing us with copies of several unpublished student research papers.

Gratitude and Gratitude:
A Meta-Analysis of Research on the Service-Tipping Relationship

ABSTRACT

The relationship between tip size and evaluations of the service was assessed in a meta-analysis of 7 published and 6 unpublished studies involving 2,547 dining parties at 20 different restaurants. Consistent with theories about equity motivation and the economic functions of tipping, there was a positive and statistically significant relationship between tip size and service evaluations. However, that relationship was much smaller than is generally supposed. The confounding effects of customer mood and patronage frequency as well as the reverse-causality effects of server favoritism toward big tippers were all examined and shown to be insufficient explanations for the correlation between tipping and service evaluations. These findings suggest that tippers are concerned about equitable economic relationships with servers, but that equity effects may be too weak for tip size to serve as a valid measure of server performance or for tipping to serve as an effective incentive for delivering good service.

Gratitude and Gratitude:

A Meta-Analysis of the Service-Tipping Relationship

Consumers of hospitality and travel services often give sums of money above and beyond the contracted prices of those services to the workers who have served them. In the United States, these gifts (called “tips”) amount to approximately \$16 billion a year (Seligman, 1998). Tipping is an interesting behavior because tips are voluntary payments given after services have been rendered. Consumers rarely aspire to pay more than necessary for goods and services. In order to get the most from their limited resources, consumers usually try to obtain things for the lowest available price. Tipping represents a multi-billion dollar exception to this general rule. It is an exception that raises questions about why people tip.

When asked why they leave tips, people most often reply that they tip to reward workers for services rendered (Speer, 1997). This response is consistent with theories about the economic functions of tipping. Economists believe that tipping exists because it is the most efficient way of monitoring and rewarding the efforts of service workers (Bodvarsson & Gibson, 1994; Hemenway, 1984; Jacob & Page, 1980). The intangible and highly customized nature of many services makes it difficult for firms to monitor and control the quality of services delivered by their employees (Shamir, 1984; Zeithaml, Berry & Parasuraman, 1988). Tipping is thought to be a way of enlisting the customers’ help in performing these quality control functions.

The idea that people tip as a reward for good service is also consistent with equity theory (Adams, 1965; Walster, Berscheid & Walster, 1973). According to this theory, people are socialized to feel anxiety or distress when their relationships with others are inequitable. A relationship is inequitable when the participants’ outcomes from the relationship are disproportionate to their respective inputs to the relationship. Services and tips are inputs and outcomes in exchange relationships between service workers and their customers. Since inequitable relationships are distressing, service customers should

attempt to maintain equity in their relationships with servers by leaving larger tips when they receive better service (Lynn & Grassman, 1990; Lynn & Graves, 1996; Snyder, 1976).

Despite the apparent obviousness of the "reward for service" explanation for tipping, there are good reasons to question its validity. First, researchers have demonstrated that people are poor at identifying the causes of their own behavior (Nisbett & Wilson, 1977), so consumers' reports that they tip as a reward for good service must be regarded with skepticism. Second, researchers have found that people are more concerned with their own outcomes than with others' outcomes in commercial exchanges (Oliver & Swan, 1989), so equity motivations may be weak in the commercial exchanges between servers and tippers. Finally, many tipped employees are paid a substandard wage and this fact has created social pressures on consumers to tip even when the service is bad (see May, 1978).

Several researchers have tested the "reward for service" explanation for tipping by examining the relationship between restaurant tip sizes and evaluations of the service. If tipping is motivated by a desire to reward servers for good service, then higher ratings should be associated with larger tips. Unfortunately, many tests of this hypothesis are unpublished and those tests that have been published are scattered across several disciplines -- i.e., economics, hospitality management, and social psychology. Moreover, those tests have produced conflicting results. Two studies have found significant positive relationships between tip sizes and service evaluations (Lynn & Grassman, 1990; Lynn & Graves, 1996), but many others have not (Bodvarsson & Gibson, 1994; Crusco & Wetzel, 1984; Lynn, 1988; Lynn & Latane, 1984; May, 1978). The present study takes another look at this hypothesis by bringing together published and unpublished studies in a meta-analysis of research on the service-tipping relationship. It also assesses several alternative explanations for this relationship.

METHOD

Identification of Studies

An attempt was made to identify published and unpublished studies of restaurant tipping that used dining parties as the units of analysis in a between subjects examination of the relationship between tip

size and evaluations of the service. Studies falling within this domain were identified in several ways. First, computerized searches of ABI Inform, ERIC, Dissertation Abstracts, Psych Abstracts and the Cornell Hospitality Database were conducted. Second, the references of the already identified studies were examined for citations of prior studies. Finally, the authors of published studies were contacted and asked for any additional studies they had conducted. A total of 13 studies (7 published and 6 unpublished) were identified and retained for analysis in this review (see Table 1). Two additional studies were identified but excluded from the review. One unpublished study was excluded because we had serious questions about the authenticity of the student-collected data. A second unpublished data set was excluded because it is being prepared for separate publication elsewhere and the authors did not want to jeopardize the publication of that data by including it in this review.

Unit of Analysis

Meta-analysts usually use studies as their units of analysis. However, restaurants are a more appropriate unit of analysis for this meta-analysis, because customer expectations of service vary across restaurants and this makes cross-restaurant comparisons of service ratings less meaningful than within-restaurant comparisons. In addition, consumers' tipping propensities vary across restaurants and this makes between-restaurant tests of the service-tipping relationship less sensitive than within-restaurant tests. Furthermore, the social function of tipping as an incentive/reward for service depends on the within-restaurant relationship between tip size and service evaluations more than on their between-restaurant relationship. For all of these reasons, restaurants were selected as the unit of analysis in this review, meaning that separate effects were obtained from each restaurant in each study being reviewed. The effects from different restaurants were independent in the sense that they were based on different subjects or dining parties, so they meet the requirements of the fixed-effects, meta-analytic, significance tests reported below.

Re-Analysis of Study Level Data

We were able to obtain the raw data from 6 of 7 published studies and 2 of 6 unpublished studies by contacting the authors of published research articles on tipping. This data was re-analyzed in order to

maximize the comparability of analyses and effects across studies as well as to provide tests of relationships that were not reported by the original investigators. In these re-analyses, tips as a percentage of bill size were used as the dependent measure. For studies with data from multiple restaurants, separate analyses were conducted on the data from each restaurant.

Effects and Statistics

The primary effect examined in this meta-analysis was the relationship between tip sizes (percent tips or bill-adjusted residual tips) and service evaluations. Tests of this relationship were obtained from each study in the review. In most cases, zero-order correlations were obtained. However, in a few cases where the zero-order correlations were not available, partial-correlations between tip sizes and service evaluations that statistically controlled for one or more other variables were obtained and used in the analyses of this relationship.

The secondary effects examined in this meta-analysis included:

(1) the correlation between tip sizes and evaluations of the food, (2) the correlation between customers' evaluations of the service and their evaluations of the food, (3) the partial correlation between tip sizes and service evaluations after statistically controlling for food evaluations, (4) the difference between the correlation of tip size with service evaluations and the correlation of tip size with food evaluations, (5) the zero-order (or , in some cases, partial) correlation between patronage frequency and tip sizes, (6) the correlation between patronage frequency and service evaluations, (7) the partial correlation between tip sizes and service evaluations after statistically controlling for patronage frequency (and, in some cases, other variables), and (8) the interaction effect of patronage frequency and service evaluations on tip size. Tests of these relationships were obtained from all of the studies in the review for which the appropriate primary-level data and/or analyses were available.

Separate tests of the above relationships were obtained for each restaurant in each of the relevant studies. The only exception was a study by Crusco and Wetzel (1984) that combined data from two restaurants -- only one test of the service-tipping relationship could be obtained from this study. In cases where evaluations of multiple dimensions of service were obtained from each subject, an average of the

different service ratings was calculated and used in tests of the effects involving service evaluations. Similarly, in cases where evaluations of multiple dimensions of the food were obtained from each subject, an average of the different food evaluations was calculated and used in tests of the effects involving food evaluations. For studies with continuous measures of patronage frequency, log transformations of these measures were obtained whenever possible (to deal with outliers) and used in tests of effects involving patronage frequency.

For each restaurant-level test of a relationship in the meta-analysis, we calculated two statistics – a correlation coefficient r that reflects the size of the effect and a z -score that reflects the statistical significance of the effect. These statistics were calculated using formulas specified in Mullen (1989) and Rosenthal (1991). The information used in these calculations came from our re-analyses or (if the raw data was not available) from the original reports of the studies.

Coding of Study Characteristics

We categorized each measure of service quality used in the meta-analysis as either: (1) a customer rating of service on a multi-item scale, (2) a customer rating of service on a single-item scale, (3) a server or third-party rating of service, or (4) a rating of the general dining experience (this category included indices that combined ratings of service with ratings of other aspects of the dining experience). In addition, we categorized each measure of patronage frequency used in the meta-analysis as either binomial or continuous. Two judges made each coding decision - the senior author did the initial coding and highlighted relevant information in the studies' methods and results sections, which he then gave to a second judge for coding. The judges agreed on all coding decisions.

RESULTS AND DISCUSSION

The data in this meta-analysis were analyzed using formulas and procedures advocated by Mullen (1989) and Rosenthal (1991). The effect sizes and z -scores from each of the restaurants in each study being reviewed were weighted equally in these analyses.

The Service-Tipping Relationship

If consumers tip in order to reward servers, then tips should be positively related to evaluations of the service. The studies in this review provided 23 tests of this hypothesis, involving 2,547 dining parties at 20 different restaurants (see Table 2). These tests produced a mean effect size r of .11 and a combined z of 5.26 (one-tailed $p < .0001$). It would take another 212 studies averaging null results to bring the combined significance of this effect below the .05 level. A marginally significant test of effect size heterogeneity, $\chi^2(22) = 33.89$, $p < .06$, indicated that the variability in effect sizes was unlikely to be produced by chance alone, so additional analyses were conducted to test some potential sources of this variability.

To see if the effect sizes varied with the measures of service quality employed in different studies, separate analyses were performed for those effects based on different types of service measures. The results of these analyses are summarized in Table 3. There are three things worth noting about these results. First, the effects from studies using customer ratings of service on multi-item scales were significantly larger than the effects from studies using less valid and/or reliable measures of service quality (\bar{r} 's = .22 vs .11, .04, and .05: all z contrasts > 1.76 , one-tailed $p < .04$). However, even in these studies, the average service effect was quite small -- accounting for only about five percent of the variability in bill-adjusted tip sizes. Second, tipping was not significantly related to servers' or third-parties' evaluations of the service ($\bar{r} = .04$, $z = .95$, $p > .50$). This is important because servers are likely to be motivated by tips only if they perceive a strong relationship between their tip receipts and their service delivery (as self-evaluated). Finally, none of the effect sizes within each classification of the service measure were significantly heterogenous (all p 's $> .12$). This means that once methodological characteristics were controlled for, the relationships between tip sizes and service ratings were fairly consistent across restaurants and studies.

Overall, the above results indicate that there is a small, but reliable and positive relationship between service evaluations and tip sizes and that the relationship generalizes across several types of

restaurants. This relationship is consistent with the "reward for service" explanation for tipping outlined in the introduction. However, there are many other potential explanations for the relationship. Several of these alternative explanations are evaluated in the paragraphs that follow.

Customer Mood

The service-tipping relationship reported above is correlational, so it may be a spurious by-product of some third variable's effects on both tips and service evaluations. One likely candidate for such a confound is customer mood. Researchers have found that environmental determinants of mood affect both consumers' evaluative judgements (see Gardner, 1985) and their tipping behaviors (Cunningham, 1979; Fisher, 1992; Rind, 1996). Thus, it is plausible that the service-tipping relationship is due to the effects of customers' moods on service evaluations and tips rather than to a direct effect of service quality on tip size. Mood effects are not specific to evaluations of service, so this explanation suggests that tip sizes will be related to a variety of customer evaluations. Specifically, customers' moods should affect evaluations of food as well as service, so the relationship between tip sizes and food evaluations should be comparable to that between tip sizes and service evaluations.

Tests of the preceding hypothesis were available from five of the studies in this review. In these studies, which involved over 1,250 dining parties at 12 different restaurants, food ratings were positively correlated with service ratings, $\bar{r} = .40$, $z = 14.50$, one-tailed $p < .0001$, and with tip sizes, $\bar{r} = .06$, $z = 1.91$, one-tailed $p < .03$. However, service ratings were more strongly related to tip sizes than were food ratings, $\bar{r} = .14$ vs $.06$. Restaurant-level t-tests of the difference between these dependent correlations combined to produce a z of 2.76 (one-tailed $p < .01$), so the difference is reliable. Furthermore, the service tipping relationship remained significant after statistically controlling for food evaluations, $\bar{r} = .13$, $z = 4.42$, one-tailed $p < .0001$. These findings indicate that the process underlying the service-tipping relationship is specific to service evaluations and that customer mood is an unlikely explanation for this relationship.

Patronage Frequency

A second potential confound that may explain the service-tipping relationship is patronage frequency. Researchers have found that regular customers of a restaurant tip more than infrequent patrons (Bodvarsson & Gibson, 1991; Lynn & Grassman, 1990). Regular customers may also rate service more highly than do infrequent patrons, because people who like the service at a restaurant are more likely to become regular patrons and/or because servers are likely to deliver better service to customers with whom they are familiar. Thus, it is possible that the service-tipping relationship is due to the effects of customer' patronage frequency on tips and service evaluations rather than to a direct effect of service quality on tip size.

Eight studies in this review provided 18 tests each of patronage-frequency effects on tipping and service effects on tipping after statistically controlling for patronage frequency. Five of these studies also provided 12 tests of the correlation between patronage frequency and service evaluations. Meta-analyses of these tests indicated that patronage frequency was significantly, positively related to tip sizes, $\bar{r} = .08$, $z = 3.46$, one-tailed $p < .0004$, and marginally related to service evaluations, $\bar{r} = .03$, $z = 1.35$, one-tailed $p < .09$. However, both relationships were weak (\bar{r} 's $< .10$) and the relationship between patronage frequency and service evaluations was significantly heterogeneous, $\chi^2(11) = 22.40$, $p < .03$. These findings suggest that patronage frequency is neither a strong or consistent confound of the service-tipping relationship. Reinforcing this conclusion, the relationship between tip sizes and service evaluations remained significant after statistically controlling for patronage frequency, $\bar{r} = .10$, $z = 4.01$, one-tailed $p < .0001$. This latter finding held true even when only those tests ($n = 6$) involving continuous measures of patronage frequency were analyzed, $\bar{r} = .19$, $z = 4.42$, one-tailed $p < .0001$. Clearly, the effects of service observed in this meta-analysis are not dependent on the confounding effects of patronage frequency.

Server Favoritism

A third alternative explanation for the correlation between tip sizes and service evaluations is that tip size may affect service quality. Servers may deliver better service to those customers known to be generous tippers. If such a "server favoritism" process does underlie the service-tipping relationship, then that relationship should be stronger for regular customers, whose tipping habits are better known, than for infrequent customers.

Five studies in this review provided primary-level tests of 12 interactions between patronage frequency and service ratings. Two tests of this interaction were excluded from analysis because small cell sizes ($n < 12$) at one level of the binomial measure of patronage frequency made the tests meaningless. The remaining tests, involving 1,094 dining parties at 10 restaurants, produced a mean r of .02 and a non-significant combined z of 0.84 ($p > .10$). A test of effect size heterogeneity was also non-significant, $\chi^2(9) = 12.50$, $p > .10$. These findings indicate that patronage frequency does not moderate the relationship between tip sizes and service evaluations and that server favoritism is unlikely to account for this relationship.

GENERAL DISCUSSIONS

In summary, this meta-analysis found a statistically significant, positive relationship between tip size and evaluations of service. The relationship was not moderated by patronage frequency and it remained significant after statistically controlling for patronage frequency and food evaluations. These findings are inconsistent with explanations based on the confounding effects of customer mood and patronage frequency as well as with an explanation based on the reverse-causality effect of server favoritism toward big tippers. The results are consistent with the idea that consumers use tips to reward waiters and waitresses for services rendered. This supports the equity theory notion that people are concerned about fairly compensating others for their inputs to exchange relationships. It also suggests, in contrast to some previous research (Oliver & Swann, 1989), that equity concerns about the fairness of others' outcomes apply to consumer transactions as well as to other social exchanges.

Although the average relationship between tip size and service evaluations was statistically significant in this review, it was also quite small – accounting for less than two percent of the variability in tip percentages. The weakness of this relationship belies many consumers’ beliefs that their primary motivation for tipping is to reward servers for good service (Speer, 1997). It also raises questions about the efficacy of tipping as an incentive for delivering good service and about the validity of tip size as a measure of server performance. These latter two issues are discussed below.

Tipping as an Incentive

All but one of the 23 service effects in this review were smaller than the .30 correlation that Cohen (1992) described as an effect size “likely to be visible to the naked eye of a careful observer” (p. 156). This fact raises questions about the incentive value of tipping, because it suggests that servers are unlikely to perceive the relationship between the services they give to tables and the tips those tables leave in return (Lynn & Graves, 1996). Of course, servers may be able to notice the service-tipping relationship at the server level of aggregation (Lynn & Simons, 2000). If servers compare their service delivery and tip receipts with those of other servers, they may notice that better servers tend to earn more tips by the end of the evening. Such a server-level relationship between service and tipping could be noticeable even when the customer-level relationship is not, because the size of relationships between variables can differ at different levels of aggregation (Ostroff, 1993). Given this possibility, our meta-analytic results do not permit definitive conclusions about the incentive value of tipping, but they do call that value into question and point to the need for more research on this issue.

Tip Size as a Measure of Server Performance

The intangible and customized nature of services makes it difficult for managers to evaluate their service employees’ performances (Shamir, 1984; Zeithaml, Berry & Parasuraman, 1988). Customers are in a better position than managers to make these performance appraisals. In fact, economists argue that this is the reason tipping exists (Bodvarsson & Gibson, 1994; Hemenway, 1984; Jacob & Page, 1980). Given these considerations, managers may be tempted to use servers’ tip averages as a measure of the

servers' job performances. An internal document at one restaurant in Houston, Texas illustrates such a use of tip information. It reads:

“This program has been set up to assist you in better serving your guests. It will be monitored by your charge tip averages...Tip averages are the most effective way to measure a server's capabilities and progress within the restaurant.”

The weak relationship between tipping and service in this meta-analysis challenges the wisdom of such performance appraisal practices. Tips do not strongly reflect service quality at the customer level of analysis. While the service-tipping relationship may be stronger at the server level of analysis (see previous comments), it may also be weaker. In the absence of more server level data, our meta-analytic findings suggest that managers should be cautious in using tip averages as a measure of server performance. Instead, managers should personally observe their servers' work, hire “mystery diners” to provide feedback about server performance, and/or solicit explicit evaluations of server performance from customers.

CONCLUSION

In conclusion, the results of this meta-analysis support the idea that restaurant patrons reward better service with larger tips. This suggests that equity motivations operate in commercial as well as social exchanges. However, the relationship between tip size and service evaluations was weaker than most people would expect. Given the small size of this relationship, restaurant managers should not rely on tips as the sole incentive for their employees to deliver good service. Nor should managers use server tip averages as the only measure of server performance. Tips are a means by which consumers reward good service, but (in restaurant settings at least) they may not be the management tool that economists and some restaurant managers believe them to be.

REFERENCES

- Adams, S. (1965). Inequality in social exchange. In: L. Berkowitz (ed.), Advances in Experimental Social Psychology, Vol.2. New York: Academic Press.
- Baune, C. (1992). The economics of tipping at Waldo's Pizza. Unpublished student paper, St. Cloud State University, St. Cloud, MN.
- Bem, D.J. (1972). Self perception theory. In L. Berkowitz (Ed.), Advances in experimental social psychology. Vol. 6. New York: Academic Press.
- Bodvarsson, O.B. & Gibson, W.A. (1994). Gratuities and customer appraisal of service: Evidence from Minnesota restaurants. The Journal of Socio-Economics, 23, 287-302.
- Crusco, A.H. & Wetzel, C.G. (1984). The Midas touch: The effects of interpersonal touch on restaurant tipping. Personality and Social Psychology Bulletin, 10, 512-517.
- Fisher, R. (1992). The effect of feeling good and gratuity generosity. Unpublished student paper. Cornell University, Ithaca, NY.
- Gardner, M. P. (1985). Mood states and consumer behavior: A critical review. Journal of Consumer Research, 12, 281-300.
- Hemenway, D. (1980). Prices and choices: Microeconomic vignettes. Cambridge, MA: Ballinger.
- Jacob, N. & Page, A. (1980). Production, information costs and economic organization: The buyer monitoring case. American Economic Review, 70, 476-478.
- Kilkelly, K. (1992). An economic study on restaurant tipping. Unpublished student paper, St. Cloud State University, St. Cloud, MN.
- Lynn, M. (1988). The effects of alcohol consumption on restaurant tipping. Personality and Social Psychology Bulletin, 14, 87-91.
- Lynn, M. & Grassman, A. (1990). Restaurant tipping: An examination of three "rational" explanations. Journal of Economic Psychology, 11, 169-181.

- Lynn, M. & Graves, J. (1996). Tipping: An incentive/reward for service? Hospitality Research Journal, 20, 1-14.
- Lynn, M. & Latane, B. (1984). The psychology of restaurant tipping. Journal of Applied Social Psychology, 14, 549-561.
- Lynn, M. & Petrick, G. (1996). Untitled and unpublished data set, Cornell University, Ithaca, NY.
- Lynn, M. & Simons, T. (2000). Predictors of male and female servers' average tip earnings. Journal of Applied Social Psychology, in press.
- Lynn, M. & Strong, P. (1992). Untitled and unpublished data set, University of Houston, TX.
- May, J.M. (1978). Tip or Treat: A Study of Factors Affecting Tipping Behavior. Unpublished master's thesis, Loyola University of Chicago.
- Mullen, B. (1989). Advanced BASIC Meta-Analysis. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Nisbett, R.E. & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.
- Oliver, R.L. & Swan, J.E. (1989). Consumer perceptions of interpersonal equity and satisfaction in transactions: A field survey approach. Journal of Marketing, 53, 21-35.
- Ostroff, C. (1993). Comparing correlations based on individual-level and aggregated data. Journal of Applied Psychology, 78, 569-582.
- Rind, B. (1996). Effect of beliefs about weather conditions on tipping. Journal of Applied Social Psychology, 26, 137-147.
- Rosenthal, R. (1991). Meta-Analytic Procedures for Social Research. Newbury Park, CA: Sage Publications.
- Seligman, D. (1998). Why do you leave tips? Forbes, December 14, 138-140.
- Shamir, B. (1984). Between gratitude and gratuity: An analysis of tipping. Annals of Tourism Research, 11, 59-78.
- Snyder, M.L. (1976). The inverse relationship between restaurant party size and tip percentage: Diffusion or equity? Personality and Social Psychology Bulletin, 2, 308.

- Speer, T.L. (1997). The give and take of tipping. American Demographics, February, 51-54.
- Walster, E., Berscheid, E. & Walster, G.W. (1973). New directions in equity research. Journal of Personality and Social Psychology, 25, 151-176.
- Zeithaml, V.A., Berry, L.L. & Parasuraman, A. (1988). Communication and control processes in the delivery of service quality. Journal of Marketing, 52, 35-48.