NÚMERO PUBLICACIÓN 246

DAVID CROW AND GERARDO MALDONADO

Are We Measuring Institutional Trust Right? Evidence from a Mexican Survey Experiment

SEPTIEMBRE 2013

CIDE

www.cide.edu

Dirección de Publicaciones
publicaciones@cide.edu
Tel. 5081 4003

Acknowledgements

This paper was presented at the 68th Annual Conference of the American Association for Public Opinion Research (AAPOR) in Boston, USA, and the Seminario de la División de Estudios Internacional (CIDE), in Mexico City, in June 2013. The authors thank the participants at these presentations for their comments, and especially Pablo Parás for his expert advice.

Abstract

Does using different scale produce substantively different evaluations of institutions? If so, why? Are there grounds for preferring one scale to another? We attempt to answer these questions by marshalling a rich set of survey experiments from the 2010 and 2012 Mexico, the Americas, and the World *surveys, carried out in Mexico (and other countries) under the aegis of the Centro de Investigación y Docencia Económicas (CIDE). The survey experiments asked split samples questions on institutional trust, as well as on migration and foreign policy (73 questions in all), providing one half-sample with a 4-pt. response scale and the other, a 7-pt. scale. Comparing the (rescaled) means of the half-samples, we find that the scales do, indeed, produce significantly different evaluations of institutions. But the size of the differences between the scales varies not only from institution to institution, but also among other topics we asked about. We show that the differences are greater when the topic is controversial, people know less about it, and attitudes are neither very positive nor very negative, but middling. Finally, we argue that the 4-pt. scale appears to measure institutional trust better than the 7-pt. scale, at least in the context of Mexico and the* Mexico, the Americas, and the World *survey.*

Key words: institutional trust, response scale, survey experiments, public opinion, foreign policy, Latin America.

Resumen

¿Usar diferentes escalas de respuesta produce evaluaciones sustantivamente diferentes de las instituciones? Si sí, ¿por qué? ¿Y existen razones para preferir un tipo de escala a otra? Pretendemos responder a estas preguntas utilizando un conjunto fructífero de experimentos en encuestas del proyecto de opinión pública México, las Américas y el Mundo *2010 y 2012, realizado en México (y otros países en Latinoamerica) bajo la coordinación del Centro de Investigación y Docencia Económicas (CIDE). Los experimentos de encuestas se hicieron preguntando a dos muestras divididas algunas preguntas sobre confianza institucional, así como migración y política exterior (73 preguntas en total), dando a una mitad de la muestra una escala de respuesta de 4 categorías y a la otra mitad una escala de 7 categorías. Comparando las medias (re-escaladas) de ambas muestras, encontramos que las escalas de respuesta, de hecho, producen diferencias significativas en la confianza en instituciones. Pero el tamaño de las diferencias entre escalas varía no solo de institución a institución, sino también entre los otros temas preguntados. Mostramos que las diferencias*

entre las medias son más grandes cuando el tópico es controversial, cuando los individuos sobre menos sobre éste, y cuando las actitudes no son muy positivas ni muy negativas, pero intermedias. Al final, sostenemos que la escala de respuesta de 4 categorías/puntos hace una mejor medida de la confianza institutional que la escala de 7 puntos, al menos en el contexto de México.

Palabras claves: confianza institucional, escalas de respuesta, experimentos de encuestas, opinión pública, política exterior, América Latina.

Introduction

Good institutions are vital to a healthy democracy. When they work right, institutions mediate between private individuals and the state, and between private individuals; express citizen interests; carry out government policy effectively (Citrin 1974); and curb political leaders' personal ambitions and autocratic tendencies (Dalton 2004; Norris 1999). All democracies need good institutions, but establishing fair, effective ones is especially important in the new democracies of Latin America and elsewhere, in which inept or corrupt institutions often placed themselves at the service of authoritarian rulers and their allies (Cleary and Stokes 2006; Svolik 2012). Political parties, legislatures, bureaucracies, courts, police forces—in a word, institutions—are often the loci of acrimonious conflict as new democracies struggle to overcome institutional deficits and achieve rule of law.

Gauging citizen perceptions of institutions rightly occupies a crucial place in cross-national survey research on democracy—in both old democracies and new democracies (Catterberg and Moreno 2006; Dalton 2004; Norris 2011). But there is no general agreement how to do this. Among other things, question wording and response sets differ from poll to poll. Two major cross-national surveys carried out in Latin America, the Latin American Public Opinion Project (LAPOP) and Latinobarometer, both have extensive batteries of questions about a wide variety of institutions, but use different response scales to measure citizen opinion toward them. Does using different scale produce substantively different evaluations of institutions? If so, why? Are there grounds for preferring one scale to another?

Here, we attempt to answer these questions by marshalling a rich set of survey experiments from the 2010 and 2012 *Mexico, the Americas, and the World* surveys, carried out in Mexico (and other countries) under the aegis of the Centro de Investigación y Docencia Económicas (CIDE).[1] The survey experiments asked split samples questions on institutional trust, as well as on migration and foreign policy (73 questions in all), providing one half-sample with a 4-pt. response scale and the other, a 7-pt. scale. Comparing the (rescaled) means of the half-samples, we find that the scales do, indeed, produce significantly different evaluations of institutions. But the size of the differences between the scales varies not only from institution to institution, but also among other topics we asked about. We show that the differences are greater when the topic is controversial, people know less about it, and attitudes are neither very positive nor very negative, but middling. Finally, we argue that the 4-pt. scale appears to measure institutional trust better

[1] Databases, questionnaires, and technical documentation are available for free at http://mexicoyelmundo.cide.edu.

than the 7-pt. scale, at least in the context of Mexico and the *Mexico, the Americas, and the World* survey.

We thus tie together two strands of scholarship—institutional trust and democratization, on the one hand, and survey methodology research on response scales, on the other—to suggest that some surveys might be giving us a distorted picture of citizens' confidence in institutions.

Of course, the 73 experimental items cover more ground, topically, than simply institutional. Some of our analyses do involve all the variables—if often as a foil to institutional trust—but we focus on trust for several reasons. First, as we emphasize above, the degree of confidence citizens have in institutions is a vital component of democratic development and, consequently, of comparative research on democratic development. Our findings here may have important ramifications for both how to conduct this research and what it has found. Second, precisely because institutional trust occupies a place of honor in comparative political research, multi-national surveys have tried to measure it, but each in its own way (two of which we test in our experiments). More than any other of the items included in our experiments, then, the institutional trust items afford the possibility for checking the results of our research against those of previous studies. Finally, as we detail below, response scale effects matter the most for the institutional trust items, in the sense of inducing the greatest differences between the 4- and 7-pt. scales.

The paper proceeds as follows. In the first section, we highlight the importance of institutions in research on democratization and review studies on response scale effects to theorize on why different response scales might produce answers to survey questions on institutional trust (and others). We then describe the *Mexico, the Americas, and the World* survey experiments, and compare the two experimental groups—the half-sample provided the 4-pt. response set and that offered the 7-pt. set—to demonstrate that they are equivalent. The third section describes the differences between mean responses for the two groups and show that the differences are greatest on the institutional trust items, and moderate to inexistent on other topics. In the following, fourth section, we explore reasons why some differences are greater than others, providing evidence that polarizing topics, and ones that fewer people have an opinion on, increase the magnitude of the differences between the scales. Fifth, based on the scales' distributional properties, internal consistency with other items in the surveys, and external consistency with other surveys, we posit that the 4-pt. scale gives us better information on what Mexicans really think about institutions in their countries. We conclude with some speculative remarks about the implications of our study on institutional trust research generally (and on survey methodology) and sketch an agenda for future research.

Measuring Institutional Trust Matters

Institutional Trust and Democracy

Since Tocqueville trust has been placed as the center of stable democracies. However, trust can be horizontally direct to fellow citizens or can be vertically referred to institutions. In modern societies, trust in institutions is, in general, relevant because it normalizes social interactions between people who neither know one another nor share a common social background. This trust is also associated with a "wide range of institutional mechanisms of social control that are designed, among other things, to reinforce trustworthy behavior" (Newton 2007: 351). Institutional trust refers to the citizen's assessment of core political democratic institutions; therefore "it has been said that entails an evaluation of the most relevant institutional attributes: credibility, fairness, transparency, and openness" (Torcal, Muñoz and Bonet 2012: 140).

What kind of attitude is institutional trust? It is not clear whether it is—using Easton's (1965, 1975) theoretical distinction—a sort of diffuse or specific political support. On the one hand, it is possible to affirm that political institutions are a core part of the political system of a country, so it is a systemic or diffuse support. Miller (1974) understood political trust as systemic confidence. On the other hand, it is possible that trust in institutions is conditional or contingent to performance of the regime, either political or economical. Citrin (1974) proposed trust in institutions as an indicator of the satisfaction with the performance of the system. The literature continues with the debate about diffuse and specific support (Price and Romatan 2004).

In order to work properly, political institutions need to be supported and, in that sense, trusted by the citizens (Fukuyama 1995; Clausen et al. 2009). In addition to that, institutional trust in democratic societies is relevant for both parts of the equation: the citizens and the regime. For citizens, institutional trust is important because it facilitates the control and accountability of elected politicians, making the flow of information at a lower cost as long as individuals feel confidence in the normal behavior of institutions (Lupia and McCubbins 1998). For regimes, institutional trust is important because that facilitates the compliance of individuals' behavior, eliminating the necessity of coercion (Catterberg and Moreno 2006: 32)—as it is true for non-democratic regimes (Svolik 2012).

In the societies with new and emerging democracies, as the one studied in this paper, institutional trust is more imperative for at least two reasons. On the one hand, new democracies are characterized basically by the emergence and transformation of political institutions, especially representative institutions—such as executive power, parliaments, and political parties. Therefore, it is reasonable to analyze the citizens' confidence in those institutions. On the other hand, although declining

confidence in institutions is a global phenomenon affecting almost all democracies (Dalton 2004; Norris 2011), the consequences for new democratic regimes could be dangerous. In established democracies, lower levels of institutional trust have motivated politicians and government to change and reform the functioning of institutions as a product of a more critical citizenry (Citrin and Luks 2001: 26). However, in new democracies the declining levels of institutional trust are related with a disillusionment and disaffection of individuals, leading to a problems of disaffection, abstention and anti-systemic behavior (Catterberg and Moreno 2006; Torcal and Lago 2006).

The study of trust in democratic regimes, especially in emerging ones, is important not only because it can demonstrate that there are differences in the levels (high or low) of institutional trust, but also because the nature of the trust is different among institutions, not all are explained by the same conditions (Kelleher and Wolak 2007). In that sense, the causes of institutional trust are seen as a reflection of disparities in society. There are three different causes or approaches to explain trust in institutions: social capital, performance, and values.

For many scholars, *social capital* is the most obvious cause of institutional trust. People with higher levels of social capital (either social trust or membership to social organizations) develop a higher sense of community and, as consequence, have a propensity to trust institutions (Keele 2007). This correlation was proposed first by Putnam (2000) for the United States, and then translated into different context like Europe (Zmerli et al. 2007) and Asia (Kim 2005), where social networks and associations are related negatively with trust in institutions. However, it has been shown that social trust has not a clear association with political trust. "And political trust is better explained by political rather than social variables—support for the governing party or coalition, national pride, interest in politics, and belief in open government" (Newton 2007; see also: Anderson and LoTempio 2002; Newton 1999; Newton and Norris 2000).

Based on the later, there is the *performance* explanation. For this approach, trust in institutions is explained by the economic results (Kelleher and Wolak 2007), personal satisfaction (Listhaug 1984), the State capacity to deliver goods (Edlung 2006) and corruption permissiveness (Anderson and Tverdova 2003; Clausen et al 2009), and all these has a consequence to economic development. In general, the idea is that while scarcity produces suspicious, well-being increases confidence in others, especially impersonal institutions.

Finally, there is the *values* explanation, inspired by the modernization theory. The general assumption is that generational changes in values system have an impact on the way citizens view institutions. In particular, the emergence of "post-materialist" values (emphasizing the life quality and self-expression behavior over physical and psychological needs) is a condition in

the change of support and trust in institutions (Inglehart 1997; Inglehart and Welzel 2005).

Response Scale Effects

Institutions, and popular trust in them, are undoubtedly a central concern of comparative political research. Thus, whether we are measuring trust correctly is also a topic of vital importance. In particular, we know that different response scales induce respondents to provide different substantive answers to the same question. In our case, we are interesting in comparing 4-pt. with 7-pt. scales to see if, effectively, there is any difference between the scales. For a scale to bias a response away from its true value, there must be both a *motive* for the respondent to answer other than truthfully, and an *opportunity* for her to do so. We look for a motive in "courtesy bias", a subset of social desirability theory. In addition to wanting to portray themselves as engaging in normatively desirable behavior—which could lead an interviewee to claim he had voted when he had not respondents also want to appear upbeat and agreeable. Thus, they may answer more positively, cheerfully, or optimistically than they really feel, or than the subject warrants.

We find that *opportunities* for differences to occur between response scales are provided by features inherent in the question topics and in the scales themselves. As we explain in greater detail below, controversial items on which respondents disagree among themselves or items about which few respondents know or have opinions (but on which they nonetheless provide responses) provide greater opportunities for the two scales to produce differences than do items on which respondents concur, or about which they know. Studies on "non-opinions" have established that where response scales provide an explicit "no opinion" or "don't know" response option have lesser variance than items that provide no such out for respondents (Krosnick). This is presumably because, where respondents are not told they can simply say they have no opinion when they do not have one, they engage in "satisficing" behavior and oblige the enumerator by giving a response that really masks the absence of an opinion.

Features of the scale itself might also slant answers one way or another. Scale features known to affect responses include the number of response options provided, whether the scale includes an explicit neutral midpoint, the range of the scale and whether it includes 0 or negative values, and whether the response categories are labelled ("anchored") semantically or numbered. More response categories may produce higher means (Moreno 2010). In addition, there is probably an upper limit to the number of categories between which a respondent can meaningfully distinguish beyond which responses become unreliable. Scholars are divided on whether a neutral midpoint makes for better responses by providing a theoretically

relevant response option (it is possible for people to be in the middle of two extremes) and avoiding forcing people to answer positively or negatively, as on an even-numbered scale, or whether it *worsens* responses by tacitly encouraging respondents to give a neutral non-opinion and group answers on the midpoint. Finally, semantic anchors can make responses more reliable by making scales more intelligible to respondents. In applying response scale effect theory to institutional trust, we note that there is no "typical measure" of trust in institutions. Citrin and Muste (1999) registered that from the 1960s until the end of the 1990s at least 25 different forms and indicators of trust.

We had hypothesized that courtesy bias would interact with features of the items (knowledge, polarization) and the scales (range, midpoint, semantic labels) to make the 7-pt. scale better than the 4-pt. scale. In theory, the 7-pt. scale provides respondents the opportunity—including a neutral midpoint—to shade responses and enable them to differentiate more finely between responses than did the coarser 4-pt. scale. On this point, at least, we were wrong, for reasons we discuss below. It may be that courtesy bias actually worked the opposite way: the 7-pt. scale seems to have induced respondents to provide more extreme responses.

Mexico, the Americas, and the World: Data and Experiments

Mexico, the Americas, and the World is a large-scale (N=2,400 in Mexico), ongoing, multi-national survey on public opinion and foreign policy, international affairs, national and supranational identities, migration, and other issues. It has been conducted in Mexico and other countries in the Americas (including, in different years, Brazil, Chile, Colombia, Ecuador, Peru, and the United States) biannually since 2004. Researchers at the Centro de Investigación y Docencia Económicas (CIDE, in Mexico City) coordinate the survey in collaboration with national teams in each of the participating countries.

Since 2006, the Mexican version of the survey has included an increasing number of survey experiments that vary question wording or response options to determine whether, and how, different ways of posing questions elicit substantively different answers. In 2010 and 2012, there were 10 experimental batteries of questions comprising 73 questions altogether on three broad thematic areas—institutional trust, migration, and foreign policy. The batteries consisted of series of sub-items prefaced by a common question stem and grouped together by topic, to wit: 1) confidence in Mexican and international institutions (2012); 2) rights that Mexicans living abroad should enjoy (2012); 3) perceptions of immigrants' contributions to Mexican society and the burdens they represent (2012); 4) a mirror image of (2) above—i.e., the same rights as for Mexicans abroad, but asked about foreign immigrants living in Mexico (2012); 5) priorities and goals for Mexican foreign policy

(2012); 6) the preferred means (military, diplomatic, cultural, and trade) for exerting Mexican influence in world affairs; 7) regional integration in Latin America (trade, migration, currency unification, etc.); 8) a slightly longer version of (2), rights for Mexicans abroad, asked in 2010; 9) the same as (3), perceptions of immigrants in Mexico, but in 2010; and 10) a slightly longer version of (4), rights for immigrants in Mexico, in 2010.

The experiments were simple. Each consisted of asking half of the 2,400 respondents to rate or agree with an item on a 4-pt. scale (for convenience's sake, the "control group") and the other half, on a 7-pt. scale (the "treatment group"). The 4-pt. scales are Likert-type items in which all four categories are anchored by semantic labels—"Not at all", "A little", "Somewhat", or "A lot"—a scheme common in both the United States and in cross-national surveys including World Values Survey (WVS), Gallup International, the International Social Survey Project (ISSP), and the Latinobarometer. In contrast, only the poles on the 7-pt. scale were anchored semantically, with the labels "Not at all" and "A lot"—also a common set-up used by LAPOP. [2] For example, the 4-pt. version of the institutional trust battery asked, "Please tell me how much you trust each of the following institutions, groups, or people. Do you trust them a lot, somewhat, a little, or not at all?" The 7-pt. version read, "On a scale of 1 to 7, where 1 means 'not at all', and 7 means 'a lot', please tell me how much you trust each of the following institutions, groups, or people." As in other surveys[3], support cards were provided for the 7-pt scale. The interviewer would then enumerate the institutions ("most people", "the police", "the Army", etc.) and record respondents' answers.[4]

Balance between Control and Treatment Groups

The experimental protocol called for randomization of survey respondents into control and treatment groups. Interviewers carried out ten interviews in each primary sampling unit (PSU, the polling place or *sección electoral*); in each packet of ten questionnaires, field supervisors collated the control (Version A) and treatment (Version B) versions to alternate such that in each PSU, an interviewer administered five of each type of questionnaire.[5]

[2] The European Social Survey (ESS) uses a larger scale: the 10-pt. categories of answer. The question is the follow: "Using this card, please tell me on a score of 0-10 how much you personally trust each of the institutions I read out. 0 means you do not trust an institution at all, and 10 means you have complete trust".

[3] E.g., LAPOP provides a support card for its 7-pt. scale.

[4] For an English translation of all question stems and response sets, please see Appendix A. Questions in the original Spanish may be consulted on the survey Web site: http://mexicoyelmundo.cide.edu.

[5] Strictly speaking, this process should be described as "quasi-randomization", since a respondent's probability of being assigned to the control or treatment group was not independent, but, on the contrary, depended on assignment of the previous respondent to one of these two groups. As we show, though, this assignment procedure achieved such a remarkable balance between the control and treatment groups that it was, for all intents and purposes, random.

Interviewers appeared to have followed the randomization protocol, and the protocol resulted in equivalent control and treatment groups. Unfortunately, the database does not contain interview dates, start and finish times, or interviewer codes that would enable us to ascertain whether interviewers did, indeed, alternate the control and treatment versions of the questionnaire—though successive observations in the database do alternate between the control and treatment questionnaires. But cross-tabulating the 240 PSUs with questionnaire type revealed that in all but one PSU, the ten interviews were divided evenly between the control and treatment versions, consistent with interviewers' carrying out the randomization instructions faithfully.

At all events, the control and treatment groups' compositions are homogenous; we can safely say that these groups were drawn from the same population. The purpose of randomization—and the reason randomized experiments are considered the "gold standard" of scientific research (in both the natural and, increasingly, the social sciences, see: Mutz 2011)—is, of course, to ensure that the control and treatment groups are equal in every respect, save for exposure to experimental treatment. Thus, any differences in outcomes between the two groups owe wholly to the experimental treatment, and not to some pre-existing difference between the groups that may have predisposed the experimental group to respond differently to treatment than the control group would have. If, in our case, women were inclined to evaluate institutions more generously than men, and a greater proportion of women were in the experimental group than in the control group, this would positively bias estimates of institutional trust on the 7-pt. scale—and, consequently, overestimate differences between the scales.

Experimental studies typically include evaluations of "covariate balance" to discard the possibility that treatment effects are really attributable (at least to some extent) to differences between the control and treatment groups. To determine whether these two groups are, effectively, the same here we carried out independent samples t-tests for differences of proportions and means on 14 sociodemographic and attitudinal covariates conceivably related to institutional trust, as we mentioned in the previous section. These were: sex, age, education (in years of schooling completed), income (categorized into 11 ranges), subjective perceptions of whether respondents' income allowed them met their needs (4-pt. scale), ideology (a left-right scale from 1 to 10), retrospective pocketbook economic evaluations (4-pt. scale from 1 to 4), interest in financial and economic news, in news on foreign affairs, and in political news (all on a 4-pt. scale), and general interest in international affairs (also a 4-pt. scale). In addition, we assessed whether the control and treatment groups differed on party identification for Mexico's three main parties (Institutional Revolutionary Party, PRI; National Action Party, PAN; and Party of the Democratic Revolution, PRD) by performing a Pearson's x^2 test on a cross-tabulation of questionnaire type and

party affiliation. Table 1 presents the results of our tests for covariate balance.

TABLE 1: DIFFERENCE OF MEANS BETWEEN EXPERIMENTAL AND CONTROL GROUPS ON SOCIODEMOGRAPHIC AND ATTITUDINAL VARIABLES

Group of variables	7-pt. Mean	4-pt. Mean	Difference 7pt. - 4pt.	*p*-value	Difference significant?
	(1)	(2)	(3)	(4)	(5)
Sex	0.494	0.495	-0.002	0.935	No
Age	41.528	41.366	0.162	0.807	No
Education	8.944	9.050	-0.106	0.568	No
Income (objective)	5.456	5.175	0.281	0.010	Yes
Income (subjective)	2.228	2.319	-0.091	0.008	Yes
Ideology (left-right)	5.764	5.842	-0.078	0.565	No
Retrospective Pocketbook Evaluation	2.172	2.201	-0.029	0.551	No
Interest in Economic News	2.547	2.551	-0.004	0.927	No
Interest in World Affair News	2.600	2.660	-0.060	0.166	No
Interest in Political News	2.687	2.734	-0.046	0.301	No
Interest in Foreign Affairs	2.436	2.406	0.030	0.448	No
Party Identification					
PAN					
(Frequency)	208	230			
(Expected Frequency)	219.2	218.8			
(Col %)	17.5%	19.4%			
PRI					
(Frequency)	374	372			
(Expected Frequency)	373.3	372.7			
(Col %)	31.5%	31.4%			
PRD					
(Frequency)	124	136			
(Expected Frequency)	130.1	129.9			
(Col %)	10.5%	11.5%			
Otro					
(Frequency)	481	447			
(Expected Frequency)	464.4	463.6			
(Col %)	40.5%	37.7%			
		X^2	2.908	0.406	No

The two groups are virtually identical. On all variables, save income (both the objective, monetary and subjective measures), there are no significant differences between the control and treatment groups. Even the observed differences on the two income variables, though statistically significant, were hardly overwhelming: an absolute difference of 0.281 (scale of 1 to 11, or 2.8% of the range) for "objective" income and an absolute difference of 0.090 (scale of 1 to 4, or 3.0% of the range) for "subjective" income. Differences between the two scales may be ascribed not to disparities between the control and treatment groups, but only to effects produced by the scales themselves.[6] We now describe those effects.

Different Response Choices, Different Answers

Do respondents answer questions about institutional trust differently when given a 7-pt. response scale than they would have when constrained to four response categories? That is, do different response scales induce different substantive evaluations of critical Mexican institutions? Yes. In fact, evaluations of institutional trust are, with few exceptions, higher on the 7-pt. scale than on the 4-pt. scale—a result that obtains across all questions, not just those on institutional trust.[7] But differences are lesser (moderate to non-existent) for questions on migration and foreign policy, which we attempt to explain in the following section.

To assess differences between the scales, we first rescale the 4- and 7-pt. items, projecting them on a common scale from 0 to 1.[8] Then, we conduct simple independent samples t-tests for differences of means for all 73 pairs of variables. Table 2 summarizes the results of the t-tests, in the aggregate, for each of the 10 experimental batteries of questions.

[6] Recently, a number of scholars have concerned themselves with increasing the validity of causal inferences in both experimental and observational studies. Estimates of treatment effects can be made more robust by both regressing outcome variables on auxiliary covariates (in addition to a treatment indicator) and by "matching" observations in the control group to one or more observations in the experimental group closest to the corresponding control group observations on a series of measured covariates. Treatment effects are calculated as the differences only between matched observations, rather than between the control and treatment groups on the whole. Though our covariate balance analysis suggests that these techniques are probably unnecessary, extensions of our research could correct for income imbalances by regression, matching, or both.

[7] This confirms a previous descriptive analysis. Moreno (2010: 41), using information from the Latinobarometer, LAPOP, WVS, and the Mexican ENCUP, shows that the 7-point and 10-point questions tend to produce higher levels of trust, perhaps because the more open scales (and higher scores), than the conventional 4-point questions.

[8] The rescaling formula is $\frac{x-x_{min}}{x_{max}-x_{min}}$, where x is the observed value, x_{min} is the minimum of the 4- or 7-pt. scale (in both cases, 1), and x_{max} is the maximum. Other studies that equate items measured on different scales standardize both scales (i.e., zero mean and variance of one) and then combine the values from the items on the standardized scale to carry out analyses. However, since our objectives included comparing means and standard deviations between the 4- and 7-pt. scales, standardizing the scales would clearly defeat the purpose of our experiment.

TABLE 2: AVERAGE DIFFERENCES FOR 10 GROUPS OF VARIABLES

Group of variables	Number of Questions	Average Difference 7pt. - 4pt.	Min. Difference	Max. Difference	Avg. *p*-value	Min. *p*-value	Max. *p*-value	Number of *p*-values Less than 0.05	Percent of Significant Differences
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Survey 2012									
Institutional Trust	15	0.061	0.030	0.089	0.001	0.000	0.013	15	100.0%
Rights for Mexicans Abroad	4	0.037	0.033	0.039	0.003	0.000	0.007	4	100.0%
Rights for Immigrants in Mexico	4	0.018	0.000	0.029	0.109	0.018	0.219	2	50.0%
Attitudes Toward Immigrants in Mexico	5	0.023	0.004	0.061	0.378	0.000	0.816	2	40.0%
Priorities for Mexican Foreign Policy	16	0.020	-0.010	0.055	0.123	0.000	0.812	7	43.8%
Means for Projecting Mexican Influence	4	0.053	0.043	0.068	0.000	0.000	0.000	4	100.0%
Regional Integration	8	0.038	0.000	0.081	0.187	0.003	0.849	6	75.0%
Survey 2010									
Rights for Mexicans Abroad	6	-0.007	-0.017	0.002	0.519	0.205	0.827	0	0.0%
Rights for Immigrants in Mexico	6	0.017	-0.024	0.049	0.080	0.000	0.151	2	33.3%
Attitudes Toward Immigrants in Mexico	5	0.022	0.000	0.049	0.303	0.001	0.986	2	40.0%

In general, the institutional trust items produce the greatest differences between the 4- and 7-pt. scales, followed by the foreign policy questions. Greatest agreement between the scales generally occurs on the migration variables. Differences on the 15 institutions rated by respondents trust averaged .061, ranging from .030 to .089.[9] All were statistically significant at $p < .000$. We observe smaller, but still important, differences for two of the three batteries on foreign policy: Mexicans' preferred means of wielding international influence (average difference of .053, with four of four variables significant at $p < .05$) and support for specific steps toward regional integration (average difference of .038, six of eight variables significant). One of the six experimental migration batteries, rights that Mexicans demand for their countrymen living abroad (2012), displays a similar level of differences (average difference of .037, four of four variables significant).

Differences for four (of six) migration batteries were distinctly fewer and smaller. They averaged .023 for 2012 perceptions of immigrants' contributions (two of five variables significant); .022 for rights Mexicans are willing to grant immigrants living in Mexico, 2012 (two of four variables significant); .022 for 2010 perceptions of immigrants contributions (two of five variables significant); and .017 for rights Mexicans are willing to grant immigrants living in Mexico, 2010 (two of six variables significant). Items in the remaining foreign policy battery, ratings of potential priorities for Mexican foreign policy, were in this same range, with average differences of .020 (seven of 16 variables significant). Finally, none of the differences for items in the remaining migration battery—rights Mexicans demand for their co-nationals living abroad, 2010—achieved statistical significance.

Honing in on institutional trust, Table 3 presents the rescaled means for the 7- and 4-pt. scales, the differences of means, and the p-values of the corresponding t-tests for each of the 15 institutions asked about in *Mexico, the Americas, and the World*. The table ranks institutions from greatest to least difference. The means are closest together for the police (.030), politicians (.041), and the Army (.041), and farthest apart for the United States government (.089), Mexican human rights organizations (.083), and multi-national corporations (.079).

[9] All differences are the rescaled mean of the 7-pt. scale minus the rescaled mean of the 4-pt. scale. Since the 7-pt. scale means are invariably higher than the 4-pt. means, all differences are positive.

TABLE 3: DIFFERENCES IN TRUST FOR 15 INSTITUTIONS, RANKED BY ABSOLUTE DIFFERENCE (MOST TO LEAST)

Rank	Name of Institution	7-pt. Mean	4-pt. Mean	Difference 7pt. - 4pt.	*p*-value
	(1)	(2)	(3)	(4)	(5)
1	US Government	0.421	0.510	0.089	0.000
2	Mexican HROs	0.552	0.635	0.083	0.000
3	MNCs	0.417	0.496	0.079	0.000
4	Intl. HROs	0.536	0.615	0.079	0.000
5	EU	0.422	0.501	0.079	0.000
6	UN	0.529	0.601	0.072	0.000
7	Congress	0.414	0.478	0.064	0.000
8	Banks	0.424	0.484	0.060	0.000
9	Domestic Business	0.583	0.642	0.059	0.000
10	President	0.517	0.565	0.048	0.000
11	Church	0.658	0.705	0.048	0.000
12	People	0.540	0.585	0.045	0.000
13	Army	0.597	0.638	0.041	0.002
14	Politicians	0.277	0.317	0.041	0.001
15	Police	0.370	0.400	0.030	0.013

Interestingly, agreement between the scales is greatest for some of the least trusted groups and institutions (police and politicians), but also for some of the most trusted (the Church and the Army). Disagreement appears to be greatest for controversial institutions (like the U.S. government and multi-national corporations) or lesser-known ones (like Mexican and international human rights organizations, and the European Union). We elaborate on and test these intuitions in the next section, in which we essay some hypotheses about why the differences of means vary across institutions and in general.[10]

Explaining Differences between the Scales

Why are differences between the scales larger for some questions than for others? We offer three explanations. First, the more controversial an institution, policy, or attitude is, the greater the difference between the 4-

[10] There is another sense (at least) in which the 4- and 7-pt. scales could yield different results: the rank order of the institutions, by means, could differ between scales. To determine whether rank orders are different, we carried out Wilcoxon signed rank tests both within each of the three broad thematic categories (institutions, migration, and foreign policy) and for all 73 variables together. The evidence is mixed. The scales produced different rank orders for institutions (z = 2.674, p = .008) and the migration variables (z = -2.204, p = .028), but not the foreign policy items (z = 1.154, p = .248) or overall (z = 0.370, p = .711).

and 7-pt. scales. Second, the less people know about an institution, policy, or attitude, the greater the difference between the scales. Third, differences between the scales will be less for institutions, policy, and attitudes with ratings (or levels of agreement) toward the extremes—i.e., either high or low—than in the middle of the range. We now formalize and test each of these hypothesizes.

H1: Greater polarization leads to greater differences.

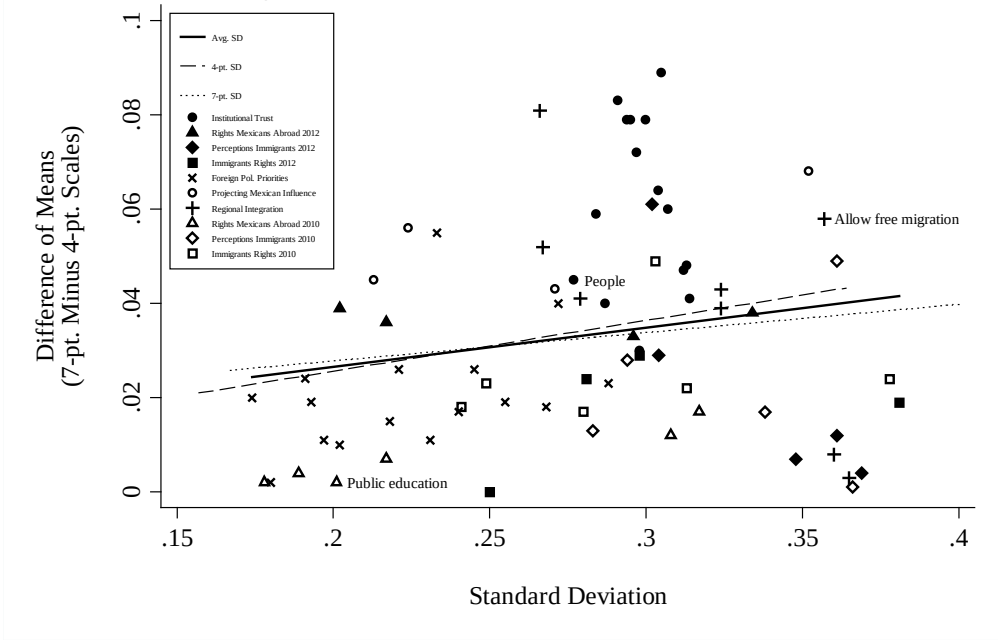
Where public opinion is divided on an institution or policy proposal, we hypothesize that this will lead to greater differences between the scales. A context of polarized opinion could, conceivably, magnify the effects induced by one of the scales relative to the other. We operationalize polarization as the standard deviation of an item. Where opinions diverge on a given proposition, opinions will spread out toward the extremes of the scale more than for propositions on which there is greater agreement. This spread translates into greater measures of dispersion, including the standard deviation.

We test this hypothesis with by simply regressing the differences between the 4- and 7-pt. scales for all 73 paired experimental items on their standard deviations. The hypothesis is $\beta_{SD} > 0$. Figure 1 is a scatterplot of differences against the standard deviations, averaged between the 4- and 7-pt. items, with the regression line ($\beta_{SDAvg} = .083$, $p = .051$)[11], shown as solid, fitted through the points. However, since it is not necessarily clear that the averaged standard deviations make for a fairer test of our hypothesis than do the standard deviations for just the 4- or just the 7-point items, we regressed the differences on the latter two as a robustness check. The results are virtually the same: the coefficient for the 4-pt. standard deviations, represented in Figure 1 as the dashed regression line, is $\beta_{SD4\text{-}pt} = .108$, $p = .025$, and for the 7-pt. standard deviations (dotted line), $\beta_{SD7\text{-}pt} = .060$, $p = .095$.[12]

[11] Since our hypotheses in this section are directional, this test—and all hypothesis tests in this section—are one-tailed.

[12] The horizontal, x-axis positions of the points for the 4- and 7-pt. standard deviations, of course, would differ slightly from the locations of the points shown in Figure 1, but we omit them from the scatterplot for clarity's sake.

Figure 1. Scatterplot of Differences of Means by Std. Devs. (with Regression Lines)



We label selected points on the scatterplot to give a more concrete feel for the import of our results. The vast majority of Mexicans believe that their compatriots living abroad (overwhelmingly in the United States) should have access to public education, as indicated by a low average standard deviation of around .20. There is virtually no disagreement, therefore, between the 4- and 7-pt. scales. At the other extreme, Mexicans are relatively divided about whether the unfettered flow of persons across international borders as part of a series of potential steps toward regional integration (standard deviation of about .37)—the so-called "North American" model of integration, which allows for free movement of capital, but not people (Gonzalez, Schiavon, et al. 2013). The difference between the scales is high (.059). In the middle is individual trust (standard deviation of .29). Mexican opinions on whether one can generally trust other people is more concentrated than for regional integration measures (represented by "+"), but less so than on rights for Mexicans abroad (the hollow triangles). The difference of means is also middling (.045).

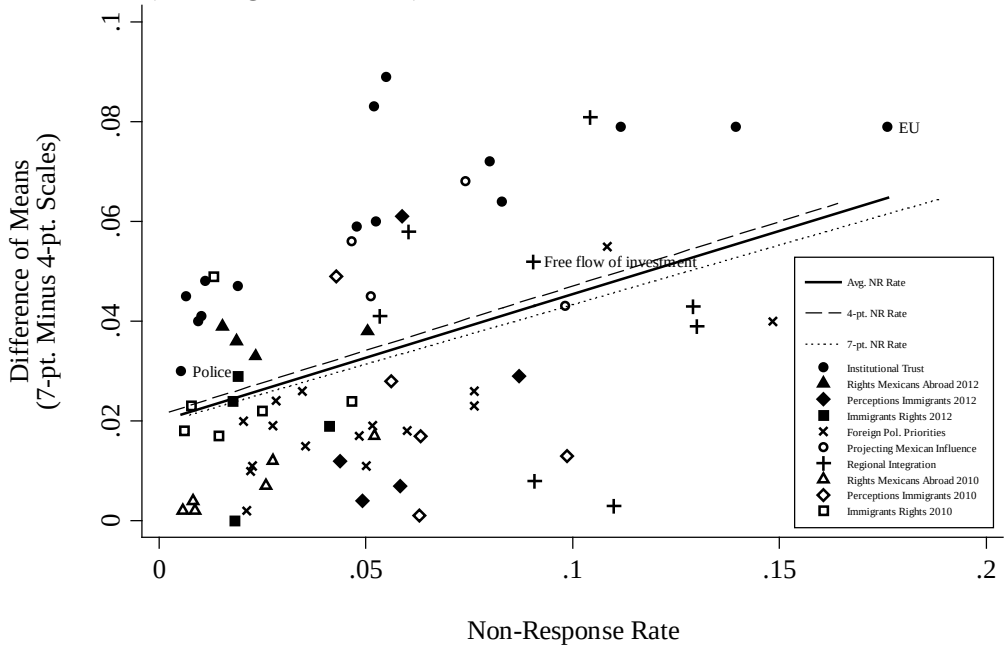
H2: Less knowledge leads to greater differences.

The less people know about an institution they are asked to rate, or attitude they are asked to agree with, the more variable their responses. Here, we take the item non-response rate—the percentage of people who answered "Don't Know" (DK), or failed to answer altogether—as a proxy for

average knowledge of a given question. Of course, item means are calculated only over respondents who answered the question (in the sense of giving a substantive response), and item non-response is a summary precisely of respondents who *did not* answer. How is it, then, that the percentage of DK responses (and "No Answers") might approximate the level of knowledge of people who did respond? We assume that for items with high non-response, the proportion of respondents with "non-attitudes"—quasi-random responses that mask the lack of an opinion on a given subject—is also likely to be high. More non-attitudes mean greater variability of responses (as indicated by measures of dispersion), and the lack of a firm opinion could, theoretically, make respondents more susceptible to scale effects.

We test this hypothesis by regressing differences of means on item non-response rates. The hypothesis is $\beta_{NR} > 0$. Figure 2 is a scatterplot of mean differences against average item non-response rates. As with Hypothesis 1, there are three possibilities for measuring non-response rates: the rates for items on the 4-pt. scale, those for the 7-pt. scale, and the average of the two. Again, we estimated regressions of mean differences on each of these to ensure our results are robust, and do not depend on using one measure or the other. The regression coefficients are large and significant in all cases: for the averaged non-response rates (solid regression line), $\beta_{NRAvg} = .254$, $p = .000$; for the 4-pt. scale non-response rates (dashed line), $\beta_{NRAvg} = .259$, $p = .000$; and for the 7-pt. non-response rates (dotted line), $\beta_{NRAvg} = .239$, $p = .000$.

Figure 2. Scatterplot of Differences of Means by Item Non-Response (with Regression Lines)



Again, we have labeled some points to help make the results more intuitively graspable. Police forces, about which nearly every Mexican has an opinion based on personal contact or, indirectly, through others' experiences (average item non-response rate of 0.5%), is the institution that registered the lowest difference of means (.030). The European Union, on the other hand, is removed from Mexicans' everyday experiences (non-response of 17.6%); it registers one of the highest differences of all variables, not just the institutional ones: .079. In the middle is the "free flow of investment" across national borders as one of the goals of regional integration. The topic is somewhat familiar to Mexicans through news reporting on the North American Free Trade Agreement, but not foremost among their concerns (non-response of 9.0%). The difference of means is .052.

H3a: Polarization is lowest for items with high or low means, and highest for items with means in the middle.

H3b: Differences of means are greatest for items with high or low means, and lowest for items with means in the middle.

Here, we venture two related hypotheses. First, Mexicans will agree more on institutions they ratings very well or very poorly, on policy measures they either support or reject overwhelmingly, and on attitudes they either accept or disavow strongly. In contrast, Mexicans will tend to *disagree* on their ratings of institutions, policies, and attitudes that fall between these extremes. In short, polarization (measured by an item's standard deviation, as in H1 above) will be greater for items with means at either end of the rating scale, and lower for those in the middle of the spectrum. If opinions are more concentrated around their means when these means are either high or low, we might expect the differences of means to be similarly lower. Where there is less disagreement, there is less room for scale effects to occur. In both cases, the points in a scatterplot of standard deviations (on the one hand) and differences of means (on the other) against item means should exhibit an inverted-U shape: low standard deviations, and differences of means, at low and high mean values, and high standard deviations, and differences of means, for midrange means. To test these hypotheses, we regressed standard deviations and differences of means, respectively, on the means and a squared term of the means. If there is, in effect, the hypothesized inverted-U shape, the first-order effect in each regression should be positive (i.e., $\beta_{MEAN} > 0$) and the quadratic term, negative (i.e., $\beta_{MEAN}^2 < 0$).

Figure 3a shows the results of the first regression—or, rather, set of regressions, since (as above) it is unclear whether the item means on the 4-

pt. scale, on the 7-pt. scale, or their average makes for the correct test. So, once again for the sake of robustness, we estimated three regressions: the average standard deviations on the average means (solid line), the 4-pt. standard deviations on the 4-pt. means (dashed line), and the 7-pt. standard deviations on the 7-pt. means (dotted line). The results are strong: for the average standard deviations $\beta_{MEANAvg} = 1.059$ ($p = .000$) and $\beta_{MEANAvg}^2 = -1.01$ ($p = .000$); 4-pt. scale, $\beta_{MEAN4\text{-}pt} = .949$ ($p = .000$) and $\beta_{MEAN4\text{-}pt}^2 = -.907$ ($p = .000$); and 7-pt. scale, $\beta_{MEAN4\text{-}pt} = 1.180$ ($p = .000$) and $\beta_{MEAN4\text{-}pt}^2 = -1.102$ ($p = .000$).

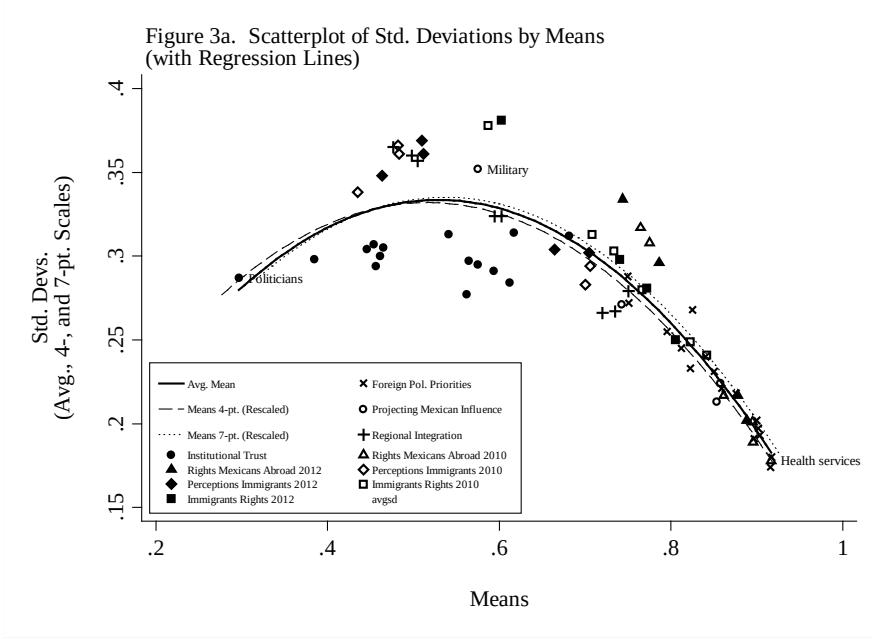
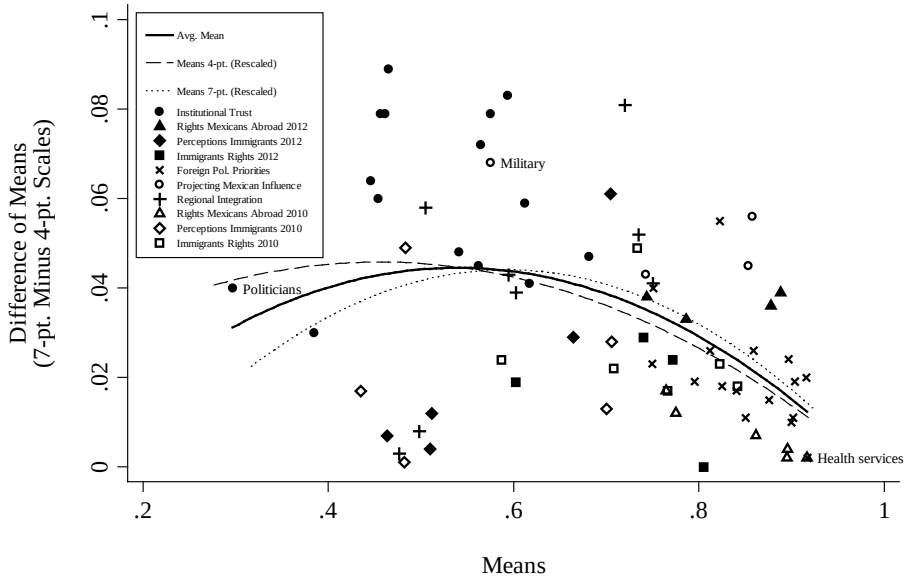


Figure 3b tells a familiar story. The figure plots difference of means on three different measures of means, the average means (solid line), the 4-pt. means (dashed line), and the 7-pt. means (dotted line). The ending is familiar as well: clear evidence of the inverted-U, though less pronounced than for the previous hypothesis. The coefficient values are: $\beta_{MEANAvg} = .242$ ($p = .042$) and $\beta_{MEANAvg}^2 = -.226$ ($p = .017$); 4-pt. scale, $\beta_{MEAN4\text{-}pt} = .148$ ($p = .121$) and $\beta_{MEAN4\text{-}pt}^2 = -.163$ ($p = .000$); and 7-pt. scale, $\beta_{MEAN4\text{-}pt} = .337$ ($p = .000$) and $\beta_{MEAN4\text{-}pt}^2 = -.284$ ($p = .000$).

Figure 3b. Scatterplot of Differences of Means by Means (with Regression Lines)



In both graphs, we highlight several representative points. Mexicans almost universally revile politicians (mean of .297). Judgments are relatively united around how bad they are (standard deviation of .287) and the difference of means is middling (.040), and among the smallest of the institutional trust items. The same happens with the idea that Mexican émigrés should have access to health services in the host countries, but at the other extreme: assent to this proposition is high (mean of .917); polarization, low (standard deviation of .179); and difference between the scales, inexistent (.002). Finally, the idea that Mexico should use its military to project Mexican foreign policy objectives in the world is a controversial idea (standard deviation of .352) that garners only lukewarm average support (mean of .575). This combination gives rise to a relatively large difference in means (.068).

Which Scale is better?

Given that the 4- and 7-pt. scales produce, under certain circumstances, very different judgments concerning institutions and policies—especially when opinion on a matter is divided, or when few people know about it—is there any way to know which scale gives us better results? We propose four criteria: 1) a good scale should produce answers with desirable distributional properties; 2) it should be easy to grasp cognitively; 3) it should "make sense" when compared to other data in the survey ("internal coherence"); and 4) it should produce results comparable to those of other surveys ("external

validity"). We address each of these criteria in turn and conclude that the 4-pt. scale measures institutional trust—and other variables—slightly, but distinctly, better.

Distributional Properties

A first cut at which scale produced better data implied looking at some of the univariate descriptive statistics to see whether responses were "better distributed" on the 4- or 7-pt. scale. What does "better distributed" mean? Responses on the better scale will more closely approximate a Gaussian distribution and congregate more tightly around the mean. In technical terms, responses will be less dispersed, less skewed, and less kurtotic. They should also "look good" intuitively, avoiding any glaring irregularities like bi- or multimodality, sharp spikes, and the like.

The 4-pt. scale appears superior on all these scores. As measured by standard deviations, the 4-pt. responses on the institutional trust items are slightly less dispersed than the 7-pt. ones. As Table 4 reveals, the average standard deviation across the 15 items for the 7-pt. scales was .309 and for the 4-pt. scales, .287. So, the 7-pt. scales' standard deviations were 2.2% higher than that for the 4-pt. scales, meaning estimates are slightly more precise for these.

TABLE 4: VARIANCE, SKEWNESS, AND KURTOSIS FOR INSTITUTIONAL VARIABLES

Name of Institution	Std. Dev.			Skewness			Kurtosis		
	7-pt.	4-pt.	Difference 7pt. - 4pt.	7-pt.	4-pt.	Absolute Difference 7pt. - 4pt.	7-pt.	4-pt.	Absolute Difference 7pt. - 4pt.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
People	0.299	0.255	0.044	-0.469	0.191	0.278	2.294	2.705	0.411
Police	0.309	0.287	0.022	0.269	-0.276	-0.007	1.976	2.269	0.293
Politicians	0.297	0.276	0.021	0.559	-0.615	-0.056	2.224	2.469	0.245
Army	0.319	0.309	0.010	-0.586	0.394	0.192	2.258	2.324	0.066
Mexican HROs	0.298	0.283	0.015	-0.569	0.201	0.368	2.429	2.453	0.024
President	0.328	0.297	0.031	-0.407	0.188	0.219	2.004	2.292	0.288
UN	0.306	0.288	0.018	-0.515	0.183	0.332	2.318	2.387	0.069
US Government	0.319	0.290	0.029	-0.224	-0.129	0.095	1.972	2.261	0.289
Banks	0.317	0.296	0.021	-0.075	-0.200	-0.125	1.905	2.286	0.381
Congress	0.316	0.292	0.024	-0.016	-0.141	-0.125	1.933	2.229	0.296
MNCs	0.305	0.282	0.023	-0.179	-0.085	0.094	2.072	2.272	0.200
EU	0.318	0.281	0.037	-0.202	-0.047	0.155	1.977	2.262	0.285
Intl. HROs	0.301	0.289	0.012	-0.519	0.206	0.313	2.376	2.396	0.020
Domestic Business	0.295	0.272	0.023	-0.725	0.244	0.481	2.693	2.686	-0.007
Church	0.315	0.308	0.007	-0.979	0.588	0.391	2.906	2.489	-0.417
Average			0.022			0.174			0.163

Skewness measures the "lopsidedness" of a distribution—that is, whether observations are grouped symmetrically around the mean or tilted leftward or rightward. Here, we take into account only "absolute" skewness [abs(0-skewness)], meaning that for purposes of our analysis a right-skewed distribution (i.e., where the median is to the *left* of the mean, shifting the probably mass leftward and making the right tail thicker) is as bad as a distribution equally skewed, but to the left (i.e., where the median is to the *right* of the mean, which shifts the peak rightward and thickens the left tail). Here, average absolute skewness was much higher, by .174 points, across the 15 trust items for the 7-pt. scale (average absolute skewness of .420) than for the 4-pt. scale (.246) (see Table 4). The differences are even greater across all 73 experimental items: 7-pt. average of 1.214, 4-pt. average of .838, difference of .376.

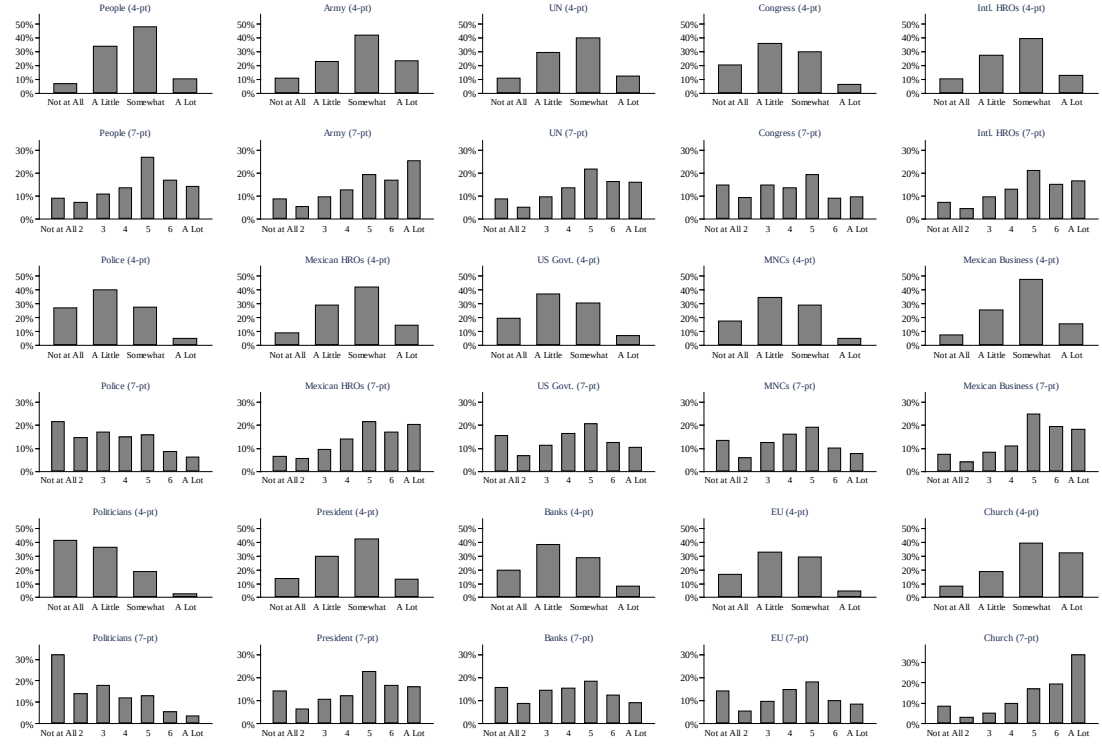
For its part, kurtosis measures the "peakedness" of a distribution. In a Gaussian distribution, the peak should be neither too high nor too flat, the tails neither too thick nor too thin. Again, we measure kurtosis in "absolute" terms [abs(3-kurtosis)], meaning that we give equal weight to *platykurtotic* distributions (in which observations from the peak and tails of a putatively Gaussian curve shift to the "shoulders", resulting in a broad, flat distribution) as we do to a *leptokurtotic* distribution (in which observations from the shoulders of the distribution migrate to the peak and tails, resulting in a high, thin peak and thick tails) in evaluating the goodness of a scale. For the 15 institutional trust items, taking the absolute value proved unnecessary, since "excess kurtosis" (values above or below 3) was only negative: all 15 variables in both scales were platykurtotic. But the 7-pt. scale was flatter—more platykurtotic—than the 4-pt. scale: average (negative) kurtosis for the former was -.778 and for the latter, .615, resulting in an absolute difference of .163. The absolute difference for all 73 variables was over double that: .380.

"Interocular inspection" of the univariate frequency distributions gives an idea of why the 7-pt. distributions have less desirable distributional properties. Figure 4 contains bar charts with the frequency distributions of the 15 institutional trust variables for the 4-pt. scales and, immediately below, on the 7-pt. scales for the same question. In general, responses on the 4-pt. scales are concentrated in the intermediate categories of "a little" or "some" trust, with fewer respondents in the extreme categories of "none" or "a lot".[13] The distributions resemble the familiar normal, "bell" curve. In contrast, the 7-pt. scales are comparatively misshapen—at times severely so. For most items, responses tend to fall at the extremes of "no" or "a lot" of trust. Trust in the United States government is especially illustrative. For the

[13] The sole exception is "politicians", which seem to be the object of special scorn in Mexico (as elsewhere). For this item, the model category is no trust, followed closely by "a little" trust. Even for the Church, an especially revered institution in Mexico, responses are concentrated in the middle category.

4-pt. scale, most response fall in the intermediate categories; on the 7-pt. scales, there were a lot of dissenters who did not trust the U.S. at all, relatively few people who used the intermediate scale options of "2" and "3", more who answered "4" and "5", and the frequencies taper off at "6" and "7" (a lot of trust). This leads to a bimodal distribution in with high frequencies at the negative extreme, and a normal-looking scale from "2" onward. In another pattern, respondents concentrate at the extremes, as for the Church and the army, where respondents located themselves overwhelmingly at the positive extreme of "7" ("a lot"), but also at the negative extreme of "no" trust—more than other categories ("2" and "3") to the left of the scale's midpoint ("4"). Thus, distributions for the 7-pt. scale are often bi- or even multi-modal (as for the Congress), comparatively flat (*platykurtotic*), and, therefore, more widely dispersed.

Figure 4. Frequency Distributions of Institutional Trust (4- and 7-pt. Scales)



Our expectation that the 7-pt. scale would give respondents more opportunities to nuance their responses, including by opting for the neutral midpoint. This should have yielded better distributional properties, we reasoned, for the 7-pt. than the 4-pt. scale and reduce the propensity toward

courtesy bias. It did not. Why? We're not sure. It could be that "primacy" or "recency" effects led respondents to answer the first or last options they were given in the question stem. It may also be that providing semantic labels for only the first and last categories afforded respondents a greater chance to vent their spleen, or express their enthusiasm, for institutions for which the 4-pt. response scale (again, contrary to our expectations) would have induced a more tempered response. Whatever the case, the 7-pt. scale—at least as we implemented it in *Mexico, the Americas, and the World*—seemed to pull respondents toward the extremes of the scale.

Cognitive Intelligibility

A good response scale should also be easy for respondents to understand. The number of points in the scale influences its intelligibility. Some research suggests that there's an upper limit to the number of categories On a scale between which respondents may meaningfully distinguish. The presence of semantic labels on all points, or explicitly on the midpoint (if there is one) may make a scale more comprehensible to an interviewee. Using support cards with graphic depictions of the scale may also make scales easier to understand. We operationalize intelligibility using item non-response rates: easy-to-understand scales should have low item non-response, whereas more people will decline to answer cognitively difficult scales. Item non-response also depends on the nature of the question being asked, of course, but pairing experimental items holds question design and content constant.

In general, the 4-pt. scales have slightly lower item non-response than the 7-pt. scales. Non-response for the 4-pt. scale ranged from .25% to 16.40%, with a mean of 4.64%; and for the 7-pt. scale, .75% to 18.85%, with a mean of 5.76%. The difference, 1.13%, is significant at $p = .000$. Of 73 items, differences in the non-response scale were significant for 22 (30%). All of these were positive, and ranged from nearly .91% to 5.36%.

Focusing on the institutional trust items, Table 5 presents the item non-response rates for each of the scales on the 15 institutional trust questions. On the 7-pt. scale, non-response ranged from .75% to 18.84% and averaged 6.34%; and on the 4-pt. scale, from .25% to 16.40%, averaging 5.13%. The difference of means across the 15 items, 1.22%, is significant at $p = .000$. Of the individual items, item non-response is significantly different for two (the Church and Mexican businesses) and close to it for another two (the police and international human rights organizations). In short, the 4-pt. scale elicits slightly higher response rates in the *Mexico, the Americas, and the World* survey, suggesting that it is more graspable for respondents than the 7-pt. scale.

Table 5: Response Rates for Institutional Trust Variables

Rank	Name of Institution	Item Non-Response (7-pt.)	Item Non-Response (4-pt.)	Difference 7pt. - 4pt.	Average Non-Response	*p*-value
		(1)	(2)	(3)	(4)	(5)
1	EU	18.8%	16.4%	2.4%	17.6%	0.116
2	MNCs	15.1%	12.8%	2.3%	14.0%	0.108
3	Intl. HROs	12.3%	10.0%	2.4%	11.2%	0.067
4	Congress	9.2%	7.4%	1.8%	8.3%	0.117
5	UN	8.8%	7.2%	1.7%	8.0%	0.129
6	US Government	5.8%	5.2%	0.5%	5.5%	0.585
7	Banks	5.8%	4.7%	1.0%	5.3%	0.268
8	Mexican HROs	5.2%	5.2%	-0.1%	5.2%	0.934
9	Domestic Business	6.0%	3.6%	2.4%	4.8%	0.005
10	Church	2.8%	1.1%	1.7%	1.9%	0.003
11	President	1.4%	0.8%	0.6%	1.1%	0.174
12	Army	1.2%	0.9%	0.3%	1.0%	0.544
13	Politicians	1.3%	0.7%	0.6%	1.0%	0.142
14	People	0.8%	0.6%	0.2%	0.7%	0.614
15	Police	0.8%	0.2%	0.6%	0.5%	0.051
	Average	6.3%	5.1%	1.2%		

Internal Coherence

The information yielded by posing a question a certain way another should be consistent with other data in the survey. That is, the associations, correlations, conditional distributions, etc., of the variable in question with other, related variables in the survey ought to conform (more or less) to intelligible patterns established by theory or other empirical work. Where an item is interpretable by other items in the survey—that is, where there is "internal coherence" among survey items—this may be a sign that the item was well-designed.

To test the internal coherence of the 15 institutional trust items with other variables in *Mexico the Americas, and the World*, we regressed each institution on the 14 variables contained in Table 1 (the covariate balance tests) thought to be related to institutional trust—recapitulating: sex, age, education, income, subjective perceptions of income, ideology, retrospective pocketbook economic evaluations, interest in economic news, in news on foreign affairs, and in political news, interest in international affairs, and sympathizing with the National Action Party (PAN), Party of the Democratic Revolution (PRD), or Institutional Revolutionary Party (PRI), Mexico's three main political parties. We postulate that the scale for which a higher number of these variables were significant across all 15 regressions is more internally coherent with other survey items.

Table 6 synthesizes the results of these regressions, showing the number of times (out of 15) a variable was significant at $p \leq .05$ across the 15 regressions for both the 4- and 7-pt. scales. Of 210 (14 x 15) total coefficients, 41 (or 19.5%) were significant for the 4-pt. scale, but only 24 (11.4%) were for the 7-pt. scale. Economic variables such as income and evaluations of one's own economy, relative to the past, have been signalled as important determinants of institutional trust. If one is better off, one is likelier to perceive that the political system and its institutions are trustworthy and work well (Kelleher and Wolak 2007; Mishler and Rose 2001). Both income and economic evaluations proved more predictive in the regressions on the 4-pt. versions of the trust variables than the 7-pt. versions: income was significant in six (of 15) 4-pt. regressions but only three of the 7-pt. regressions, and economic evaluations were significant in nine of the 4-pt. regressions, but only three of the 7-pt. regressions. Political orientations also influence institutional trust. Research on personality traits and ideological orientation shows that "conscientiousness" (self-discipline, orderliness, etc.) is associated with conservatism (Gallego and Oberski 20; Mondak 2012). Those toward the right of the 1-to-10 left-right ideology scale should therefore be more supportive of institutions and, indeed, are so, but far more in the 4-pt. regressions (six of 15 variables significant) than in the 7-pt. ones (just one of sixteen). It is unsurprising that those who identify as

sympathizers of the National Action Party (PAN)—conservative, confessional, laissez-faire—are more favorably inclined toward institutions, both as rightists and adherents of the party in power nationally at the time of the survey. *Panismo* is a significant determinant of institutional trust in four of the 4-pt. regressions, but just one of the 7-pt. regressions. In sum, the criterion of internal consistency tips the scales further toward the 4-pt. scale.

Table 6: Regression of Institutional Trust 15 Institutions (4-pt. Rescaled & 7-pt. Rescaled on Sociodemographic and Attitudinal Variables in Table 1)

	7 pt.	4 pt.
Name of Institution	Number of *p*-values less than 0.05	Number of *p*-values less than 0.05
Sex	0	0
Age	2	2
Education	1	4
Income (objective)	3	6
Income (subjective)	1	2
Political Affiliation		
PAN	1	4
PRI	3	1
PRD	1	0
Ideology (left-right)	1	6
Retrospective Pocketbook Evaluation	6	9
Interest in Economic News	2	2
Interest in World Affair News	2	1
Interest in Political News	1	0
Interest in Foreign Affairs	0	4
Total	24	41

External Validity

As we said before, another assessment of consistency of the 4-pt. versus 7-pt. scales is its comparability to other surveys that used the same or similar questions and scales. If our expectation is right, the 4-pt scale should produce results more comparable to those of other surveys than the 7-pt scale. In order to do so, we have analyzed the external validity of the *Mexico, las Americas y el Mundo* scales with other surveys using the same procedures. Fortunately to us, the measures of institutional trust are very often repeated

in cross-national investigations. It is a topic that has been gauge in different national and international, political and social surveys. However, not all surveys inquire about the same institutions: some emphasizes political and national institutions; others accentuate social and sub-national institutions; and in the case of the MAM, the interest is on international and national institutions.

Other regional projects in Latin America about public opinion give us the opportunity to compare both scales with their data. On the one hand, Latinobarometer has normally used the 4-pt scale of answers; on the other hand, the Latin American Public Opinion Project (LAPOP) has used the 7-pt scale for institutional trust. However, those projects have not asked individuals in Latin America about the same set of institutions as MAM. In the case of Latinobarometer, only seven institutions (police, politicians, army, President, Congress, church, and banks) were asked in both surveys; in the case of LAPOP, we can only compare only six institutions (the same as Latinobarometer, expect for banks). Moreover, surveys have not asked people about exactly the same institutions, but the results are very comparable. This is the case for a couple of institutions. First, whilst MAM asked about "politicians", Latinobarometer and LAPOP asked for "political parties". And second, whilst MAM asked for the "army", Latinobarometer and LAPOP asked for "armed forces". Finally, we need to mention that whist we compare the results of MAM and LAPOP for the same year (2012), we were not able—due to data availability—to do the same with Latinobarometer (we use the 2009 dataset).

To make a proper analysis we have collected information from the three surveys and merged them into a new single dataset. We do so in order to perform two different tests. First, we conduct simple independent sample t-tests for differences of means for the available institutions in both surveys. Second, we assessed whether both surveys samples differed on the distribution of responses by performing a Pearson's x^2 test on a cross-tabulation of type of answers and survey. The results are shown in Table 7.

As we can see, comparing the results of the MAM and the Latinobarometer, five out of seven institutions have statistically different means ($p < 0.000$) in both tests; and, in general, the MAM produces higher means that the Latinobarometer. However, it is important to mention that the institutions with the highest (church) and the lowest mean (politicians/parties) have not differences, consistently with previous results in this paper about polarizing institutions. Comparing the results of MAM and LAPOP, three out of six institutions have statistically different means ($p < 0.000$) in both tests. But, contrary to the previous comparison, those institutions where the difference is not statistically significant ($p > 0.000$)—police, President, church—have higher means in MAM than in LAPOP. What it is relevant to see here is that when we calculate the average absolute

differences in means and the Person's x^2 coefficients in both tests, we find two results. On the one hand, the 4-pt scale (used in MAM and Latinobarometer) has higher results in means than the 7-pt scale (used in MAM and LAPOP). It indicates that the 7-pt scale produces results relatively more comparable with other surveys than the 4-pt, contrary to what we expected. (We need to mention here that we are not comparing the same year samples here.) On the other hand, the results for the Person's x^2 coefficients show that the distributions are more comparable in the 4-pt scale than the 7-pt scale—although the difference in absolute terms of the coefficients mean is not high.

Table 7: Differences in Trust in Institutions, by different Latin American surveys

Name of Institution	4-pt Mean MAM	4-pt Mean Latino barometer	Difference	*p*-value	Pearson's χ^2	*p*-value
	(1)	(2)	(3)	(4)	(5)	(6)
Police	0.373	0.296	0.077	0.000	47.515	0.000
Politicians/Parties	0.277	0.265	0.011	0.315	3.407	0.333
Army	0.595	0.548	0.046	0.000	23.419	0.000
President	0.514	0.389	0.125	0.000	102.282	0.000
Congress	0.414	0.363	0.051	0.000	20.911	0.000
Church	0.657	0.642	0.015	0.261	33.684	0.000
Banks	0.419	0.373	0.046	0.001	19.591	0.001
Average			*0.053*		*35.829*	

Name of Institution	7-pt. Mean MAM	7-pt Mean LAPOP	Difference	*p*-value	Pearson χ^2	*p*-value
Police	0.400	0.397	0.003	0.797	13.610	0.034
Politicians/Parties	0.317	0.397	-0.079	0.000	70.092	0.000
Army	0.638	0.700	-0.061	0.000	36.429	0.000
President	0.565	0.561	0.004	0.728	25.792	0.000
Congress	0.478	0.536	-0.058	0.000	63.242	0.000
Church	0.705	0.689	0.015	0.209	10.790	0.095
Average			*0.036*		*36.659*	

Conclusions

Response scale options make a difference. In effect, the choice of one scale can induce respondents to answer differently than they would have on another scale. In our case, the 7-pt. scale yielded higher means on not only institutional trust questions, but also on questions about foreign policy and immigration, when compared to the 4-pt. scale. But these effects—biases, really—are especially pronounced for institutional trust, rightly a topic at the center of cross-national public opinion research. By using the 7-pt. scale, *Mexico, the Americas, and the World* may be overestimating institutional trust.[14]

Of course, it may be the case that rather than the 7-pt. scale's overestimating institutional trust, the 4-pt. scale is *underestimating* it. That is, the 7-pt. scale gives truer estimates that the 4-pt. scale biases downward. We do not think this is the case, though. For the reasons we set forth above, the 4-pt. scale seems to generate more trustworthy data. The data are more normally and precisely distributed; the scale is slightly easier to understand; the 4-pt. data are more consistent than the 7-pt. data with other information in the survey; and the 4-pt. scale is a little closer to what we observe on other surveys than the 7-pt. data.

The import of these findings is both methodological and substantive. Methodologically, the clear implication is a preference for the 4-pt. over the 7-pt. scale. However, we are reluctant to recommend wholesale abandonment of the 7-pt. scale without further testing, simply because we may not have employed it as well as we might have. Several options could be to employ an explicitly labelled (or "semantically anchored") midpoint; to provide support cards in face-to-face interviewers with some visual representation of the scale (as LAPOP does, to good effect, apparently); to alter the range so that rather than "no" and "a lot" of trust, the extremes are truly polar opposites (e.g., a numeric response scale from -3 to 3, or semantic labels of "completely disagree" to "completely agree"); and, finally, question stems could be modified—or interviewers instructed—to encourage respondents to use the entire range of available responses to express their opinions.

Of deeper concern, though, are our research's repercussions for previous studies on institutional trust. Is research based on the 7-pt. scale wrong? If our conjectures are correct, it may be that some studies have

[14] We are not sure whether this conjecture obtains for other surveys using 7-pt. scales. In particular, the Latin American Public Opinion Project (LAPOP) surveys did not suffer from bimodal distributions with respondents concentrated at the extremes of the scales, at least on comparable institutional trust items in *Mexico, the Americas, and the World* (MAM). One difference in survey administration was the LAPOP used visual support cards, which may have encouraged respondents to use the middle range of the scale more than they did on MAM.

overestimated the extent to which Mexicans, Latin Americans, and citizens worldwide trust in their institutions. This would be an especially grave finding in countries where institutional trust is low even on the 7-pt. scale; if already low levels of trust turn out to be *overestimated* the situation is perhaps more worrisome than we thought, especially in new democracies that are struggling to overcome deficits of institutional performance and establish effective, fair governance. Furthermore, our study seems to point toward reevaluating research that explores the determinants of institutional trust, or its effects on other aspects of democracy—particularly political behavior such as voting, protest, and the like. If institutional trust is mismeasured, do we need to revise results based on this mismeasurement? In particular, well known results of measurement theory establish that measurement error increases item variance—as is the case with the 7-pt. scale here—and leads to "attenuation bias" (overestimating coefficients by failing to account for measurement error) for coefficients of independent variables plagued by measurement error. Thus, one area of future research should be reanalysis of previous results by simulating trust variables on a 4-pt. scale based based on known (or assumed) conditional distributions of these on independent variables.

In short, we view this study not as a revolutionary overthrowing of a robust body of research on institutional trust, but as a starting point for asking whether we survey researchers are measuring institutional trust well and, if not, what the consequences are for what we know—or think we know—about institutional trust.

Annexes

Survey 2012

(78A) Please tell me, how much trust do you have in each of the following institutions, groups or people? Do you have A lot, some, a little, or no trust in...?

	A lot	Some	A little	No trust	DK	NA
The Majority of People	(1)	(2)	(3)	(4)	(8)	(9)
The police	(1)	(2)	(3)	(4)	(8)	(9)
Politicians	(1)	(2)	(3)	(4)	(8)	(9)
The Army	(1)	(2)	(3)	(4)	(8)	(9)
Mexican Human Rights Organizations	(1)	(2)	(3)	(4)	(8)	(9)
The President	(1)	(2)	(3)	(4)	(8)	(9)
The United Nations	(1)	(2)	(3)	(4)	(8)	(9)
The United States Government	(1)	(2)	(3)	(4)	(8)	(9)
The Banks	(1)	(2)	(3)	(4)	(8)	(9)
The Congress	(1)	(2)	(3)	(4)	(8)	(9)
Multinational Corporations	(1)	(2)	(3)	(4)	(8)	(9)
The European Union	(1)	(2)	(3)	(4)	(8)	(9)
International Human Rights Organizations	(1)	(2)	(3)	(4)	(8)	(9)
Domestic Business	(1)	(2)	(3)	(4)	(8)	(9)
The Church	(1)	(2)	(3)	(4)	(8)	(9)

(78B) On a scale of 1 to 7, where 1 means Nothing and 7 A lot, please tell me, how much trust do you have in each of the following institutions, groups or people?

	Grade	DK	NA
The Majority of People		(8)	(9)
The police		(8)	(9)
Politicians		(8)	(9)
The Army		(8)	(9)
Mexican Human Rights Organizations		(8)	(9)
The President		(8)	(9)
The United Nations		(8)	(9)
The United States Government		(8)	(9)
The Banks		(8)	(9)
The Congress		(8)	(9)
Multinational Corporations		(8)	(9)
The European Union		(8)	(9)
International Human Rights Organizations		(8)	(9)
Domestic Business		(8)	(9)
The Church		(8)	(9)

(10A) How much would you agree or disagree with Mexicans who live abroad having the right to... Would you say you Strongly agree, somewhat agree, somewhat disagree or strongly disagree?

	Strongly Agree	Somewhat Agree	Somewhat	Strongly Disagree	DK	NA
Take their family to live with them	(1)	(2)	(3)	(4)	(8)	(9)
Have access to public education	(1)	(2)	(3)	(4)	(8)	(9)

Have access to public healthcare	(1)	(2)	(3)	(4)	(8)	(9)
Get a job under the same conditions as that country's citizens	(1)	(2)	(3)	(4)	(8)	(9)
Vote in their country of residence	(1)	(2)	(3)	(4)	(8)	(9)

(10B) On a scale of 1 to 7, where 1 means Completely Disagree and 7 Completely Agree. Please tell me, how much would you agree or disagree with Mexicans who live abroad having the right to…

	Grade	DK	NA
Take their family to live with them		(8)	(9)
Have access to public education		(8)	(9)
Have access to public healthcare		(8)	(9)
Get a job under the same conditions as that country's citizens		(8)	(9)
Vote in their country of residence		(8)	(9)

(13A) In general, how much do you agree or disagree with the following statements regarding foreigners living in México? Do you strongly agree, somewhat agree, somewhat disagree or strongly disagree? Foreigners who live in México…

	A lot	Some	A little	No trust	DK	NA
Contribute to Mexican economy	(1)	(2)	(3)	(4)	(8)	(9)
Take jobs away from Mexicans	(1)	(2)	(3)	(4)	(8)	(9)
Compromise security	(1)	(2)	(3)	(4)	(8)	(9)
Bring innovative ideas	(1)	(2)	(3)	(4)	(8)	(9)
Weaken our traditions and customs	(1)	(2)	(3)	(4)	(8)	(9)

(13B) On a scale of 1 to 7, where 1 means Completely Disagree and 7 Completely Agree, how much do you agree or disagree with the following statements regarding foreigners living in México? Foreigners who live in México…

	Grade	DK	NA
Contribute to Mexican economy		(8)	(9)
Take jobs away from Mexicans		(8)	(9)
Compromise security		(8)	(9)
Bring innovative ideas		(8)	(9)
Weaken our traditions and customs		(8)	(9)

(26A) How much do you agree or disagree with foreigners who live in México being allowed to… Would you say you strongly agree, somewhat agree, somewhat disagree or strongly disagree?

	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree	DK	NA
Bring their family to live with them	(1)	(2)	(3)	(4)	(8)	(9)
Have access to public education	(1)	(2)	(3)	(4)	(8)	(9)
Have access to public healthcare	(1)	(2)	(3)	(4)	(8)	(9)
Get a job under the same conditions as Mexicans citizens	(1)	(2)	(3)	(4)	(8)	(9)
Vote in México	(1)	(2)	(3)	(4)	(8)	(9)

(26B) On a scale of 1 to 7, where 1 means Completely Disagree and 7 Completely Agree. Please tell me, how much would you agree or disagree with foreigners who live in México being allowed to...

	Grade	DK	NA
Bring their family to live with them		(8)	(9)
Have access to public education		(8)	(9)
Have access to public healthcare		(8)	(9)
Get a job under the same conditions as Mexicans citizens		(8)	(9)
Vote in México		(8)	(9)

(32A) How important should each one of the following objectives be for México's foreign policy: very important, somewhat important, barely important or not important?

	Very important	Somewhat important	Barely important	Not important	DK	NA
Strengthening the United Nations Organization (UN) (1)	(1)	(2)	(3)	(4)	(8)	(9)
Strengthening the Organization of American States (OAS) (2)	(1)	(2)	(3)	(4)	(8)	(9)
Combating international terrorism (3)	(1)	(2)	(3)	(4)	(8)	(9)
Helping spread democracy to other countries (4)	(1)	(2)	(3)	(4)	(8)	(9)
Promoting sales of Mexican products in other countries (5)	(1)	(2)	(3)	(4)	(8)	(9)
Helping improve the standard of living in less-developed countries (6)	(1)	(2)	(3)	(4)	(8)	(9)
Preventing the spread of nuclear weapons (7)	(1)	(2)	(3)	(4)	(8)	(9)
Fighting drug trafficking and organized crime (8)	(1)	(2)	(3)	(4)	(8)	(9)

Protecting Mexican interests in other countries (9)	(1)	(2)	(3)	(4)	(8)	(9)
Bringing foreign investment to México (10)	(1)	(2)	(3)	(4)	(8)	(9)
Protecting our land and sea borders (11)	(1)	(2)	(3)	(4)	(8)	(9)
Protecting the environment (12)	(1)	(2)	(3)	(4)	(8)	(9)
Promoting regional integration (13)	(1)	(2)	(3)	(4)	(8)	(9)
Protecting and promoting human rights in other countries (14)	(1)	(2)	(3)	(4)	(8)	(9)
Attracting tourists (15)	(1)	(2)	(3)	(4)	(8)	(9)
Promoting Mexican culture (16)	(1)	(2)	(3)	(4)	(8)	(9)

(32B) On a scale of 1 to 7, where 1 means Not Important and 7 Very Important, please tell me, how important should each one of the following objectives be for México's foreign policy?

	Grade	DK	NA
Strengthening the United Nations Organization (UN) (1)		(8)	(9)
Strengthening the Organization of American States (OAS) (2)		(8)	(9)
Combating international terrorism (3)		(8)	(9)
Helping spread democracy to other countries (4)		(8)	(9)
Promoting sales of Mexican products in other countries (5)		(8)	(9)
Helping improve the standard of living in less-developed countries (6)		(8)	(9)
Preventing the spread of nuclear weapons (7)		(8)	(9)
Fighting drug trafficking and organized crime (8)		(8)	(9)
Protecting Mexican interests in other countries (9)		(8)	(9)
Bringing foreign investment to México (10)		(8)	(9)
Protecting our land and sea borders (11)		(8)	(9)
Protecting the environment		(8)	(9)

(12)			
Promoting regional integration (13)		(8)	(9)
Protecting and promoting human rights in other countries (14)		(8)	(9)
Attracting tourists (15)		(8)	(9)
Promoting Mexican culture (16)		(8)	(9)

(37A) In order to increase México's influence in the world, how much do you agree with México utilizing the following resource... would you say you strongly agree, somewhat agree, somewhat disagree or strongly disagree?

	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree	DK	NA
Military	(1)	(2)	(3)	(4)	(8)	(9)
Diplomatic	(1)	(2)	(3)	(4)	(8)	(9)
Cultural	(1)	(2)	(3)	(4)	(8)	(9)
Commercial	(1)	(2)	(3)	(4)	(8)	(9)

(37B) On a scale of 1 to 7, where 1 means Completely Disagree and 7 Completely Agree. Please tell me, in order to increase México's influence in the world, how much do you agree with México utilizing the following resource...?

	Grade	DK	NA
Military		(8)	(9)
Diplomatic		(8)	(9)
Cultural		(8)	(9)
Commercial		(8)	(9)

(51A) How much do you agree or disagree with the following actions being taken to favor Latin American integration? Do you strongly agree, somewhat agree, somewhat disagree or strongly disagree with...?

	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree	DK	NA
Allowing the free movement of goods and services across	(1)	(2)	(3)	(4)	(8)	(9)

the region						
Allowing the free movement of investments across the region	(1)	(2)	(3)	(4)	(8)	(9)
Allowing the free movement of people across the region, without border controls	(1)	(2)	(3)	(4)	(8)	(9)
Creating a Latin American currency	(1)	(2)	(3)	(4)	(8)	(9)
Building roads and bridges to connect the region	(1)	(2)	(3)	(4)	(8)	(9)
Creating a Latin American parliament or congress to write common laws	(1)	(2)	(3)	(4)	(8)	(9)
Adopt a common Latin American external policy	(1)	(2)	(3)	(4)	(8)	(9)
Forming a Latin American army	(1)	(2)	(3)	(4)	(8)	(9)
Create joint business with other Latin American countries	(1)	(2)	(3)	(4)	(8)	(9)

(51B) On a scale of 1 to 7, where 1 means Completely Disagree and 7 Completely Agree. How much do you agree or disagree with the following actions being taken to favor Latin American integration?

	Grade	DK	NA
Allowing the free movement of goods and services across the region		(8)	(9)
Allowing the free movement of investments across the region		(8)	(9)
Allowing the free movement of people across the region, without border controls		(8)	(9)
Creating a Latin American currency		(8)	(9)
Building roads and bridges to connect the region		(8)	(9)
Creating a Latin American parliament or congress to write common laws		(8)	(9)
Adopt a common Latin American external policy		(8)	(9)
Forming a Latin American army		(8)	(9)
Create joint business with other Latin American countries		(8)	(9)

Survey 2010

(6A) How much would you agree or disagree with Mexicans who live abroad having the right to... Would you say you Strongly agree, somewhat agree, somewhat disagree or strongly disagree?

	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree	DK	NA
Take their family to live with them	(1)	(2)	(3)	(4)	(8)	(9)
Have access to public education	(1)	(2)	(3)	(4)	(8)	(9)
Have access to public healthcare	(1)	(2)	(3)	(4)	(8)	(9)

Get a job under the same conditions as that country's citizens	(1)	(2)	(3)	(4)	(8)	(9)
Vote in their country of residence	(1)	(2)	(3)	(4)	(8)	(9)

(6B) On a scale of 1 to 7, where 1 means Completely Disagree and 7 Completely Agree. Please tell me, how much would you agree or disagree with Mexicans who live abroad having the right to…

	Grade	DK	NA
Take their family to live with them		(8)	(9)
Have access to public education		(8)	(9)
Have access to public healthcare		(8)	(9)
Get a job under the same conditions as that country's citizens		(8)	(9)
Vote in their country of residence		(8)	(9)

(10A) In general, how much do you agree or disagree with the following statements regarding foreigners living in México? Do you strongly agree, somewhat agree, somewhat disagree or strongly disagree? Foreigners who live in México…

	A lot	Some	A little	No trust	DK	NA
Contribute to Mexican economy	(1)	(2)	(3)	(4)	(8)	(9)
Take jobs away from Mexicans	(1)	(2)	(3)	(4)	(8)	(9)
Compromise security	(1)	(2)	(3)	(4)	(8)	(9)
Bring innovative ideas	(1)	(2)	(3)	(4)	(8)	(9)
Weaken our traditions and customs	(1)	(2)	(3)	(4)	(8)	(9)

(10B) On a scale of 1 to 7, where 1 means Completely Disagree and 7 Completely Agree, how much do you agree or disagree with the following statements regarding foreigners living in México? Foreigners who live in México…

	Grade	DK	NA
Contribute to Mexican economy		(8)	(9)
Take jobs away from Mexicans		(8)	(9)
Compromise security		(8)	(9)

Bring innovative ideas		(8)	(9)
Weaken our traditions and customs		(8)	(9)

(21A) How much do you agree or disagree with foreigners who live in México being allowed to... Would you say you strongly agree, somewhat agree, somewhat disagree or strongly disagree?

	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree	DK	NA
Bring their family to live with them	(1)	(2)	(3)	(4)	(8)	(9)
Have access to public education	(1)	(2)	(3)	(4)	(8)	(9)
Have access to public healthcare	(1)	(2)	(3)	(4)	(8)	(9)
Get a job under the same conditions as Mexicans citizens	(1)	(2)	(3)	(4)	(8)	(9)
Vote in México	(1)	(2)	(3)	(4)	(8)	(9)

(21B) On a scale of 1 to 7, where 1 means Completely Disagree and 7 Completely Agree. Please tell me, how much would you agree or disagree with foreigners who live in México being allowed to...

	Grade	DK	NA
Bring their family to live with them		(8)	(9)
Have access to public education		(8)	(9)
Have access to public healthcare		(8)	(9)
Get a job under the same conditions as Mexicans citizens		(8)	(9)
Vote in México		(8)	(9)

References

Anderson, Christopher J. and A.J. LoTempio (2002). "Winning, Losing, and Political Trust in America", *British Journal of Political Science*, 32: 335-351.

Anderson, Christopher J. and Y.V. Tverdova (2003). "Corruption, Political Allegiances, and Attitudes toward the Government in Contemporary Democracies", *American Journal of Political Science*, Vol. 47: 91-109.

Catterberg, Gabriela and Alejandro Moreno (2006). "The Individual Bases of Political Trust: Trends in New and Established Democracies", *International Journal of Public Opinion Research*, Vol. 18: 31-48.

Citrin, Jack (1974). "Comment: The Political Relevance of Trust in Government", *American Political Science Review*, Vol. 68: 973-988.

Citrin, Jack and Christopher Muste (1999). "Trust in Government", in John P. Robinson, Phillip R. Shaver and Lawrence S. Wrightsman, eds., *Measures of Political Attitudes*, New York: Academic Press.

Citrin, Jack and Samantha Luks (2001). "Political Trust Revisited: Déjà vu All Over Again", in J. R. Hibbins and E. Theiss-Morse, eds., *What is It about Government that American Dislike*, New York: Cambridge University Press.

Clausen, Bianca, Kraay Aart and Nyiri Zsolt (2009). "Corruption and Confidence in Public Institutions: Evidence from a Global Survey", Development Research Group-World Bank, Policy Research Paper 5157.

Cleary, Matthew R. and Susan Stokes (2006). *Democracy and the Culture of Skepticism: Political Trust in Argentina and Mexico*, New York: Russell Sage Foundation.

Dalton, Russell (2004). *Democratic Challenges, Democratic Choices: The Erosion of Political Support in Advanced Industrial Democracies*, Oxford: Oxford University Press.

Easton, David (1965). *A System Analysis of Political Life*, New York: Wiley.

Easton, David (1975). "A Re-assessment of the Concept of Political Support", *British Journal of Political Science*, 5: 435-457.

Edlung, Jonas (2006). "Trust in Capability of the Welfare State and General Welfare State Support: Sweden 1997-2002", *Acta Sociologica*, Vol. 49: 395-417.

Fukuyama, Francis (1995). *Trust: Social Virtues and the Creation of Prosperity*, New York: Free Press.

Gallego, Aina and Daniel Oberski (2011). "Personality and Political Participation: The Mediation Hypothesis", *Political Behavoir*.

Inglehart, Ronald (1997). *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*, Princeton: Princeton University Press.

Inglehart, Ronald and Christian Welzel (2005). *Modernization, Cultural Change, and Democracy: The Human Development Sequence*, New York: Cambridge University Press.

Keele, Luke (2007). "Social Capital and the Dynamics of Trust in Government", *American Journal of Political Science*, Vol. 51: 241-254.

Kelleher, Christine A. and Jennifer Wolak (2007). "Explaining Public Confidence in the Branches of State Government", *Political Research Quarterly*, Vol. 60: 707-721.

Kim, Ji-Young (2005). "Bowling Together' Isn't a Cure-All: The Relationship between Social Capital and Political Trust in South Korea", *International Political Science Review*, Vol. 26: 193-213.

Listhaug, Ola (1984). "Confidence in Institutions: Findings from the Norwegian Values Study", *Acta Sociologica*, Vol. 27: 111-122.

Lupia, Arthur and M. McCubbins (2008). *Democratic Dilemma: Can Citizens Learn What They Need to Know?*, New York: Cambridge University Press.

Miller, Arthur (1974). "Political Issues and Trust in Government, 1964-1970", *American Political Science Review*, Vol. 68: 951-972.

Mishler, William and Richard Rose (2001). "What are the Origins of Political Trust? Testing Institutional and Cultural Theories in Post-Communist Societies", *Comparative Political Studies*, Vol. 34: 30-62.

Mondak, J. J. (2010). *Personality and the foundations of political behavior.* Cambridge: Cambridge University Press

Moreno, Alejandro (2010). "Introducción", in Alejandro Moreno, ed., *Confianza en las instituciones: México en perspectiva comparada*, Mexico City: CESOP-ITAM.

Mutz, Diana C. (2011). *Population-Based Survey Experiments*, Princeton: Princeton University Press.

Newton, Kenneth (1999). "Social and Political Trust in Established Democracies", in Pippa Norris, ed., *Critical Citizens*, New York: Oxford University Press.

Newton, Kenneth (2007). "Social and Political Trust", in Russell Dalton and Hans-Dieter Klingemann, eds., *Oxford Handbook of Political Behavior*, New York: Oxford University Press.

Newton, Kenneth and Pippa Norris (2000). "Confidence in Public Institutions: Faith, Culture, or Performance?", in Susan Pharr and Robert Putnam, eds., *Disaffected Democracies: What's Ailing the Trilateral Democracies?*, Princeton: Princeton University Press.

Norris, Pippa (1999). "Introduction: The Growth of Critical Citizens?", en Pippa Norris, ed., *Critical Citizens: Global Support for Democratic Governance*, New York: Oxford University Press.

Norris, Pippa (2011). *Democratic Deficit: Critical Citizens Revisited*, New York: Cambridge University Press.

Price, Vincent and Anca Romatan (2004). "Confidence in Institutions Before, During, and After Indecision 2000", *Journal of Politics*, Vol. 66: 939-956.

Putnam, Robert (2000). *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon & Schuster.

Svolik, Milan (2012). *The Politics of Authoritarian Rule*, New York: Cambridge University Press.

Torcal, Mariano and Ignacio Lago (2006). "Political Participation, Information, and Accountability: Some Consequences of Political Disaffection", in Mariano Torcal and José Ramón Montero, eds., *Political Disaffection in Contemporary Democracies: Social Capital, Institutions, and Politics*, London: Routledge.

Torcal, Mariano, Jordi Muñóz, and Eduard Bonet (2012). "Trust in the European Parliament: From Affective Heuristics to Rational Cueing", in David Sanders, Pedro C. Magalhaes, and Gabor Toka, eds., *Citizens and the European Polity: Mass Attitudes toward the European and National Polities*, Oxford: Oxford University Press.

Zmerli, Sonja, Kenneth Newton, and José Ramón Montero (2007). "Trust in People, Confidence in Political Institutions, and Satisfaction with Democracy", in Jan W. van Deth, Jose Ramon Montero and Anders Westholm, eds., *Citizens and Involvement in European Democracies*, London: Routledge.