Measuring and Predicting National Differences in Tipping Customs: A Critical Review

Michael Lynn*

Cornell Hotel School

International Journal of Hospitality Management (2025), 131, doi:104343

*Michael Lynn is the Michael D. Johnson and Family Professor of Services Marketing at the School of Hotel Administration, Cornell University. Direct correspondence to him using email at wml3@cornell.edu. Author declares no funding or conflict of interest.

Abstract

A review of research on the predictors of national tipping customs found that they varied with national characteristics likely to affect motivations for and against tipping and/or exposure to other pro-tipping cultures. However, numerous methodological problems undermine confidence in the theoretical processes underlying these effects. In addition, inconsistencies in findings raise questions about which effects are real and how generalizable they are. Measures of national customary tip sizes based on travel guides show weak convergent validity, so some of the inconsistencies may be attributable to weak measurement. Other potential sources of inconsistency are low statistical power and the effects of untested moderator variables. Better measures of national customary tip amounts based on worldwide surveys of national populations are needed as are larger sample sizes and more tests of moderation effects. Such measures and research would benefit researchers, hospitality/tourism businesses, and international tourists alike.

Measuring and Predicting National Differences in Tipping Customs: A Critical Review

1. Introduction

In many countries around the world, it is customary for consumers to give a voluntary sum of money (called a drick, mancia, pourboire, propina, and tip, among other names) to hospitality and other service providers. Whom it is customary to tip and how much they should be tipped varies from country to country to country. For example, in the U.S. it is customary to tip many different service providers - including taxicab drivers who are typically tipped 15 to 20 percent of the bill; while in Sweden it is customary to tip only a few service providers -- including taxicab drivers who are typically tipped 10 percent of the bill (Lynn, 1997; Lynn and Lynn, 2004). Understanding the drivers of these national differences in tipping customs is theoretically interesting because those drivers are likely to provide insights into the functions of and motivations for tipping as well as into the origins, evolution, and spread of social norms more generally. Such an understanding would also help businesses in the hospitality and tourism industries to better anticipate, understand, and address resistance to potential counter-normative tipping policies, such as those adopted by some airlines, hotels, restaurants, and resorts (c.f., Berger, 2019; Evans and Dave, 1999; Richards and Rosato, 1995; Wells, 2013). Finally, such knowledge would provide a deeper understanding and richer experience of the nations they visit to international tourists seeking to learn about different cultures.

The theoretical and practical value of such knowledge has motivated a number of quantitative studies examining the predictors of national differences in tipping customs. Those studies have appeared in behavioral economics (Lynn, 2006; Lynn and Starbuck, 2015), consumer behavior (Lynn, Zinkhan, and Harris, 1993), cross-cultural studies (Minkov, 2008; Mansfield, 2016), hospitality management (Lynn, 1997, 2000a, Lynn and Lynn, 2004), marketing

(Ferguson, Meghee and Woodside, 2018), personality and individual differences (Lynn, 1994, 2000b), social psychology (Torffason, Flynn and Kupor, 2012), and tourism (Ferguson, Meghee and Woodside, 2017; Schwartz and Cohen, 1999) journals. The current paper brings this scattered research together in a critical review of the literature. Publicly available quantitative studies examining predictors of national differences in tipping customs were identified from personal knowledge of the tipping literature as well as a Google Scholar search of "national differences in tipping," and an examination of the references of already identified papers. Excluded from this review are numerous unpublished analyses of the current author – including some that will be presented in a soon to be published book on tipping.

2. Casual Review of Theory and Findings

Research on national differences in tipping has largely followed the lead of Lynn, Zinkhan and Harris (1993), who theorized that they stem from national differences in the value placed on tipping's consequences. The consequences of tipping usually considered in this research include: (i) better wages for service workers, (ii) better future service for consumers, (iii) more social approval/esteem for consumers, (iv) maintenance of equity/reciprocity between consumers and service workers, and (v) fulfillment of a social obligation or duty to tip, (vi) creation or exacerbation of status and power differences between consumers and service workers, and (vii) loss of resources by tippers. Lynn (2015) has connected these consequences to altruistic, future-service, social-esteem, reciprocity, and duty motives for tipping and to egalitarian and cost-saving motives for not tipping respectively. The only other general explanation for national differences in tipping customs was offered by Mansfield (2016), who argued that they reflect national differences in travel related exposure to other pro-tipping cultures. Thus, researchers have examined the effects on tipping customs of national characteristics considered likely to

affect (i) altruistic, future service, reciprocity, social-esteem, and duty motivations for tipping, (ii) egalitarian and cost-saving motivations for not tipping, and/or (iii) exposure to other pro-tipping cultures. Table 1 provides a list of the studied predictors of national differences in tipping and an indication of which underlying process or processes each has been theorized (and/or seems to the current author as being likely) to affect, as well as an indication of the direction of that predictor's statistically significant relationships (if any) with tipping customs. More detailed summaries of the identified studies examining national differences in the prevalence of tipping are in Table A1 and those of studies examining national differences in customary tip sizes are in Table A2.[1]

Insert Table 1 about here

A casual reading of this literature seems to provide strong support for existing theory about the origins/causes of national differences in tipping customs. Specifically, the literature indicates that tips are more common and/or larger in countries to the extent that they have characteristics thought likely to strengthen population desires for tipping's consequences – i.e., for (i) improving service workers' wages, (ii) receiving better future service, (iii) gaining or keeping social approval/esteem, (iv) maintaining equity/reciprocity with service workers, and (v) fulfilling a social obligation or duty to tip (see Table 1). The literature also indicates that tips

---

[1] A note on the terminology used here: Tipping prevalence and tipping frequency are used interchangeably to refer to how often (across both occupations and situations) tipping is practiced by the population of a nation. Tip amount refers to how much money the people in a nation tend to tip specified service providers on a given tipping occasion. Tipping customs or norms typically specify who typically is (or should be) tipped and how much money they typically are (or should be) tipped, so these terms are used to refer to tipping prevalence and/or various customary tip amounts.

are less common and/or smaller in countries to the extent that they have characteristics thought to foster stronger population concerns about (vi) the power and status implications of tipping, and (vii) the monetary costs of tipping. Finally, the literature indicates that tips are larger in countries with characteristics that are thought to expose their populations to other cultures with large normative tips. However, a more critical reading of the literature is needed to have confidence in these findings and conclusions.

3. Critical Analysis of the Literature

Several considerations suggest that one or more of the significant predictors of tipping customs in the literature may be Type 1 errors. First, many of the significant effects have failed one or more attempts at replication (see Tables A1 and A2). Second, the use of small samples, frequent removal of outliers, and uncertainty about how many non-significant predictors have been examined but not reported or published raises concerns about Type-1 errors and their handmaidens p-hacking and selection-bias. Given that there have been serious allegations that much larger literatures, such those on ego-depletion (Friese, et. al., 2019) and power-posing (Simmons and Simonsohn, 2017), are just collections of Type 1 errors, it is plausible that the national tipping customs literature is also just noise-mining with no real evidential value about the predictors of tipping customs. Indeed, the plethora of easily obtained data on national differences exacerbates the potential file drawer problem (unpublished null effects) and makes the likelihood that the literature is just a collection of Type 1 errors even greater than is true of more experimental literatures.

3.1. Test that the Literature is Just a Collection of Type 1 Errors

The possibility that the vast majority of effects in the literature on national differences in tipping are Type 1 errors was assessed by analyzing the distribution of p-values from those significant (two-tailed $p < .05$) results summarized in Tables A1 and A2.[2] Note that multiple significant tests of the same relationship between specific variables were included in the main analysis, but that all such tests differed at least a little in terms of sample and/or control variables. This analysis is based on the idea that when the null-hypothesis of no effect is true, then significant results are found by chance, which should produce a linear trend when regressing significant p-values on their rank order (with no constant in the model). A positive quadratic trend in that regression provides evidence that a reasonably large number of the observed effects are real and not just Type 1 errors. The logic is similar to that underlying Simonsohn, Nelson and Simmons' (2014) p-curve analysis, but the test of significance used here should be less sensitive to dependencies among the p-values than is the Stoufer test used in p-curve analysis. Graphical results of the analysis of significant p-values in the national tipping customs literature are presented in Figure 1. A significant positive quadratic trend in the regression of p-values on their rank order ($\beta = 1.39$, $p < .001$, $n = 70$) indicates that there are significantly more very small p-values than moderately small ones, which would not happen if a vast majority of the effects were Type 1 errors.

As a robustness check, two additional analyses examined (i) p-values only from the first significant test of each predictor's relationships with tipping prevalence in TableA1, and (ii) p-values only from the first significant test of each predictor's relationships with tip size (or change

[2] The significant effects of attitude toward tipping were omitted from this analysis in order to keep this obviously valid relationship from carrying the results and hiding the absence of other real (non-chance) effects. Also dropped were a few effects whose calculated exact p-values were not significant at the two-tailed .05 level.

in tip size) in Table A2. These analyses are more conservative than the one reported above because they involve fewer significant p-values being analyzed and weaker dependencies among them. Again, significant effects of attitude toward tipping were omitted from these analyses. The results of these more conservative analyses were similar to the one reported above - significant positive quadratic trends were found in the relationships between p-values and their rank orders in both analyses ($\beta = 1.79$, $p < .001$, $n = 22$ and $\beta = 1.83$, $p < .001$, $n = 20$ respectively). Thus, it is clear that the literature does identify real predictors of both tipping prevalence and customary tip size.

3.3. Identification of More and Less Reliable Predictors

While the literature as whole does contain information about valid predictors of tipping customs, that does not mean all the findings are valid. In fact, inconsistent findings about the effects of numerous predictors across analyses and/or studies (see Tables A1 and A2) suggest that at least some of the findings cannot be accepted at face value. Assuming that effects replicated at least once are more reliable than unreplicated ones, then it is reasonable to infer from the literature that the national prevalence of tipping and/or customary size of tips increase with national power distance, uncertainty avoidance, masculinity, extraversion, religiosity, commitment to duty, income inequality, and bribery while they decline with national psychoticism, trust, purchasing power parity, tax burden and political/civil rights (see Table 1). If you accept significant effects that no one has yet attempted to replicate, then the existing literature also indicates that the prevalence of tipping increases with the value placed on recognition and esteem, monumentalism, the need for achievement and the need for power and that the customary size of restaurant tips increases with national travel to the U.S. and the size of

tips in neighboring countries while it decreases with the customary application of service charges. These seemingly trustworthy effects provide some support for all of the theorized processes underlying national tipping customs (see shaded cells of Table 1), so the earlier casual reading of the literature is not as misleading as it could have been.

Nevertheless, this support for theory is far from definitive because it consists only of unmoderated, correlational relationships with few control variables -- meaning that the effects could have any number of other explanations. In fact, many of the significant predictors are associated with more than one of the theorized underlying processes (see Table 1), so it is unclear precisely what process or processes are responsible for the effects. The use of correlational methods cannot be avoided in studying this topic, but future researchers need to use more control and moderator variables (and hence, larger samples too) in order to better test the theorized explanations for national differences in tipping customs.

3.3. Need for Tests of Moderation

The numerous failures to replicate effects in the national tipping customs literature raise questions about the generalizability as well as the reliability of those effects. Unfortunately, the current literature contains almost no test of moderation effects. Starbuck (2009) tested the interaction of income inequality (GINI) with egalitarianism (reversed: PDI) finding non-significant effects on different measures of tipping customs, but that is the only interaction effect tested in the literature. Furguson, et. al.'s (2017, 2018) qualitative comparative analysis, with its emphasis on combinations (or recipes) of determinants for membership in a set of high tipping nations and separate combinations for membership in a set of low tipping nations seems at first glance to resemble tests of interactions. Indeed, they contrast their methodology with more traditional methods arguing that "the main effect relationships in these prior studies do not reflect

the complex interactions between cultural elements and their effect on tipping behavior" that their analyses provide (Feguson, et. al., 2018). However, their qualitative comparative analyses involve no statistical tests of significance and, hence, no way of separating chance from non-chance configurations of cultural drivers of tipping customs. Furthermore, they examine Boolean combinations (or "intersections") of elements that together impact tipping customs, but this is not quite the same thing as moderation effects in which one variable strengthens or weakens the effects of another. Thus, there is a clear need for more tests of moderation effects in the literature.

One obvious candidate moderator deserving of attention is service type. Lynn (2018, 2021) found that the motives underlying individual differences in tipping vary by occupation or service type, so it is plausible that the causes and predictors of national differences in tip size also vary by occupation or service type. Supporting this possibility, researchers have found inconsistent effects when examining the predictors of customary tip sizes given to waiters/waitresses and taxicab drivers (see Table A2). For example, individualism has been found to be positively related to national taxicab, but not restaurant, tip sizes (Lynn and Lynn, 2004; Starbuck, 2009). Unfortunately, it is not clear if this and other service-type differences in the effects of national characteristics on customary tip sizes are due to chance or reflect the reliable moderation effects of service type on the drivers of customary tip sizes. This issue deserves more research attention going forward.

Of course, theory about underlying processes should be used to identify other potential moderators too. There are too many potential interactions of theoretical interest to enumerate, but there is a clear class or category of interaction that I believe should be given priority – namely situation X disposition interactions. Some national characteristics may affect the motives underlying tipping because they describe the circumstances faced by the people in a country. For

example, income inequality (GINI) is mostly a situational variable (though it may also say something about the egalitarian values of a country too), as are national differences in GDP per-capita, purchasing power parity (PPP), tax burden, number of tourist to and from the U.S., and having customary service charges added to bills. Other national characteristics may affect the motives underlying tipping because they describe more general values and dispositions of the people in a country. For example, national differences in power distance, uncertainty avoidance, masculinity, individualism, extraversion, neuroticism, psychoticism, and needs for achievement all reflect underlying values, motives, and/or behavioral tendencies. At the individual level, dispositional effects on behavior are often activated or strengthened by relevant situations – see research on Trait Activation Theory (Tett, Toich and Ozkum, 2021) – and similar interaction effects are likely to occur at the national level too. Starbuck's (2009) test of GINI X PDI interactions is an example of what I am advocating. His tests proved unproductive, but other interactions of this type may bear more fruit. For example, national benevolence, empathy or related dispositions may predict tipping customs more strongly in countries with large income inequality, people living in poverty, or other situational triggers of interpersonal caring. Identifying and testing situation X disposition interactions like these may help explain some of the inconsistencies in the existing literature and would provide stronger tests of theory about the underlying causes of national differences in tipping.

3.4. Measurement Reliability and Validity

Inconsistencies in the findings of the national tipping customs literature could also be due to poor measurement quality in at least some studies. Researchers have typically used international tipping guides as sources of data about tipping customs. Many have reported weak and inconsistent agreement between guides about the recommended tips sizes for restaurant servers

and for taxicab drivers. With respect to recommended tip amounts in restaurants, Lynn and Lynn (2004) reported correlations from three different guides that ranged from .34 to .61 with a mean of .51; Starbuck (2009) reported correlations from five different guides that ranged from .22 to .77 with a mean of .47; Lynn and Starbuck (2015) reported correlations from four different guides that ranged from .48 to .67 with a mean of .59; and Mansfield (2016) reported correlations from five different guides that ranged from .50 to .80 with a mean of .65. The consistency of recommendations for taxicab tip sizes is a little better, but still not great – Lynn and Lynn (2004) reported correlations from three different guides that ranged from .45 to .70 with a mean of .57, while Starbuck (2009) reported correlations from five different guides that ranged from .51 to .93 with a mean of .72. Thus, it appears that the reliability of data on customary tip amounts from different tipping guides varies but is generally low. The researchers cited above all averaged the values from different guides into indices, perhaps in the hope that it would cancel out much of the error variance in each guide's recommendations, but it is not clear if the numbers of guides used were large enough to make this strategy effective.

To further assess the quality of the measures used for analyses in this literature, the correlations among those measures are examined below and recommendations about what measures to trust and use in future research are provided. National values for all the measures of tipping customs used for analysis in the tipping customs literature (i.e., those in Tables 2-5) are available in a data repository at https://doi.org/10.7910/DVN/EUQ2A4, but note that the TP2 and TP3 labels were switched in this paper from their use in that data file so that TP2 here is TP3 in the file and vice versa.

3.4.1. National Prevalence of Tipping

There are five measures of the prevalence of tipping in the existing literature, with one of those being an index derived from three of the others. These measures generally add or average the frequency of tipping across multiple occupations. The most commonly used measure (TP1) was developed by Lynn, Zinkhan, and Harris (1993) from a 1988 international tipping guide by Nancy Star and covers 32 nations. Two measures were developed as part of a dissertation by Starbuck (2009). One of those measures (TP2) was a binomial measure based on five different international tipping guides from 2003 that covers 160 nations. The other (TP3) was based on surveys of 35 to 50 consumers from each of 30 nations in 2002. The final measure (TP4) was developed by Lynn based on surveys of 63 Intercontinental Hotel Group concierges from 41 nations in 1998. The later measure was never used by itself, but it was averaged with TP1 and TP3 and the resulting index (TP5) was used by Lynn and Starbuck (2015). The intercorrelations among the measures are presented in Table 2.

Insert Table 2 about here

Given the relatively weak correlation of TP4 with the other measures of the same construct and the fact that it was based on the judgment of only one to seven concierges per country, it appears to be the least valid of the lot. This means that Lynn and Starbuck's (2015) use of the measure as a component of their index TP5 is likely to have undermined the validity of that index to some degree. Thus, tests of relationships involving TP4 and TP5 should be viewed with more skepticism than those involving TP1, TP2 and TP3. An average of the standardized values of TP1 and TP3 would be a better measure of national tipping prevalence than is the index TP5.

Such a new index (TP6) covers 18 fewer nations, but it correlates more highly with TP2 ($r = .78$, $n = 40$, $p < .001$) than does TP5 ($r = .60$, $n = 57$, $p < .001$). Going forward, researchers looking to use existing measures should choose between TP2 (which is only binomial, but covers many more nations) and TP6 (which covers fewer nations, but is much less crude). Note that the high correlation of TP1 (from a 1988 guide) with TP2 (from 2003 guides) provides evidence that this dimension of tipping customs does not change much over time and its high correlation with TP3 (from 2002 surveys) not only reinforces that point but also supports the validity of tipping guides as sources of information about this dimension of tipping customs.

3.4.2. Customary Restaurant Tip Amounts

There are seven different measures of customary national restaurant tip percentages in the existing literature. All but one of them are based on data from international tipping guides, with the specific guides being used varying from one researcher to another. The lone exception was a measure developed by Starbuck (2009) from surveys of 35 to 50 consumers from each of 30 nations in 2002. The correlations among all of these measures are reported in Table 3.

Insert Table 3 about here

The correlations among the different measures of customary restaurant tip amounts were positive, but mostly modest in size -- ranging from .24 to .88 with a mean of .57. It is tempting to attribute the weak convergent validity of these measures to differences in the way the data were coded. Some measures included tips of zero while others did not; some treated small denomination-based tip recommendations as 3% while others dropped those denomination-based

tip recommendations; some excluded customary tip amounts given on top of service charges while others did not. Consistent with this critique, the highest correlations were between measures with similar coding rules -- RTA2, RTA4, and RTA7, which all included tips of zero, small denomination-based tips, and tips on top of service charges. However, that is not the whole story because, as already discussed, similar weaknesses in convergent validity have been reported by authors who used a consistent coding procedure to create indices from individual tipping guides.

Another possible explanation for weak convergent validity among these measures is that different tipping guides were produced at different times and (unlike the prevalence of tipping) customary tip amounts may change over time. In fact, Lynn (2025) provides evidence that restaurant tipping percentages in the U.S. have increased by about 1.5 points each decade over the past half century. Similar temporal effects might explain some of the inconsistencies across different tipping guides. However, even guides published within one year of one another have shown weak convergent validity (see Lynn and Lynn, 2004), so the problem is most likely to be attributable to poor data collection by the guides' authors. Overall, the evidence suggests that the reliability of data on customary restaurant tip amounts from different tipping guides is generally low and that averaging the data across guides does not appreciably increase its reliability.

Arguably, Starbuck's (2009) survey-based measure (RTA4) has the greatest prima facie validity. If you accept it as the standard, then only RTA2 provides a larger samples of nations together with good convergent validity. Thus, tests of relationships involving RTA1, RTA3, RTA5, RTA6, and RTA7 should be viewed with more skepticism than those involving RTA2 and RTA4. Furthermore, researchers interested in using existing measures of customary restaurant tip amounts should use one of the latter measures -- choosing between them based on their

preferences regarding the tradeoff between validity and number of nations covered. However, there is clear need for better measures of this construct based on more extensive and standardized surveys of multiple nations' populations about their customary restaurant tip amounts.

3.4.3. Customary Taxicab Tip Percentages

There are five different measures of customary national taxicab tip percentages in the existing literature. All but one of them are based on data from international tipping guides, with the specific guides being used varying from one researcher to another. The lone exception was a measure developed by Starbuck (2009) from surveys of 35 to 50 consumers from each of 30 nations in 2002. The correlations among all of these measures are reported in Table 4.

Insert Tables 4 and 5 about here

The correlations among the different measures of customary taxicab tip amounts were positive, but mostly modest in size -- ranging from .33 to 1.00 with a mean of .56. [3] TTA5's perfect correlation with TTA1 stands out, but is attributable to the former taking 20 of its 30 values directly from the latter. TTA2, which was based on only one guide, shows the weakest convergent validity with all its correlations being less than .45. After dropping TTA2, TTA3 has the strongest convergent validity with all its correlations being greater than .70. Thus, findings involving TTA3 should be seen as more credible than those involving the other measures of

[3] Similar weakness in convergent validity has been reported by authors who created the indices TTA1 and TTA3 from individual tipping guides – Lynn and Lynn (2004) reported correlations from three different guides that ranged from .45 to .70 with a mean of .57, while Starbuck (2009) reported correlations from five different guides that ranged from .51 to .93 with a mean of .72.

customary taxicab tip amounts and it should be used as the preferred existing measure in future research. However, there is a clear need for better measures of this construct too, so future researchers are encouraged to create those better measures by conducting systematic surveys of many national populations about their customary taxicab tip amounts.

3.4.4. Correlations Among Different Tipping Customs

Correlations among the preferred measures of national tipping prevalence, restaurant tip amounts, and taxicab tip amounts are presented in Table 5. They are all positively correlated, which suggests that one or more common cause underlies all of these dimensions of tipping customs. Nevertheless, the two measures of tipping prevalence correlated substantially more strongly with one another than with the measures of customary tip amounts, which provides some evidence for divergent validity and justifies expectations that tipping prevalence and customary tip amounts may have different predictors and underlying causes.

4. Summary and Conclusions

Research on the predictors of national tipping customs found that they varied with national characteristics such as power distance, uncertainty avoidance, masculinity, extraversion, bribery, purchasing power parity, income inequality, tax burden, impression management concerns, religiosity, travel to the U.S., use of service charges, and tipping levels in neighboring countries. These findings are consistent with the ideas that national tipping customs are caused by (i) national differences in the value placed on tipping's consequences (Lynn, Zinkhan and Harris, 1993) and (ii) differential exposure to pro-tipping cultures (Mansfield, 2016). More specifically, they can be seen as suggesting that tipping functions as a way of helping and/or fairly compensating workers, of gaining social approval, esteem, goodwill and/or service (or

avoid their loss), and of fulfilling a social obligation. From a managerial perspective, these findings and ideas suggest that adoption of no-tipping policies in places where tipping is common should be accompanied by increasing other assurances or guarantees of these outcomes. However, correlational designs, multi-faceted predictors, small sample sizes (permitting few control variables), and a focus on main effects mean that other explanations for the observed relationships are possible, so much more research is needed to nail down the causal processes underlying these predictors' relationships national differences in tipping customs.

In addition, inconsistencies in findings raise questions about which effects in the literature are real and how generalizable they are. To better establish the reliability and generalizability of findings, future researchers need to collect data on larger samples of nations and to test interactions as well as main effects. Particularly, promising interactions to study are those between dispositional differences among nations (e.g., empathy, egalitarianism, and status-seeking) and situational differences among nations likely to activate those national dispositions (e,g., GINI, PPP, and tax burden).

Another potential source of inconsistent findings in this literature is poor measurement. Measures of national customary tip sizes based on travel-guides show weak convergent validity. Since tipping guides do not appear to be a very reliable source data about customary restaurant and taxicab tip amounts, future researchers should use systematic surveys of consumers around the world to get better measures of these constructs. International, multi-scholar efforts -- along the lines of those underlying the World Values Survey (Inglehart, Basanez and Moreno, 1998) and the GLOBE Project (House, Hanges, Javidan, Dorfman and Gupta, 2004) -- to produce such survey-based measures of customary tip amounts would allow scholars to better identify the predictors of national tipping customs and, in so doing, to test theories about causal processes

underlying those national differences. Such improved measures could also be used to improve tipping guides, which are clearly inadequate, and that would enable foreign tourists seeking to immerse themselves in different cultures to better conform with local norms in the countries they visit.

In addition to better measures, larger sample sizes, use of more control variables, and more frequent testing of interactions, the literature on national tipping customs would also benefit from a broader base of theory. The main explanations offered to date are (i) that tipping customs reflect the value that national populations place on tipping's consequences (Lynn, Zinkhan and Harris, 1993), which draws on functional and rational choice ideas, and (ii) tipping customs spread via travel related exposure to pro-tipping cultures (Mansfield, 2016), which draws on social-imitation and social-diffusion ideas. Even Starbuck (2009) who discusses game theory, and Ferguson, et. al. (2017, 2018) who reference complexity theory and fuzzy set theory, ultimately rest their explanations for observed national differences on rational choice ideas akin to those of Lynn, et. al. (1993). I am not familiar enough with other disciplines to identify specific candidates, but surely there are additional relevant theories in anthropology, political science, regional studies, sociology and/or other fields. Future hospitality and tourism researchers studying this topic should consider drawing upon (and perhaps collaborating with scholars in) these other disciplines to develop more and richer theories about national differences in tipping customs.

Cross-cultural research on tipping customs is still in its infancy, and substantial improvements in theory and measurement as well as in study sample sizes and study design are needed going forward, but some meaningful effects with practical as well as theoretical implications have been identified. Hopefully, this paper will encourage more hospitality and

tourism scholars to add to this literature and, thereby, expand our understanding of national differences in this fascinating custom.

Table 1. Theorized and/or likely effects of studied predictors on motives for and against tipping along with significant relationships of those predictors with tipping customs in the literature.

	Altruistic Motives	Future Service Motives	Social Approval/ Esteem Motives	Equity/ Reciprocity Motives	Duty Motives	Egalitarian (Anti-Tipping) Motives	Cost Saving (Anti-Tipping) Motives	Norm Diffusion via Exposure to Pro-Tipping Cultures	Significant Relationships with Tipping Customs[a]
Power Distance						-			++
Uncertainty Avoidance		+	+						++
Masculinity	-		+	+					++
Individualism	-		-		-				mixed
Extraversion		+	+						++
Neuroticism		+	+						+
Psychoticism	-		-		-				--
Bribery		+							++
GINI	+					-			++
Human Development Index	-						-		mixed
PPP/GDP per capita	-						-		--
Political/Civil Rights[b]									--
Trust		-							--
Tax Rate/Burden							+		--
Govt Spending on Welfare/Healthcare	-								none
Value Recognition			+						+
Importance of Esteem			+						+
nAchievement			+	+					+
nAffiliation	+		+	+					none
nPower						-			+
Monumentalism			+						+
Homicide Rate[b]									none
Duty to Volunteer/Help					+				++
Lie Score			+						mixed
Religiosity	+								++
Travel/Entries to US								+	+
Tourism from US								+	none
Communist (y/n)						+			none
Service Chage common				-					-
Tip in neighboring countries								+	+
Workers' rights	-			-					none
Minimum wage	-			-					none
McDonalds per 10,000 people								+	none
Internet users per 100 people								+	none
Trade with US/All international trade								+	none

[a] In this column, one plus or minus sign means that at least one significant positive or negative effect of that predictor has been found. More than one plus or minus sign means the relationship of that predictor with tipping customs was significant using two different measures of tipping customs. Mixed results mean that there were significant effects of that predictor in both a positive and negative direction. None means that none of the observed relationships involving that predictor were statistically significant. [b] Used as a control variable with no theoretical rationale provided by original authors or apparent to the current author. .Note: Shaded cells identify theorized/expected effects of predictors on tipping motives/processes that are supported by evidence of the predictors' relationships with tipping customs.

Table 2. Correlations among measures of national tipping prevalence (with n's on and below the diagonal) show reasonable convergent validity – except for TP4.

	TP1	TP2	TP3	TP4	TP5
TP1 (Lynn, et. al., 1993: from a 1988 guide)	32	.82**	.87**	.76**	.97**
TP2 (Starbuck, 2009; from five 2003 guides)	32	160	.72**	.40*	.60**
TP3 (Starbuck, 2009; from 2002 surveys)	22	30	30	.56**	.93**
TP4 (Lynn, 1998; from 1998 surveys)[a]	20	40	18	41	.93**
TP5 (Lynn & Starbuck, 2015: index of TP1, TP3, and TP4)	32	57	30	41	58

[a] Never used as a standalone measure in analyses. *p < .05, ** p < .01

Table 3. Correlations among measures of customary national restaurant tip amounts (with n's on and below the diagonal) show weak convergent validity.

Measure (Author/Year/Source of Data)	RTA1	RTA2	RTA3	RTA4	RTA5	RTA6	RTA7	Average r
RTA1 (Lynn & Lynn, 2004: 4 tipping guides)	62	.29	.39*	.60**	.62**	.24	.88**	.50
RTA2 (Lynn, 2008: 1 tipping guide)	36	75	.41**	.81**	.50**	.57**	.67**	.54
RTA3 (Starbuck, 2009: 5 tipping guides)	59	61	135	.70**	.57**	.44**	.72**	.54
RTA4 (Starbuck, 2009: consumer survey)	18	29	29	30	.69**	.79**	.70**	.72
RTA5 (Lynn & Starbuck, 2015: 4 tipping guides)	58	59	109	23	160	.29**	.44*	.52
RTA6 (Mansfield, 2016: 5 tipping guides)	60	68	130	30	131	173	.73**	.51
RTA7 (Ferguson, et. al., 2018; 3 tipping guides)	17	26	26	19	24	27	30	.69

*p < .05, ** p < .01

Table 4. Correlations among measures of customary national taxicab tip amounts (with n's on and below the diagonal) show weak convergent validity.

Measure (Author/Year/Source of Data)	TTA1	TTA2	TTA3	TTA4	TTA5	Average r
TTA1 (Lynn & Lynn, 2004: 4 tipping guides)	70	.38*	.76**	.47*	1.0**	.65
TTA2 (Lynn, 2008: 1 tipping guide)	46	76	.41*	.46*	.43*	.42
TTA3 (Starbuck, 2009: 5 tipping guides)	65	61	125	.70**	.73	.65
TTA4 (Starbuck, 2009: consumer survey)	28	29	27	30	.49*	.53
RTA5 (Ferguson, et. al., 2018; 2-3 tipping guides)	24	27	26	19	30	.66

*p < .05, ** p < .01

Table 5. Correlations among measures of different national tipping customs (with n's on and below the diagonal) provide evidence of divergent validity for measures of national tipping prevalence and national customary tip amounts.

	TP3	TP6		RTA4	RTA2	TTA3
Tipping Prevalence TP2 (Starbuck, 2009)	160	.78**		.49**	.41**	.33**
Tipping Prevalence TP6[a] (new index)	40	40		.43*	.41*	.48**
Restaurant Tip Amount RTA4 (Starbuck, 2009)	29	28		30	.81**	.73**
Restaurant Tip Amount RTA2 (Lynn, 2008)	64	35		29	75	.53**
Taxicab Tip Amount TTA3 (Starbuck, 2009)	120	35		27	60	125

[a] TP6 is a new index formed from averaging standardized scores of TP1 and TP3. *p < .05, ** p < .01

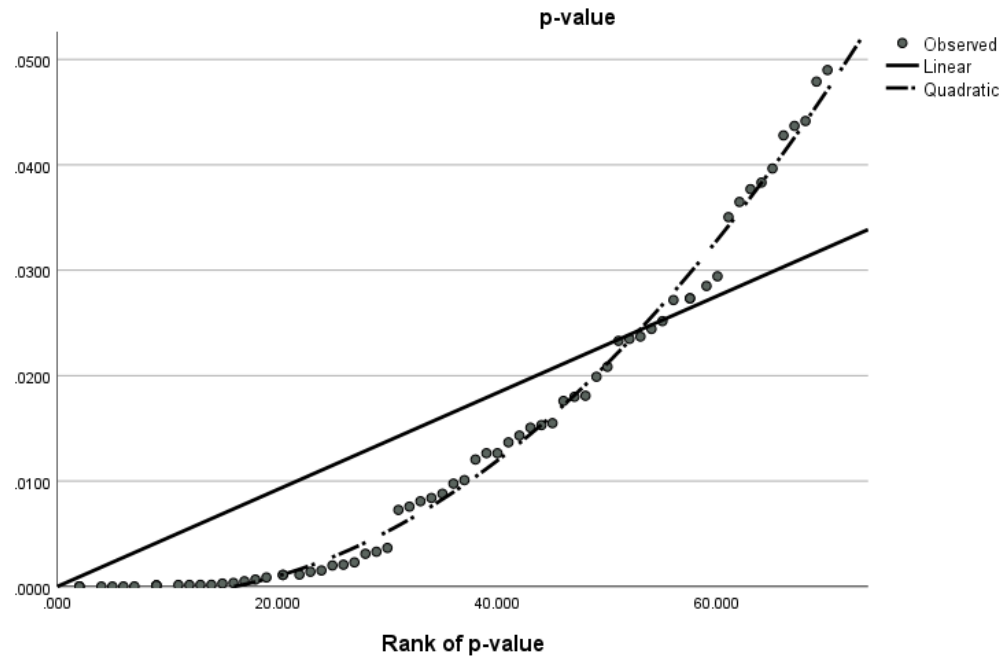

Figure 1. Significant p-values from the national tipping customs literature plotted against their rank order indicate that the relationship is non-linear – meaning that a sizeable number of those relationships are real and not attributable to chance alone.

Appendix

Table A1. Summary of studies examining the predictors of national differences in the prevalence of tipping.

Study/Tipping Measure	**Predictor**	**r**	**pr**	**ß**	**B (SE)**	**n**	**Comments**
Lynn, et. al. (1993)/Tipping Prevalence 1 (TP1)	Power Distance (PDI)	.46*	.21			29	TP1 is the number of tipped professions out of 33 in each country as coded from Star (1988). Values from Hofstede(1983). Japan removed as an outlier in analyses. Partial r's control only for other values.
	Uncertainty Avoidance (UAI)	.55*	.41*			29	
	Masculinity (MAS)	.47*	.43*			29	
	Individualism (IND)	-.39*	.00			29	
Lynn (1994)/TP1	Neuroticism	.86*				13	Neuroticism from Lynn and Hampson (1975). UK, Ireland, Canada, US and Japan removed as outliers in analyses.
Lynn (1997)/TP1	Value Recognition	.62*				15	Value of recognition from Gordon (1976). Importance of esteem from Haire, et. al. (1966). No tests for outliers.
	Importance of Esteem	.61*				13	
Schwartz & Cohen (1999)/TP1	Tax Burden			-.96*		24	Regression coefficients controlled for continent and are based on a reanalysis of the data provided in the paper's appendix, which was required to calculate exact p-values. No tests of outliers.
	Gov't Welfare Spending			-.45		18	
Lynn (2000)/TP1	n Achievement	.60*		.64*		25	Needs from McClelland's (1961) coding of children's readers published between 1946 and 1955.Brazil and New Zealand removed as outliers in analyses. Betas are from a regression with only n(ACH), N (AFFIL),and n(POW) as predictors.
	n Affiliation	-.20		.09		25	
	n Power	.31		.34*		25	
Lynn (2000)/TP1	Extraversion (E)	.32		.50*		19	Personality scores from Lynn and Martin (1995). Iceland and Japan removed as outliers in analyses. Betas are from a regression with only E, N and P as predictors.
	Neuroticism (N)	.56*		.77*		19	
	Psychoticism (P)	-.06		-.40*		19	
Lynn (2006)/TP1	Purchasing Power Parity (PPP)	-.49*				32	
	Attitude toward Tax Evasion	-.05				16	
Minkov (2008)/TP1	Monumentalism	.53*				29	Correlation is a Spearman rank order correlation. No tests for outliers.

Starbuck (2009)/ TP2	HDI	.17*		.79		154	TP2 is a binomial measure of tipping prevalence (high vs. low) based on coding of five online travel guides as reported by Starbuck (2009). N's are for correlations. Various outliers removed from different analyses. B(SE) are from a regression of TP2 on all of Hofstede's values (n = 47)
	PRI	-.21*		.78		158	
	Trust	-.48*		.94		38	
	GINI	-.01		1.15		123	
	Bribery	-.07				65	
	PDI	.10			-.02(.02)	46	
	UAI	.25*			.03(.02)	48	
	MAS	.33*			.06(.03)*	47	
	IND	-.17			-.02(.02)	48	
	Attitude toward Tipping	.67*				30	
Starbuck (2009)/ TP3	Human Development Index (HDI)	-.37*				30	TP3 is the median across 30 occupations of the average rated frequency of tipping each occupation in each nation as reported by Starbuck (2009). N's are for correlations. Various outliers removed from analyses. Betas are from a regression of TP3 on Trust and attitude toward tipping with n = 22. Partial r's regressed TP3 on all of Hofstede's values (n = 24).
	Political Rights Index (PRI)	.30				29	
	Trust	-.66*		-.51*		21	
	Income Inequality (GINI)	.19				28	
	Bribery	.27				21	
	PDI	.21	.13			25	
	UAI	.48*	.39			24	
	MAS	.46*	.49*			24	
	IND	-.27	-.03			24	
	Attitude toward Tipping	.66*		.58*		30	
Torfason, Flynn & Kupor (2012)/TP1	Bribery	.60*				32	No tests for outliers.
	IND	-.39*				32	
	PDI	.45*				32	
	GDP per capita (logged)	-.40*				32	
	GINI	.57*				32	
	Homicide Rate	.26				32	
	Civil Liberty Restrictions	.30				32	
	Highest Marginal Tax Rate	-.56*				32	
	Minimum Wage	-.19				32	
	Public Funding of Welfare	-.19				32	
Lynn & Starbuck (2015)/TP5	Extraversion	.21			-.05 (.04)	28	TP5 is an index constructed by Lynn and Starbuck (2015) by averaging TP1, TP2, and results from a survey of Intercontinental Hotel concierges from around the world (TP4). Regression coefficients are from a full information maximum likelihood analysis.
	Neuroticism	.33			.23* (.03)	28	
	Psychoticism	-.09			-.08 (.05)	28	
	Lie	.47*			.04 (.02)	28	
	Attitude toward Tipping	.72*			1.68* (.17)	30	
	Duty	.67*			.02* (.01)	24	

							N's are for correlations. N for FIML analysis was 56 but with missing values. No tests for outliers.
Ferguson, Meghee & Woodside (2017)/TP1	PDI	.40*		.10		30	Authors focus on a "configural analysis" that provides no tests of statistical significance. I focus on their standard correlation analyses. No deletion of outliers.
	UAI	.42*		.10		30	
	IND	-.33		.22		30	
	MAS	.26		-.05		30	
	Religiosity	.76*		.88*		30	
	GINI	.36*		.05		30	
	PPP	-.48*		.16		30	
* Effect reported as statistically significant at two-tailed $p < .05$.							

Table A2. Summary of studies examining the predictors of national differences in customary tip amounts.

Study/Tipping Measure	Predictor	r	ß	B (SE)	n	Comments
Lynn & Lynn (2004)						
Restaurant Tip Amount 1 (RTA1)	Power Distance (PDI)		-.17		35	Measures of RTA1 and TTA1 were coded by Lynn and Lynn (2004) from *Global Road Warrior*, *World Travel Guide* and *How to Tip*. Measures of tipping reflect average size of non-zero tips when no service charges are added to the bill. Small currency denominated tips coded as 3%. Predictors come from Hofstede (1983). Betas are from a regression with only PDI, UAI, MAS and IND as predictors.
	Uncertainty Avoidance (UAI)		.40*		35	
	Masculinity (MAS)		.66*		35	
	Individualism (IND)		.11		35	
Taxicab Tip Amount 1 (TTA1)	PDI		.28		41	
	UAI		-.03		41	
	MAS		.51*		41	
	IND		.30*		41	
Lynn (2008)						
Restaurant Tip Amount 2 (RTA2)	Extraversion (E)		.34ŧ		27	RTA2, TTA2, and PTA1 came from an online guide at Magellan.com and include tip amounts of zero. Small currency denominated tips coded as 3% in restaurants and 5% in taxicabs as reported by Lynn (2008). Tips given only under special circumstances were coded as 0. Countries where tipping was illegal were excluded. National EPQ scores came from Steel and One (2002). Betas are from a regression with only E, N and P as predictors. Was able to replicate Lynn's (2008) results involving RTA2, and reached same conclusions with different coefficients using TTA2 and PTA1.
	Neuroticism (N)		-.10		27	
	Psychoticism (P)		-.32ŧ		27	
Taxicab Tip Amount 2 (TTA2)	E		.39*		27	
	N		-.08		27	
	P		-.29		27	
Porter Tip Amount 1 (PTA1)	E		.45*		26	
	N		.08		26	
	P		.07		26	
Starbuck (2009)						
Restaurant Tip Amount 3 (RTA3)	Human Development Index (HDI)	-.33*			127	Measures of RTA3 and TTA3 were coded by Starbuck (2009) from *Columbus World Travel Guide, Fodor's, Frommer's, Global Road Warrior, and the Lonely Planet.* Measures of RTA4 and TT4 were based on surveys of people from each country and reported by Starbuck (2009). Measures of tipping are averages of reported tip percentages, including tips of zero. Small currency denominated tips coded as 3%. N's are for correlations. Various outliers removed from analyses.
	Political Rights Index (PRI)	-.18*			131	
	Trust	-.45*			34	
	Income Inequality (GINI)	.39*			106	
	Bribery	.54*			58	
	PDI	.18			44	
	UAI	.14			44	
	MAS	.08			44	
	IND	-.10			44	
	Attitude toward Tipping	.54*			29	
Taxicab Tip Amount 3 (TTA3)	HDI	.02			123	
	PRI	-.03			128	
	Trust	.04			34	

	GINI	-.16			102	
	Bribery	.00			62	
	PDI	-.25			46	
	UAI	-.27			46	
	MAS	.13			46	
	IND	.42*			46	
	Attitude toward Tipping	.58*			29	
Restaurant Tip Amount 4 (RTA4)	Human Development Index (HDI)	.19			27	
	Political Rights Index (PRI)	-.27			27	
	Trust	-.00			18	
	Income Inequality (GINI)	.10			27	
	Bribery	-.07			19	
	PDI	-.15			23	
	UAI	-.23			23	
	MAS	.06			23	
	IND	.37			23	
	Attitude toward Tipping	.71*			27	
Taxicab Tip Amount 4 (TTA4)	HDI	.29			28	
	PRI	-.19			28	
	Trust	.24			19	
	GINI	-.25			27	
	Bribery	-.32			19	
	PDI	-.40			22	
	UAI	-.27			22	
	MAS	.04			22	
	IND	.49*			22	
	Attitude toward Tipping	.57*			27	
Lynn & Starbuck (2015)						Measure of RTA5 is an index constructed by Lynn and Starbuck by averaging restaurant tip amounts, excluding tips of zero, reported in *World Travel Guide, Fodor's, Magellan, and Presents & Law.* Small currency denominated tips coded as 3%. Regression coefficients are from a full information maximum likelihood analysis that included all the predictors. N's are for correlations. N for FIML analysis was 56 but with missing values. No tests for outliers.
Restaurant Tip Amount 5 (RTA5)	Extraversion	.39*		.14 (.26)	32	
	Neuroticism	.15		.35 (.21)	32	
	Psychoticism	-.04		-.80* (.33)	32	
	Lie	-.24		-.63* (.15)	32	
	Attitude toward Tipping	.50*		2.31 (1.18)	23	
	Duty	.43*		.17* (.04)	27	

Mansfield (2016)						
Restaurant Tip Amount 6 (RTA6)	Travel/Entries to US			.58* (.23)	163	Measure of RTA6 is an index constructed by Mansfield (2016) by averaging restaurant tip amounts, including tips of zero, reported in *World Travel Guide, Global Road Warrior, Magellan, Chikyu no Arukikata, and Lonely Planet.* Small currency denominated tips were dropped/ignored. Results are from different models for different sets of variables; all models included US Entries, GDP, Communist, Service Charge, and Neighbor Tip. No tests for outliers. ΔRTA measured from Pan Am's Global Guide in 1982 and Global Road Warrior in 2010.
	Tourism from US			.16 (.23)	138	
	GDP per capita			-.00 (.00)	163	
	Communist (y/n)			-3.48 (2.43)	163	
	Service Chage common			-8.84* (.94)	163	
	Tip in neighboring countries			.70* (.15)	163	
	Workers' rights			-.24 (.78)	163	
	Minimum wage/GNI per worker			.23 (1.58)	158	
	McDonalds per 10,000 people			-0.17 (8.55)	154	
	Internet users per 100 people			.001 (.04)	154	
	Trade with US/All international trade			.66 (.66)	154	
Change in Restaurant Tip Amount (ΔRTA)	ΔTravel/Entries to US			1.49* (.67)	71	
	Service Charge in 2010			-7.48* (1.87)	71	
	Stable Service Charge (y,n)			-2.37 (1.61)	71	
	Service Charge 1982			9.62* (1.83)	71	
	ΔCommunist			.94 (3.07)	71	
	ΔGDP per capita			-.00 (.00)	71	
	ΔNeighborhood Tip			.01 (.18)	71	
	ΔWorker rights			-.02 (.79)	71	
Ferguson, Meghee & Woodside (2018)						RTA7 and TTA5 were constructed by Ferguson, et. Al. (2018) using data from Lynn & Lynn (2004), Magellan's (2009), and Conde Nast (2008) with one source per country – measures are not an index. Authors focus on a "configural analysis" that provides no tests of statistical significance. I focus on their standard correlation analyses. No deletion of outliers.
Restaurant Tip Amount 7 (RTA7)	MAS	.23			30	
	IND	-.23			30	
	PDI	.41*			30	
	UAI	.36*			30	
	Religiosity	.54*			30	
	GINI	.38*			30	
	PPP	-.35			30	
Taxi Tip Amount 5 (TTA5)	MAS	.06			30	
	IND	.24			30	
	PDI	.02			30	
	UAI	-.19			30	
	Religiosity	.16			30	
	GINI	-.01			30	

	PPP	.16			30	
[t] Effect reported as statistically significant at one-tailed $p < .05$. * Effect reported as statistically significant at two-tailed $p < .05$.						

References

Berger, S. (2019). This airline is asking passengers to tip their flight attendants. https: www.cnbc.com/2019/01/11/, accessed August 12, 2021.

Dewald, B. W. A. (2001). Tipping by foreign tourists. *Pacific Tourism Review*, *5*(1-2), 43-50.

Evans, M. and Dave, D. (1999). The thorny question of automatic service charges. *Cornell Hotel and Restaurant Administration Quarterly*, 40(4), 78-83.

Ferguson, G., Megehee, C. M., & Woodside, A. G. (2017). Culture, religiosity, and economic configural models explaining tipping-behavior prevalence across nations. *Tourism Management*, *62*, 218-233. https://doi.org/10.1016/j.tourman.2017.05.001

Ferguson, G., Megehee, C. M., & Woodside, A. G. (2018). Applying asymmetric, case-based, forecasting modeling in service research: Cultures' consequences on customers' service gratuities. *Australasian marketing journal*, *26*(4), 369-381. https://doi.org/10.1016/j.ausmj.2018.10.009

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2019). Is ego depletion real? An analysis of arguments. *Personality and social psychology review*, *23*(2), 107-131. https://doi.org/10.1177/1088868318762183

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (Eds.). (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Sage publications.

Inglehart, R. F., Basanez, M., & Moreno, A. (1998). *Human values and beliefs: A cross-cultural sourcebook*. University of Michigan Press.

Lynn, M. (1994). Neuroticism and the prevalence of tipping: A cross-country study. *Personality and Individual Differences*, *17*(1), 137-138. https://doi.org/10.1016/0191-8869(94)90270-4

Lynn, M. (1997). Tipping customs and status seeking: a cross-country study. *International Journal of Hospitality Management*, *16*(2), 221-224. https://doi.org/10.1016/S0278-4319(97)00007-8

Lynn, M. (2000). National character and tipping customs: the needs for achievement, affiliation and power as predictors of the prevalence of tipping. *International Journal of Hospitality Management*, *19*(2), 205-210. https://doi.org/10.1016/S0278-4319(00)00008-6

Lynn, M. (2000). National personality and tipping customs. *Personality and Individual Differences*, *28*(2), 395-404. https://doi.org/10.1016/S0191-8869(99)00109-9

Lynn, M. (2006). Tipping in restaurants and around the globe: An interdisciplinary review. *Handbook of contemporary behavioral Economics*, 648-666.

Lynn, M. (2008). Personality effects on tipping attitudes, self-reported behaviors and customs: A multi-level inquiry. *Personality and Individual Differences*, *44*(4), 989-999. https://doi.org/10.1016/j.paid.2007.10.025

Lynn, M. (2015). Service gratuities and tipping: A motivational framework. *Journal of Economic Psychology*, *46*, 74-88. https://doi.org/10.1016/j.joep.2014.12.002

Lynn, M. (2018). How motivations for tipping vary with occupational differences in descriptive tipping norms. *Journal of behavioral and experimental economics*, *77*, 1-10.

Lynn, M. (2021). The effects of occupational characteristics on the motives underlying tipping of different occupations. *Journal of Behavioral and Experimental Economics*, *95*, 101783.

Lynn, M. (2025). How have US restaurant tips changed over time?. *International Journal of Hospitality Management*, *124*, 103969.

Lynn, M., & Brewster, Z. W. (2020). The tipping behavior and motives of US travelers abroad: Affected by host nations' tipping norms?. *Journal of Travel Research*, *59*(6), 993-1007. https://doi.org/10.1177/0047287519875820

Lynn, M., & Lynn, A. (2004). National values and tipping customs: A replication and extension. *Journal of Hospitality & Tourism Research*, *28*(3), 356-364. https://doi.org/10.1177/1096348003256605

Lynn, M., & Starbuck, M. M. (2015). Tipping customs: The effects of national differences in attitudes toward tipping and sensitivities to duty and social pressure. *Journal of Behavioral and Experimental Economics*, *57*, 158-166. https://doi.org/10.1016/j.socec.2014.07.005

Lynn, M., Zinkhan, G. M., & Harris, J. (1993). Consumer tipping: A cross-country study. *Journal of Consumer Research*, *20*(3), 478-488. https://doi.org/10.1086/209363

Mansfield, E. D. (2016). The political economy of the itching palm: An analysis of tipping norms. *International studies quarterly*, *60*(3), 375-386. https://doi.org/10.1093/isq/sqw015

Minkov, M. (2008). Self-enhancement and self-stability predict school achievement at the national level. *Cross-cultural research*, *42*(2), 172-196. https://doi.org/10.1177/1069397107312956

Richards, R. and Rosao, D. (1995). Hottel; bans gratuities after tip from guest. *USA Today* (October 3), B9.

Schwartz, Z., & Cohen, E. (1999). Tipping and the nation's tax burden: A cross-country study. *Anatolia*, *10*(2), 135-147. https://doi.org/10.1080/13032917.1999.9686976

Simmons, J. P., & Simonsohn, U. (2017). Power posing: P-curving the evidence. *Psychological science*, *28*(5), 687-693. https://doi.org/10.1177/0956797616658563

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: a key to the file-drawer. *Journal of experimental psychology: General*, *143*(2), 534. https://psycnet.apa.org/doi/10.1037/a0033242

Starbuck, M. M. (2009). *A comparative study of tipping practices and attitudes* (Doctoral dissertation, Oxford University).

Tett, R. P., Toich, M. J., & Ozkum, S. B. (2021). Trait activation theory: A review of the literature and applications to five lines of personality dynamics research. *Annual Review of Organizational Psychology and Organizational Behavior*, *8*(1), 199-233.

Torfason, M. T., Flynn, F. J., & Kupor, D. (2013). Here is a tip: Prosocial gratuities are linked to corruption. *Social Psychological and Personality Science*, *4*(3), 348-354. https://doi.org/10.1177/1948550612454888

Wells, P. (2013). Leaving a tip: A custom in need of changing? *The New York Times*, (September 4), A1.