



Published in final edited form as:

Qual Life Res. 2016 May ; 25(5): 1131–1136. doi:10.1007/s11136-015-1147-8.

## Language Bias and Self-Rated Health Status among the Latino Population: Evidence of the Influence of Translation in a Wording Experiment

Gabriel R. Sanchez<sup>1,\*</sup> and Edward D. Vargas<sup>2</sup>

<sup>1</sup>Robert Wood Johnson Foundation Center for Health Policy at the University of New Mexico, MSC02 1645, 1 University of New Mexico, 1909 Las Lomas NE, Albuquerque, NM 87131, USA

<sup>2</sup>Center for Women's Health and Health Disparities Research, University of Wisconsin-Madison 310 N. Midvale Blvd, Suite 201; Madison, WI 53705, USA, edward.vargas@gmail.com, + 011 1 (608) 262-7573, fax +011 1 (608) 265-5947

### Abstract

A growing body of research seeks to understand how language bias in survey research impacts our abilities to make generalizations in the study of racial and ethnic disparities. This research uses a wording experiment to assess self-rated health among a representative study of the Latino population (n=1,200). Our analysis shows that by manipulating only the translation of the category fair health into Spanish we are able to directly test the hypothesis that the translation of fair to *regular* in Spanish suppresses Latino self-rated health. We find convincing evidence through the use of logistic and multinomial logistic regressions that respondents provided with the term *regular* report poorer health when compared to those who were given the alternative translation of *mas o menos*. We also find that this translation effect is driven solely by a movement of respondents to choose fair rather than good health, which can in fact explain lower than expected health status rates in studies looking to explore differences between Latinos and non-Latinos. This research informs the study of racial and ethnic disparities, providing a detailed explanation for mixed findings in the Latino health disparities literature.

### Keywords

health disparities; self-reported health; language bias; Latino populations; survey research

### INTRODUCTION

The issue of language translation has concerned survey researchers for some time.<sup>[1, 2, 3]</sup> Although much of this work has been conducted by comparative scholars interested in cross-national analysis among individuals who speak various languages,<sup>[4, 5, 6]</sup> domestic research in the United States has only recently considered the potential bias associated with translation.<sup>[7, 8, 9, 10, 11]</sup> The ramifications associated with language translation bias in survey research has become more salient with recent demographic changes. In fact, the vast

\*Corresponding author: Gabriel Sanchez, sanchezg@unm.edu, + 011 1 (505) 277-0130, fax + 1 (505) 277-0118.

majority of growth in the overall U.S. population between 2000 and 2010 is attributed to increases in the non-White population. The staggering growth of the Hispanic/Latino population is largely responsible for this diversification, as the increase of 15.2 million Latinos from 2000–2010 is more than half of the overall population growth of 27.3 million in the United States during this same time period [12]. The growth rate of the Latino population was surpassed only by that of Asian Americans, who had a growth rate of 43 percent between 2000 and 2010, the largest of any racial or ethnic group. External migration has helped to fuel this population growth, as recent report from the Center for Immigration Studies indicates that the United State’s immigrant population reached a record of nearly 38 million in 2007, with immigrants accounting for one in eight U.S. residents – the highest ratio in the past 80 years [13].

With large foreign-born populations among both Latinos and Asians, non-English language use is highly salient for scholars interested in comparisons between these populations and their English speaking co-ethnics as well as with white English speaking counterparts (who are traditionally used as the reference category to gauge disparities). In fact, according to the US Census, 75 percent of Latinos over 5 years of age indicate that they speak a language other than English at home. This trend in Spanish language use is also being experienced in large survey data collection projects. For example, 61 percent of respondents to the Latino National Survey [14], a major national telephone survey of 8,600 respondents, chose to conduct their survey interview in Spanish.[15]

The implications of survey bias due to language translation have major consequences for health disparities research and for scholars interested in making Latino and non-Latino comparisons. In contrast to research finding Latino health outcomes to be roughly equal, and in some cases better than non-Hispanic whites (Latino Health Paradox in perinatal outcomes), scholars have found Latinos to have poorer health than whites when utilizing general self-rated health status.[16, 17, 18,19] Language is at the center of the discussion of what accounts for this counter-intuitive finding in the literature. For example, research has found that language proficiency is correlated with self-rated health status, with Spanish proficient Latinos reporting lower rates of health than their English dominant counterparts. [20, 21, 22, 23]

Research has suggested that these differences in language could be driven by cultural and linguistic norms about how the response categories within the self-defined health status measure translate to Spanish.[24,25,26,27] Below are the categories of the self-rated health status measure used in the Centers for Disease Control and Prevention–Behavioral Risk Factor Surveillance System (BRFSS), as well as their Spanish language translation in parenthesis: Excellent (*Excelente*), Very Good (*Muy Buena*), Good (*Buena*), Fair (*Regular*), and (Poor (*Mala*). Scholars have suggested that the translation of the category fair to *regular* may denote a more positive meaning in Spanish than it is intended to, thus inflating self-reports of health among Spanish speaking Latinos.

While scholars have yet to directly test this hypothesis, a recent Viruell-Fuentes et al. study [27] does investigate whether language of interview impacts the rate of Latino respondents to choose the fair category across two datasets, the Chicago Community Adult Health Study

[28] and BRFSS. The authors found that conducting the survey interview in Spanish was in fact correlated with an increased likelihood of rating health as fair or poor across both datasets, even when controlling for multiple exogenous factors. The authors suggest that the translation of fair to *regular* leads Spanish language Latino respondents to report poorer health than they would if Latinos conducted the interview in English. While convincing, the authors recommend for research to further explore a more direct test of whether a translational bias exists in Latino self-rated health, preferably with an experimental design among Latino respondents. Taking this suggestion into mind, this manuscript takes advantage of a survey designed with the purpose of exploring whether randomly distributing contrasting translations of the category fair (i.e. *regular* or *mas o menos*) to Spanish speaking respondents does in fact lead to different outcomes.

Our analysis intends to shed light on whether a translational bias exists among Spanish speaking Latino respondents in the translation of fair in the self-rated health measure. Moreover, do Latinos who are assigned the translation of *regular* choose that response category more than those who are assigned *mas o menos*? Lastly, once we control for a vector of variables that impact self-rated health, does being assigned *regular* increase the likelihood of reporting poorer health? This research then provides a direct test of Spanish language translation for our understanding of racial and ethnic health disparities as they relate to Latinos and hopes to improve research by providing insight on the role language bias plays in the study of self-rated health.

## METHODS

### Data Collection

For our analysis, we took advantage of a 2011 Latino Decisions/ImpreMedia survey that was designed in collaboration with the Robert Wood Johnson Foundation Center for Health Policy at the University of New Mexico. Latino Decisions conducted the field work for the survey and worked in conjunction with the Robert Wood Johnson Foundation Center for Health Policy at the University of New Mexico to design a survey instrument focused on health and Latinos. Latino Decisions is a leading survey firm that focuses specifically on capturing the attitudes of Latino respondents, and the RWJF Center for Health Policy is centered on racial and ethnic health disparities research. The sample and design allows us to not only test the relationship between language use and self-rated health, but also allowed us to explore the heterogeneous nature of the Latino experience among a nationally representative sample of Latinos. This is therefore an ideal dataset for our research question, as the research team built-in a split sample approach in the self-rated health status measure where half of the Spanish speaking respondents were randomly given *regular* and the other half were given *mas o menos* in translating the English fair response. Taken together, this is the only nationally representative dataset of Latinos that has a built in language experiment and a host of key independent variables that predict Latino health.

A total of 1,200 Latinos were interviewed over the phone through two samples: 600 Latino registered voters and 600 non-registered Latinos. The non-voter sample was added for the purpose of ensuring that our ability to explore the relationship between language use and health including non-citizens, who are obviously not included in registered voter samples.

All phone calls were administered by Pacific Market Research in Renton, Washington. The survey has an overall margin of error of +/- 4 percent, with an AAPOR response rate of 29 percent. Latino Decisions selected the 21 states with the highest number of Latino registered voters, states that collectively account for over 95 percent of the Latino electorate. Although this sample was designed to capture a large margin of Latino voters, these states also comprise 91 percent of the overall Latino adult population. The voter sample was drawn from registered voters by using the official statewide databases of registered voters, maintained by elections officials in each of the 21 states.

A separate list of Hispanic households was used to identify respondents for the non-voter sample, which was designed to be proportionate to the overall population in those states. Probability sampling methods were employed in both samples based on the respective lists used to identify the universe of potential participants. Respondents were interviewed by telephone, and they could choose to be interviewed in either English or Spanish. A mix of cell phone only and landline households were included in the sample, and both samples are weighted to match the 2010 Current Population Survey<sup>29</sup> universe estimate of Latinos and Latino voters respectively for these 21 states with respect to age, place of birth, gender, and state. The survey was approximately 22 minutes long and was fielded from September 27, 2011 to October 9, 2011.

## Measures

The primary outcome variable of interest is self-rated health status using a single health status question within the Latino Decisions dataset. As a part of their Behavioral Risk Factor Surveillance System (BRFSS), the National Centers for Disease Control and Prevention (CDC) conducts ongoing, state-based surveys of adult health nationwide. The self-rated health status question included in the Latino Decisions survey is very close in wording to the item included in the CDC core BRFSS. The use of general self-rated health long history can be traced back to the 1950's and popularized in a Rand- Medical Outcome Study in what was called the Short-Form 36 and later Short-Form 36 version 2. [<sup>30,31,32</sup>] Both questions utilize a 1 to 5 Likert scale, with respondents rating their health status from excellent to poor. The specific survey question we utilized was "*How would you rate your overall physical health -- excellent, very good, good, fair, or poor?*" which is nearly identical to the CDC BRFSS question of "*Would you say that in general your health is-- excellent, very good, good, fair, or poor?*" The categories of the dependent variable for this study are collapsed into a binary variable for our logistic model and used as a five point nominal scale for our multinomial logistic regressions. For our logistic models, we recoded the original 5-point Likert scale, into a binary indicator of poor or fair health =1 (34.78 percent), and 0= good health, very good health, excellent health (65.22 percent). The coding of our dependent variable for our multinomial logistic model is 1= Poor, 2= Fair, 3= Good, 4= Very Good, and 5= Excellent. Approximately 7.22 percent of respondents answered "Fair"; 27.57 percent answered "Fair; 28.70 percent answered "Good;" 21.22 percent answered "Very Good"; and 15.0 percent answered "Excellent." These distributions indicate sufficient variance and representation of each of the response categories to allow reasonable estimations and tests of the hypotheses offered above. Similar to other work in this area, we are interested in estimating the probability of poor and fair health.[<sup>24,27</sup>]

Our main explanatory variables are three mutually exclusive measures of language of interview; English (n=), Spanish-*regular* (n=), and Spanish-*mas o menos* (n=). Our analysis compares respondents who took the survey in English and respondents who took the Spanish-*mas o menos* survey to our reference category Spanish-*regular*. We also include multiple covariates such as education, age, gender, income, insurance coverage, citizenship, financial stability, and Mexican origin as modeled in prior literature. Summary statistics for all variables used in this analysis are listed in Table 1.

### Statistical Analysis

The research design implemented for this project is based on an experiment embedded in the 2011 Latino Decisions/ImpreMedia survey. To test the hypothesis of whether translation impacts the results generated from a self-rated health status measure we manipulated the Spanish translation of the response category *fair* in questionnaires administered to respondents who elected to take the survey in Spanish. Half of the Spanish language respondents were randomly provided the phrase *mas o menos* as a translation for fair, the alternative suggested by Viruell-Fuentes et al.<sup>24</sup>, and the other half were randomly given the more commonly used *regular* response option. This research design requires a rich sample of Latinos to effectively implement. The 2011 Latino Decisions survey provides an ideal data source, as the survey's sample of Latinos does an excellent job of capturing the diversity within the Latino population. For example, the survey contains a large number of foreign-born (56 percent of sample), non-citizen (19 percent of sample), and uninsured Latinos (20 percent). Most important to the purposes of this study, a robust 50 percent of the survey's respondents conducted the interview in Spanish.

Our analytical approach is intended to determine the relationship between language of survey and health outcomes within a nationally representative sample of Latino adults. Our first analysis focuses on determining the effect of taking the survey in Spanish given the *regular* translation of fair health on reporting poor and fair health compared to respondents who are given *mas o menos* translation and respondents who took the English version. We therefore conduct logistic regression to examine the differences across language categories on the probability of having poor and fair health, controlling for a handful of measures that have been found to be correlated with Latino health status.

Our second analysis examines if respondents who took the survey in Spanish and given *regular* are more likely to rate their health as fair, compared to Spanish speaking respondents who are given the *mas o menos* version and English speaking respondents. We therefore conduct multinomial logistic regression to predict the odds of reporting fair (*regular*, *mas o menos*) health, across language categories, controlling for a handful of measures that have been found to be correlated with Latino health status. The variable self-rated health is a limited dependent variable that is measured using an ordinal survey indicator. Since using ordinary least squares regression to estimate this type of dependent variable can result in biased coefficients and misleading results, the preferred estimation approach is ordered logit model (OLM) or ordered probit model (OPM).<sup>33,34</sup> The OLM and OPM regression are based on the rationale of proportional odds or parallel regression equations. If the ordered models violate this assumption, then a higher order specification such as the multinomial

logit model (MNLM) is preferred. Given our theoretical framework in comparing the fair response to all other response categories, we are less concerned about violating the proportional odds assumption, so MNLM satisfies our analytical specification. In estimating our models, we weighted the data to account for the survey's complex design and all statistical analysis was conducted using Stata 12 software (StataCorp. 2011. *Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.).

Finally, we control for a handful of measures that have been found to be correlated with Latino health status. These control variables include education, age, gender, income, insurance coverage, citizenship, financial stability, and Mexican origin as modeled in prior literature.

## Results

The analysis begins with a discussion of the descriptive statistics for the self-rated health status dependent variable generated with each approach to translating fair into Spanish as well as the control variables used in our analysis. As depicted in the figure 1, the frequencies suggest that differences in translation influences observed Latino health status rates, as more respondents select the fair category when posed with the term *regular* (41 percent) as compared to *mas o menos* (31 percent). Furthermore, the mean for the health status variable is different across both approaches, indicating better overall health status for Latinos when respondents are provided the phrase *mas o menos* (2.946) as compared to *regular* (2.824). This is in line with the *Translation Effect Hypothesis* and in support of those who contend the use of the term *regular* suppresses Latino health status among Spanish speaking respondents. Interestingly, the difference in distribution due to terminology used in translation is limited to the fair and good categories. The descriptive statistics indicate that nearly all of the movement toward poorer health among respondents provided with *regular* comes at the expense of the adjacent category good or *bueno*.

Although the descriptive results provide support for the contention that translation bias might explain lower than expected self-rated health status among Latinos provided in surveys, we need to determine if the use of *regular* leads to suppressed levels of health status when other factors are accounted for. Regarding our overall sample, after dropping missing data (114 observations) we have a total sample of 1,086 respondents. A large segment of our total sample indicated that they conducted in the interview in Spanish (50 percent) and just under 60 percent our sample is female. The mean age in our sample is 52, and the majority of our sample has at least a High School education. Moreover, over half of the sample indicated that they are of Mexican ancestry, both consistent with national data on Latinos from the U.S. Census. In regards to citizenship, over 80 percent of our sample is U.S. citizens. In sum, our sample is representative of U.S. Latinos, as the U.S. Census estimates that about 65 percent of the Latino population is of Mexican origin, 63 percent of Latinos over the age 25 have a High School education, and about 74 percent of Latinos over 5 years of age speak Spanish at home.

Our logistic regression models test the differences across language categories on self-rated health among Latinos, (using Spanish-*Regular* as the reference category) controlling for education, age, gender, income, insurance coverage, citizenship, financial stability, and

Mexican origin. This modeling approach allows for a direct test of whether the translation of fair to *regular* does in fact lead to poorer observed Latino health status than what would be expected in English and for those given the *mas o menos* option.

As reflected in Table 2-model 1, we find that there is a significant relationship between the translation term *mas o menos* and the likelihood of respondents selecting the fair or poor health. Relative to Spanish speaking respondents who were given the term *mas o menos*, Spanish speaking Latinos who were given *regular* do in fact rate their health more poorly. To assess the substantive impact of this translation effect, we conducted post estimation analysis and computed predicted probabilities for values of the translation variable while holding other variables in the model at their means or modes. The probability of reporting either poor or fair health increases from 26 percent when *mas o menos* is used to 34 percent when *regular* is used to translate the fair category for Spanish speaking respondents. Therefore, even when other factors are accounted for, we find that use of the term *regular* does have a suppressing relationship on Latino self-rated health. More intuitively, the odds of reporting poor and fair health are 35 percent lower for Spanish speaking respondents given *mas o menos* relative to Spanish speaking respondents given *regular* as the Spanish translation for fair health, holding all other variables constant ( $p < 0.05$ ). We do not find differences between Spanish speaking respondents given *regular* compared to English speaking respondents in our logistic regression model.

The second component of the analysis is to assess the relative impact of a translation effect across each response category of the self-rated health measure. Here we utilize multinomial logistic regression and set the category of interest, fair, as the base category to assess whether a difference in translation term leads to a significant difference in the likelihood of respondents choosing any specific category relative to choosing fair health. The main results are depicted in Table 3, (see appendix for full model). Consistent with the descriptive results, we find that implementing a different translation for fair health for Spanish speaking respondents only yields a statistically significant difference in the likelihood of respondents choosing good (*bueno*) health. From table 3, we see that respondents provided with *mas o menos* are more likely to rate their health as good relative to the base category of fair. More specifically, respondents provided with *mas o menos* are 28 percent more likely to rate their health as good relative to the base category of fair, compared to respondents provided the *regular* option, holding all else constant ( $p < 0.05$ ). In comparing English respondents, we find that the main differences are at the extremes, in that English respondents are more likely to report poor health and excellent health, relative to fair, when comparing these respondents to Spanish speaking respondents given *regular*, holding all else constant ( $p < 0.05$ ).

To help visualize these relationships and test differences in the response categories. We use post estimation analysis after estimating our multinomial logistic regression and obtain the predicated probabilities of each self-rated response by language category. These relationships are shown in figure 2, were we list self-rated health responses on the x-axis and predicted probabilities on the y-axis. As shown in figure 2, when respondents are provided with the response category *regular* their likelihood of reporting good health is 30 percent as compared to 23 percent when they are given *mas o menos* as the translation for fair health, this difference is statistically significant at the 0.05 level. From figure 2, we can conclude

that this 7 percent difference is directly attributable to using the *regular* translation as opposed to *mas o menos* wording, as we control for age, gender, education, income, citizenship, insurance coverage, Mexican origin and financial stability (all of which were set to their mean or mode values). We also find that respondents given *regular* are statistically different than English speaking Latinos, within this fair health response category. Among other response categories, we find that there are differences in the good health response as well, as Spanish speaking respondents who are given the *mas o menos* survey are more likely than English respondents to report good health ( $p < .05$ ). Lastly, we find that English speaking respondents are more likely to report poor health compared to all non-English speaking respondents ( $p < 0.05$ ).

## CONCLUSION

Scholars have struggled to explain outcomes of studies utilizing self-rated health status for Latinos, as these measures have consistently produced lower rates of health for Latinos relative to non-Latino whites. This is surprising to many given that research using other measures of health status consistently suggest that Latinos have health outcomes that are on par, or in some cases, better than non-Hispanic whites. Among other potential explanations, language translation has been offered as a theory to explain this apparent contradiction. We have attempted to assess the impact of translation bias on Latino self-rated health in this analysis by implementing an experiment within a survey of Latinos which includes a large sample of respondents who conducted the survey in Spanish. By manipulating only the translation of the category “fair” health into Spanish we are able to directly test the theory that the translation of fair to *regular* in Spanish suppresses Latino self-rated health.

We find convincing evidence that respondents provided with the term *regular* report poorer health when compared to those who were given the alternative translation of *mas o menos*. Furthermore, we find that this translation effect is driven solely by a movement of respondents to choose fair rather than good health, which can in fact explain lower than expected health status rates in studies looking to explore differences between Latinos and non-Latinos. Given the large percentage of Latinos who prefer to conduct surveys in Spanish, the implications of this finding are significant. We believe that these findings should motivate scholars to interpret results generated with the self-rated health measure with some caution, particularly when making comparisons between Latinos and non-Latinos. We hope that this study will also lead to additional attention to the consequences of language translation in survey research, particularly within the context of health disparities research.

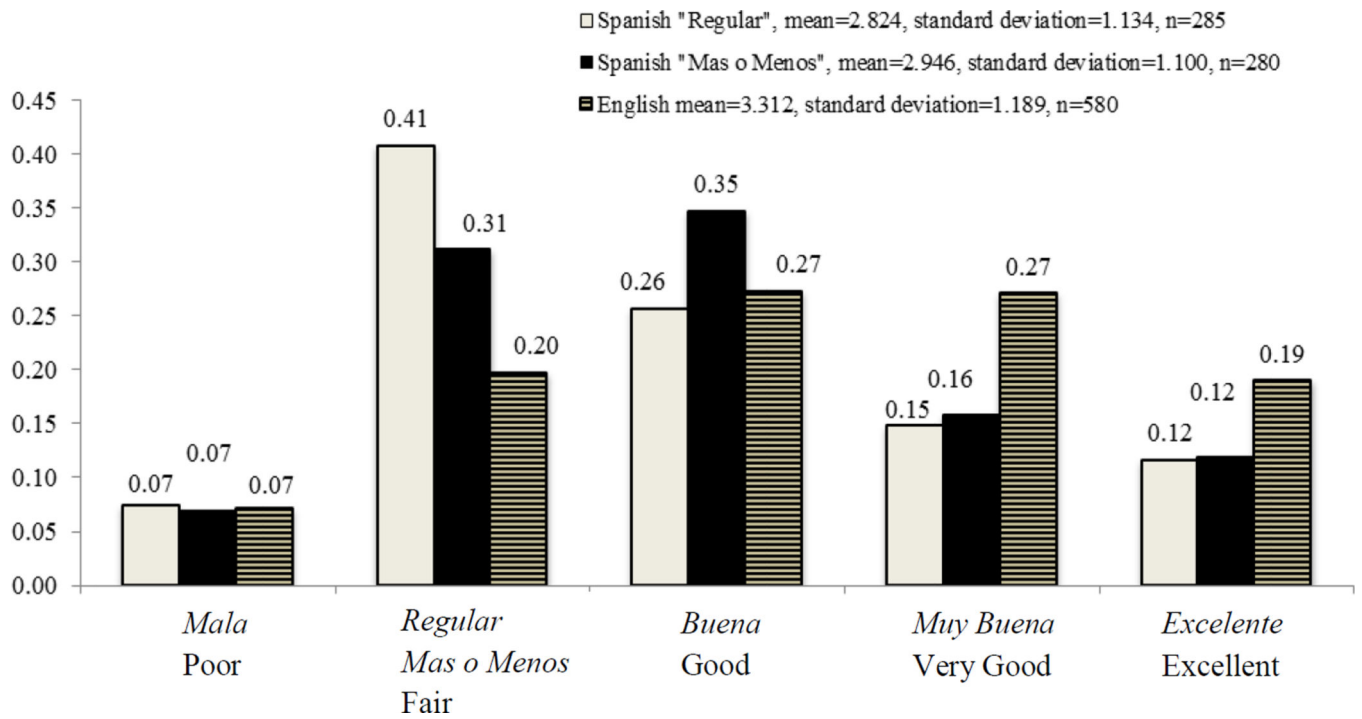
## Bibliography

1. Hopkins DJ. Translating into Votes: The Electoral Impact of Spanish-Language Ballots. *American Journal of Political Science*. 2011; 55(4):813–829.
2. Reilly S, Richey S. Ballot Question Readability and Roll-Off: The Impact of Language Complexity. *Political Research Quarterly*. 2011; 64:59–67.
3. Tucker JT, Espino R. Government Effectiveness and Efficiency? The Minority Language Assistance Provisions of the VRA. *Texas Journal on Civil Liberties and Civil Rights*. 2007; 12(2):164–231.

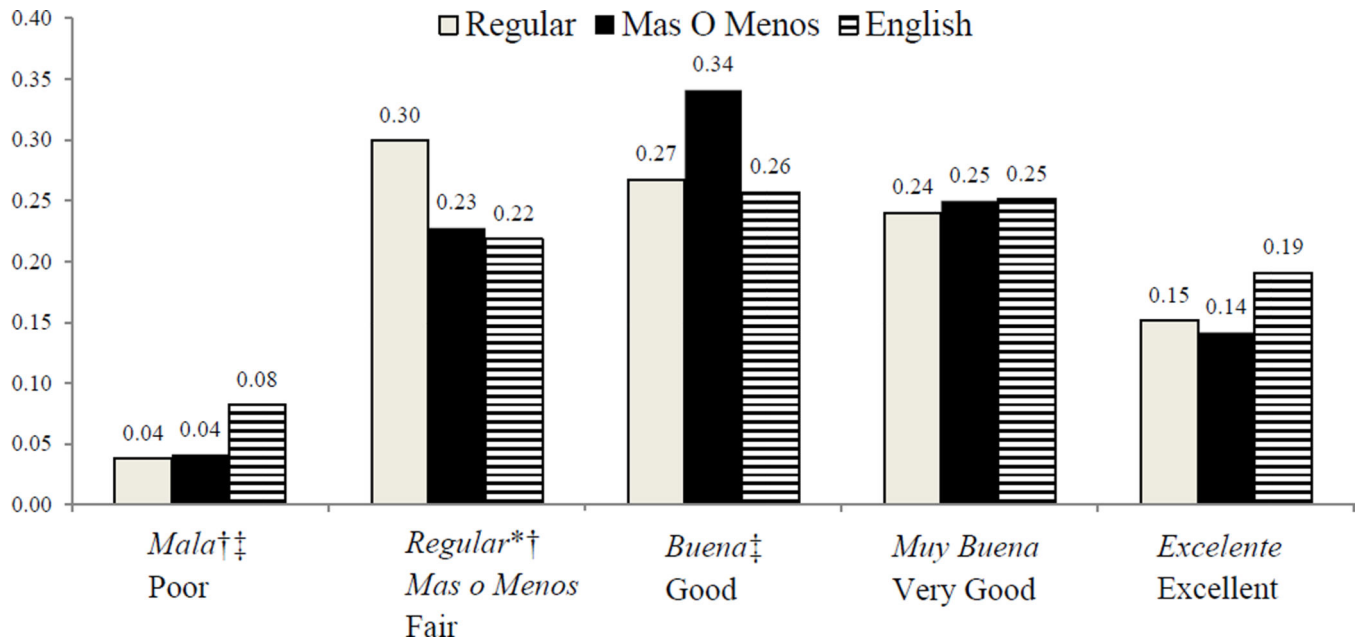


4. Liu, Amy H.; Jacob, I Ricks. Coalitions and Language Politics: Policy Shifts in Southeast Asia. *World Politics*. 2012; 64(03):476–506.
5. Patten, Alan. *Political Theory and Language Policy*. *Political Theory*. 2001; 29(5):691–715.
6. Schmid, Carol L. *The Politics of Language: Conflict, Identity, and Cultural Pluralism in Comparative Perspective*. Cary: Oxford University Press; 2001.
7. Kirkman-Liff B, Mondragón D. Language of interview: relevance for research of southwest Hispanics. *American Journal of Public Health*. 1991; 81(11):1399–1404. [PubMed: 1951794]
8. Marin, G.; Marin, B. *Research with Hispanic Populations*. Newbury Park, CA: Sage Publications; 1991.
9. Ponce, Ninez A.; Lavarreda, SA.; Yen, W.; Brown, ER.; DiSogra, C.; Satter, DE. The California Health Interview Survey 2001: Translation of a Major Survey for California's Multiethnic Population. *Association of Schools of Public Health*. 2004 Jul-Aug;(119)
10. Perez, Efrén O. The Origins and Implications of Language Effects in Multilingual Surveys: A MIMIC Approach with Application to Latino Political Attitudes. *Political Analysis*. 2011; 19(4): 434–454.
11. Wu, Shinyi M.; Susan, Ridgely; Jose, J Escarce; Leo, S Morales. Language access services for Latinos with limited English proficiency: Lessons learned from *Hablamos Juntos*. *Journal of General Internal Medicine*. 2007; 22(2):350–355. [PubMed: 17957424]
12. Humes, KR.; Jones, NA.; Roberto, R. *Overview of Race and Hispanic Origin*: 2010. Washington, DC: U.S. Department of Commerce; 2011.
13. Camarota SA. *Immigrants in the United States, 2007: A Profile of America's Foreign-Born Population*. Center for Immigration Studies. 2007
14. Fraga, LR.; Garcia, JA.; Hero, R.; Jones-Correa, M.; Martinez-Ebers, V.; Segura, GM. *Latino National Survey (LNS)*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research; 2006.
15. See LNS topline, available at: [http://depts.washington.edu/uwiser/documents/LNS\\_toplines\\_FIP\\_April16\\_2007.pdf](http://depts.washington.edu/uwiser/documents/LNS_toplines_FIP_April16_2007.pdf)
16. Karl Finch B, Vega WA. Acculturation Stress, Social Support, and Self-Rated Health among Latinos in California. *Journal of Immigrant Health*. 2003; 5(3):109–117. [PubMed: 14512765]
17. Karl Finch B, Hummer RA, Reindi M, Vega WA. Validity of Self-rated Health among Latino(a)s. *American Journal of Epidemiology*. 2002; 155(8):755–759. [PubMed: 11943694]
18. McGree DL, Y L, G C, et al. Self-reported health status and mortality in a multiethnic US cohort. *American Journal of Epidemiology*. 1999; 149:41–46. [PubMed: 9883792]
19. Shetterly SM, Baxter J, Mason LD, et al. Self-rated Health among Hispanics vs. non-Hispanic White Adults: The San Luis Valley Health and Aging Study. *American Journal of Public Health*. 1996; 86:1798–1801. [PubMed: 9003141]
20. Berkanovic E. The effect of inadequate language translation on Hispanics' responses to health surveys. *American Journal of Public Health*. 1980; 70(12):1273–1276. [PubMed: 7435745]
21. Carrasquillo O, Orav EJ, Brennan TA, Burstin HR. Impact of Language Barriers on Patient Satisfaction in an Emergency Department. *Journal of General Internal Medicine*. 1999; 14(2):82–87. [PubMed: 10051778]
22. Hendricson WD, Russell J, Prihoda TJ, Jacobson JM, Rogan A, Bishop GD. An Approach to Developing a Valid Spanish Language Translation of a Health-Status Questionnaire. *Medical Care*. 1989; 27(10):959–966. [PubMed: 2796414]
23. Morales LS, Cunningham WE, Brown JA, Liu H, Hays RD. Are Latinos Less Satisfied with Communication by Health Care Providers? *Journal of General Internal Medicine*. 1999; 14(7): 409–417. [PubMed: 10417598]
24. Shetterly SM, Baxter J, Mason LD, Hamman RF. Self-rated health among Hispanic vs non-Hispanic white adults: the San Luis Valley Health and Aging Study. *American Journal of Public Health*. 1996 Dec; 86(12):1798–1801. [PubMed: 9003141]
25. Bzostek S, Goldman N, Pebley A. Why do Hispanics in the USA Report Poor Health? *Social Science Medicine*. 2007; 65(5):990–1003. [PubMed: 17574713]

26. Kandula NR, Lauderdale DS, Baker DW. Differences in Self-Rated Health among Asians, Latinos, and non-Hispanic Whites: the Role of Language and Nativity. *Annual Epidemiology*. 2007; 17(3): 191–198.
27. Viruell-Fuentes, et al. Language of Interview, Self-Rated Health, and the Other Latino Health Puzzle. *Am J Public Health*. 2011; 101(7):1306–1313. [PubMed: 21164101]
28. Bureau of the Census. *Current Population Survey*. Washington, DC: U.S. Department of Commerce; 2010.
29. Ware, JE.; Davies-Avery, A.; Donald, C. *General Health Perceptions*. Santa Monica, CA: Rand; 1978. *Conceptualisation and Measurement of Health for Adults in the Health Insurance Study Vol. V*.
30. Brook, RH.; Ware, JE.; Davies-Avery, A. *Overview*. Santa Monica, CA: Rand; 1979. *Conceptualization and Measurement of Health for Adults in the Health Insurance Study Vol. VIII*.
31. Stewart, AL.; Ware, JE., editors. *Measuring Functioning and Well-Being. The Medical Outcomes Study Approach*. Durham: Duke University Press; 1992.
32. House, JS.; Kaplan, GA.; Morenoff, J.; Raudenbush, SW.; Williams, DR.; Young, EA. *Chicago Community Adult Health Study, 2001–2003*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research; 2011.
33. Long, John S. A volume in the Sage Series for Advanced Quantitative Techniques. Thousand Oaks, CA: Sage Publications; 1997. *Regression Models for Categorical and Limited Dependent Variables*.
34. Long, John S.; Jeremy, Freese. *Regression Models for Categorical and Limited Dependent Variables with Stata*. Second. College Station, TX: Stata Press; 2005.



**Figure 1.**  
Summary Statistics for the Distribution of Self-Rated Health



**Figure 2. Adjusted Predicted Probabilities of Multinomial Logistic Regression Model of Self-Rated Health for Latinos by Language of Interview: Latino Decisions/ImpreMedia Survey 2011 (n=1,086)**

*Note.* Controlling for age, gender, education, income, citizenship, insurance coverage, Mexican origin and financial stability (all of which were set to their mean or mode values).

\* $P < 0.05$  for the difference between Spanish-“Regular” and Spanish-“Mas o Menos-” language interviewees in the same response category, †  $P < 0.05$  for the difference between English- and Spanish-“Regular” language interviewees in the same response category, ‡  $P < 0.05$  for the difference between English- and Spanish-“Mas O Menos” language interviewees in the same response category.

**Table 1**

Summary Statistics using a 2011 Latino Decisions/ImpreMedia Survey (n=1,105).

Variable	Mean	Std. Dev.	Min	Max
Poor Health <sup>1</sup>	0.35	0.48	0	1
Health Status <sup>2</sup>	3.10	1.17	1	5
Spanish <i>Regular</i> <sup>3</sup>	0.25	0.43	0	1
Spanish <i>Mas o Menos</i> <sup>4</sup>	0.25	0.43	0	1
English <sup>5</sup>	0.50	0.50	0	1
Education <sup>6</sup>	3.47	1.55	1	6
Income less than 39k	0.19	0.39	0	1
Income: Missing	0.49	0.50	0	1
Income: 40k–60k	0.13	0.34	0	1
Income: 60k–80k	0.07	0.26	0	1
Income: 80k>	0.12	0.33	0	1
Uninsured	0.20	0.40	0	1
Female	0.59	0.49	0	1
Financial Stability	0.39	0.49	0	1
Age	51.62	17.18	18	98
U.S. Citizen <sup>7</sup>	0.81	0.39	0	1
Mexican Origin <sup>8</sup>	0.53	0.50	0	1

<sup>1</sup>Poor Health Status is coded 0= good health, very good health, excellent health and 1= poor health, fair health

<sup>2</sup>Self-Rated Health is coded 1= poor health, 2=fair health, 3=good health, 4= very good health, 5=excellent health

<sup>3</sup>Spanish Language *Regular* 0=English, Spanish *mas o menos*, 1=Spanish *regular*

<sup>4</sup>Spanish Language *Mas o Menos* 0=English, Spanish *regular*, 1= Spanish *mas o menos*

<sup>5</sup>English Language:, Spanish *regular*, 0= Spanish *regular*, Spanish *mas o menos*, 1=English

<sup>6</sup>Highest education levels completed, (1=Grade 1–8, 2=Some HS, 3=HS, 4=Some College, 5=College Grad, 6=Post-Grad)

<sup>7</sup>Mexican Ancestry: 0=Non Mexican, 1=Mexican

<sup>8</sup>U.S. Citizen: 0=non U.S. citizen, 1=U.S. citizen by birthright, naturalization, born in Puerto Rico

**Table 2**

Logistic Coefficients for Regression of Language Use on Poor or Fair Health using a 2011 Latino Decisions/ ImpreMedia Survey.

VARIABLES	Model 1	
	$\beta$	Odds Ratios
Reference: Spanish <i>Regular</i>		
Spanish <i>Mas o Menos</i>	-0.429 **	0.651 **
English	-0.233	0.793
Education <sup>1</sup>	-0.246 ***	0.782 ***
Income Reference: less than 39k		
Income: Missing	-0.248	0.78
Income: 40k–60k	-0.299	0.741
Income: 60k–80k	-0.069	0.933
Income: 80k>	-0.860 ***	0.423 ***
Uninsured	0.441 **	1.554 **
Female	0.169	1.184
Financial Stability	0.535 ***	1.708 ***
Age	0.023 ***	1.024 ***
U.S. Citizen <sup>2</sup>	-0.473 **	0.623 **
Mexican Origin <sup>3</sup>	-0.054	0.947
Constant	-0.696 *	0.499 *
Number of Observations	1,086	
Pseudo R2	0.115	

Notes:

\*\*\*  
p<0.01,

\*\*  
p<0.05,

\*  
p<0.1,

$\beta$  is a logit coefficient

<sup>1</sup>Highest education levels completed, (1=Grade 1–8, 2=Some HS, 3=HS, 4=Some College, 5=College Grad, 6=Post-Grad)

<sup>2</sup>U.S. Citizen: 0=non U.S. citizen, 1=U.S. citizen by birthright, naturalization, born in Puerto Rico

<sup>3</sup>Mexican Ancestry: 0=Non Mexican, 1=Mexican

**Table 3**

Full Multinomial Logit Regression Coefficients, Dependent Variable = Self-Rated Health, Comparison Response Category= “Fair health”, Reference Category=Spanish Respondents Given *Regular*. Latino Decisions/ImpreMedia Survey 2011 (n=1,086).

VARIABLES	Poor		Good		Very Good		Excellent	
	$\beta$	Odds Ratio	B	Odds Ratio	$\beta$	Odds Ratio	$\beta$	Odds Ratio
Reference: Spanish <i>Regular</i>								
Spanish <i>Más o Menos</i>	0.236 (0.430)	1.266	0.541** (0.233)	1.717**	0.355 (0.281)	1.426	0.242 (0.325)	1.274
English	1.008** (0.413)	2.740**	0.295 (0.239)	1.343	0.405 (0.263)	1.499	0.592** (0.290)	1.808**
Education <i>I</i>	0.042 (0.112)	1.043	0.201*** (0.067)	1.223***	0.402*** (0.075)	1.494***	0.195** (0.082)	1.220**
Income Reference: less than 39k								
Income: Missing	-0.660 (0.499)	0.517	0.071 (0.256)	1.073	0.545* (0.278)	1.725*	-0.359 (0.329)	0.699
Income: 40k-60k	-0.826 (0.545)	0.438	0.162 (0.267)	1.176	0.303 (0.293)	1.353	0.031 (0.309)	1.032
Income: 60k-80k	0.463 (0.515)	1.588	0.127 (0.379)	1.136	0.632* (0.380)	1.881*	-0.159 (0.422)	0.853
Income: 80k>	-0.638 (0.701)	0.528	0.198 (0.367)	1.219	1.219*** (0.350)	3.385***	0.597 (0.370)	1.817
Uninsured	0.163 (0.343)	1.178	-0.201 (0.210)	0.818	-0.392* (0.238)	0.676*	-0.946*** (0.280)	0.388***
Female	-0.121 (0.296)	0.886	0.035 (0.179)	1.035	-0.219 (0.194)	0.803	-0.607*** (0.212)	0.545***
Financial Stability	0.369 (0.294)	1.446	-0.251 (0.178)	0.778	-0.572*** (0.199)	0.565***	-0.823*** (0.222)	0.439***
Age	0.011 (0.009)	1.011	-0.012** (0.006)	0.988**	-0.026*** (0.006)	0.974***	-0.033*** (0.007)	0.967***
U.S. Citizen <sup>2</sup>	0.361 (0.416)	1.435	0.454** (0.230)	1.574**	0.378 (0.262)	1.459	0.978** (0.305)	2.659***
Mexican Origin <sup>3</sup>	0.492 (0.306)	1.635	0.284 (0.184)	1.329	0.104 (0.196)	1.110	-0.064 (0.212)	0.938

VARIABLES	Poor		Good		Very Good		Excellent	
	$\beta$	Odds Ratio	B	Odds Ratio	$\beta$	Odds Ratio	$\beta$	Odds Ratio
Constant	-3.264 (0.789)	0.038***	-0.656 (0.453)	0.519	-0.821* (0.494)	0.440*	0.129 (0.537)	1.138
Pseudo R2	0.0887							

Notes:

\*\*\* p<0.01,

\*\* p<0.05,

\* p<0.1,

$\beta$  is a logit coefficient

<sup>1</sup> Highest education levels completed, (1=Grade 1-8, 2=Some HS, 3=HS, 4=Some College, 5=College Grad, 6=Post-Grad)

<sup>2</sup> U.S. Citizen: 0=non U.S. citizen, 1=U.S. citizen by birthright, naturalization, born in Puerto Rico

<sup>3</sup> Mexican Ancestry: 0=Non Mexican, 1=Mexican