Qual Life Res
DOI 10.1007/s11136-015-1147-8

BRIEF COMMUNICATION

Language bias and self-rated health status among the Latino population: evidence of the influence of translation in a wording experiment

Gabriel R. Sanchez[1] · **Edward D. Vargas**[2]

Accepted: 26 September 2015

Abstract
Purpose This research uses a translation experiment to assess the Spanish translation of the "fair" response in the self-rated health measure among a representative study of the Latino population in the USA.
Methods Using a unique Latino-specific survey ($n = 1200$), researchers built in a split sample approach in the self-rated health status measure where half of the Spanish-speaking respondents ($n = 600$) were randomly given "*regular*" and the other half were given "*Mas o Menos*" in translating the English "fair" response. We first estimate a logistic regression model to estimate differences across language categories on the probability of reporting poor and fair health and then estimate a multinomial logistic regression to test whether respondents who took the survey in Spanish and given "*regular*" are more likely to rate their health as fair compared to English speakers and Spanish-speaking respondents who are given the "*Mas o Menos*" version.
Results From our logistic regression model, we find that Spanish-speaking respondents given the "*regular*" response are more likely to report poor health relative to English-speaking respondents and Spanish-speaking respondents who were randomly given "*Mas o Menos*." The results from our multinomial logistic models suggest that Spanish respondents provided with "*Mas o Menos*" are more likely to rate their health as good relative to the base category of fair and relative to both English and Spanish speakers given "*regular.*"
Conclusion This research informs the study of racial and ethnic disparities by providing a detailed explanation for mixed findings in the Latino health disparities literature. Researchers interested in self-rated health should translate the general self-rated health option "fair" to "*Mas o Menos*" as our wording experiment suggests that the current wording "*regular*" overinflates the reporting of poor health.

Keywords Health disparities · Self-reported health · Language bias · Latino populations · Survey research

Introduction

Survey bias due to language translation has major implications for health disparities research and for scholars interested in making generalizations across racial and ethnic populations. In contrast to research finding Latino health outcomes to be roughly equal, and in some cases better than non-Hispanic whites (Latino health paradox), scholars have found that Latinos report poorer health than whites when utilizing a self-rated health status [1–4]. Language is at the center of the discussion of what accounts for this counterintuitive finding [5–8].

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 are translated into Spanish [9–12].

✉ Gabriel R. Sanchez
sanchezg@unm.edu

Edward D. Vargas
edward.vargas@gmail.com

[1] Robert Wood Johnson Foundation Center for Health Policy, University of New Mexico, MSC02 1645, 1 University of New Mexico, 1909 Las Lomas NE, Albuquerque, NM 87131, USA

[2] Center for Women's Health and Health Disparities Research, University of Wisconsin-Madison, 310 N. Midvale Blvd, Suite 201, Madison, WI 53705, USA

Published online: 06 October 2015

Below are the categories of the self-rated health status measure typically used in the majority of large national and international surveys, as well as their Spanish language translations in parentheses: 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 study by Viruell-Fuentes et al. [12] finds that conducting the survey interview in Spanish was correlated with an increased likelihood of rating health as "fair" or "poor" across national datasets, even when controlling for multiple exogenous factors. The authors suggest that the translation of "fair" to *regular* (the most common translation approach) leads Spanish language Latino respondents to report poorer health than Latinos respondents who completed the interview in English. Using this theoretical framework, our analysis sheds light on whether or not a translational bias exists among US Spanish-speaking Latino respondents in the translation of "fair" in the self-rated health measure. This research provides a direct test of the effect of Spanish language translations on self-rated health to inform our understanding of racial and ethnic health disparities as they relate to Latinos, improving research by providing insight into the role language bias plays in the study of self-rated health.

Methods

Data collection

We take advantage of a nationally representative 2011 Latino Decisions/ImpreMedia telephone survey ($n = 1200$) that was designed in collaboration with the Robert Wood Johnson Foundation Center for Health Policy at the University of New Mexico. The sample and design allowed us to test Spanish language translation bias on self-rated health controlling for the heterogeneous nature of the Latino experience among a nationally representative sample of Latinos. This is therefore an ideal dataset, as the research team built in a split sample approach in the self-rated health status measure: half of the Spanish-speaking respondents ($n = 600$) were randomly presented with *regular* and the other half were presented with *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. The survey was approximately 22 min long and was fielded from September 27 to October 9 of 2011. The overall margin of error was ±4 %, with an AAPOR response rate of 29 %.

Measures

The primary outcome variable of interest is self-rated health status that was collapsed into a binary variable for our logistic models 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 indictor of "poor" or "fair" health = 1 (34.78 %) and 0 = all else (65.22 %). The distribution of our dependent variable for our multinomial logistic model is 7.22 % of respondents answered "poor," 27.57 % "fair," 28.70 % "good," 21.22 % "very good," and 15.0 % "excellent."

Our main explanatory variables are three mutually exclusive measures of language of interview: English, Spanish-*regular*, and Spanish-*Mas o Menos*. 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*.

Statistical analysis

Our first analysis focuses on determining the effect of taking the survey in Spanish and being given the *regular* translation of "fair" on reporting "poor" or "fair" health compared to respondents given the *Mas o Menos* translation and to respondents who responded to the survey in English. We therefore conduct a logistic regression to examine the differences across language categories on the probability of having "poor" or "fair" health.

Our second analysis examines whether respondents who took the survey in Spanish and who were given the translation *regular* are more likely to rate their health as "fair" compared to Spanish-speaking respondents who are given the *Mas o Menos* version and to English-speaking respondents. We therefore conduct a multinomial logistic regression to predict the odds of reporting "fair" (*regular*, *Mas o Menos*) health across language categories. Finally, we control for education, age, gender, income, insurance coverage, citizenship, financial stability, and Mexican origin, all of which have been found to be correlated with Latino health status.

Results

Table 1 shows the resulting statistics from the sample of 1086. One hundred and fourteen observations were dropped as a result of missing data. The results from our logistic models (Table 2) suggest that there is a significant

Table 1 Summary statistics using a 2011 Latino Decisions/ImpreMedia Survey ($n = 1105$)

Variable	Mean	SD	Min	Max
Poor health[a]	0.35	0.48	0	1
Health status[b]	3.10	1.17	1	5
Spanish-*regular*[c]	0.25	0.43	0	1
Spanish-*Mas o Menos*[d]	0.25	0.43	0	1
English[e]	0.50	0.50	0	1
Education[f]	3.47	1.55	1	6
Income <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
US citizen[g]	0.81	0.39	0	1
Mexican origin[h]	0.53	0.50	0	1

[a] Poor health status is coded 0 = good health, very good health, excellent health and 1 = poor health, fair health

[b] Self-rated health is coded 1 = poor health, 2 = fair health, 3 = good health, 4 = very good health, 5 = excellent health

[c] Spanish language *regular*: 0 = English, Spanish-*Mas o Menos*, 1 = Spanish-*regular*

[d] Spanish language *Mas o Menos*: 0 = English, Spanish-*regular*, 1 = Spanish-*Mas o Menos*

[e] English language: Spanish-*regular*, 0 = Spanish-*regular*, Spanish-*Mas o Menos*, 1 = English

[f] Highest education levels completed (1 = grade 1–8, 2 = some HS, 3 = HS, 4 = some college, 5 = college grad, 6 = post-grad)

[g] Mexican ancestry: 0 = non-Mexican, 1 = Mexican

[h] US citizen: 0 = non-US citizen, 1 = US citizen by birthright, naturalization, born in Puerto Rico

relationship between the translation term *Mas o Menos* and the likelihood of respondents selecting the "fair" or "poor" health options. 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 findings show that the probability of reporting either "poor" or "fair" health increases from 26 % when *Mas o Menos* is used to 34 % 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 effect on Latino self-rated health.

Table 2 Logistic coefficients for regression of language use on "poor" or "fair" health using a 2011 Latino Decisions/ImpreMedia Survey

Variables	Model 1	
	β	Odds ratios
Reference: Spanish-*regular*		
Spanish-*Mas o Menos*	−0.429**	0.651**
English	−0.233	0.793
Education[a]	−0.246***	0.782***
Income reference: <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***
US citizen[b]	−0.473**	0.623**
Mexican origin[c]	−0.054	0.947
Constant	−0.696*	0.499*
Number of observations	1086	
Pseudo R^2	0.115	

We also acknowledge the work by experimental survey researchers who argue that controlling for covariates in experiments may undermine the basis for unbiased inference based on the experimental design [16]. Given this concern, we also compare Spanish-*Mas o Menos* versus Spanish-*regular* without controlling for additional covariates and our findings remain consistent. For our binary "poor or fair" response (Spanish-*Mas o Menos* compared to Spanish-*regular*), the logit β coefficient is 0.291 ($p < 0.01$) with an odds ratio of 1.337 with a standard error of 0.176

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$, β is a logit coefficient

[a] Highest education levels completed (1 = grade 1–8, 2 = some HS, 3 = HS, 4 = some college, 5 = college grad, 6 = post-grad)

[b] US citizen: 0 = non-US citizen, 1 = US citizen by birthright, naturalization, born in Puerto Rico

[c] Mexican ancestry: 0 = non-Mexican, 1 = Mexican

Our second analysis assesses the relative impact of a translation effect across each response category of the self-rated health measure. The findings of our multinomial logistic regression that sets the category of interest, "fair," as the base category show 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. We find that respondents provided with *Mas o Menos* are 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 (Table 3).

To help visualize these relationships, we obtained the predicated probabilities of each self-rated response by

Table 3 Full multinomial logit regression coefficients, dependent variable = self-rated health, comparison response category = "fair," reference category = Spanish respondents given *regular*, Latino Decisions/ImpreMedia Survey 2011 ($n = 1086$)

Variables	Poor		Good		Very good		Excellent	
	β	Odds ratio	β	Odds ratio	β	Odds ratio	β	Odds ratio
Reference: Spanish-*regular*								
Spanish-*Mas o Menos*	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[a]	0.042 (0.112)	1.043	0.201*** (0.067)	1.223***	0.402*** (0.075)	1.494***	0.199** (0.082)	1.220**
Income reference: <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***
US citizen[b]	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[c]	0.492 (0.306)	1.635	0.284 (0.184)	1.329	0.104 (0.196)	1.110	−0.064 (0.212)	0.938
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 R^2	0.0887							

We also acknowledge the work by experimental survey researchers who argue that controlling for covariates in experiments may undermine the basis for unbiased inference based on the experimental design [16]. Given this concern, we also compare Spanish-*Mas o Menos* versus Spanish-*regular* without controlling for additional covariates, and our findings remain consistent. For our "good" response (Spanish-*Mas o Menos* compared to Spanish-*regular*), the logit β coefficient is 0.429 ($p < 0.05$) with an odds ratio of 0.651 with a standard error of 0.215

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$, β is a logit coefficient, standard errors in parentheses

[a] Highest education levels completed (1 = grade 1–8, 2 = some HS, 3 = HS, 4 = some college, 5 = college grad, 6 = post-grad)

[b] US citizen: 0 = non-US citizen, 1 = US citizen by birthright, naturalization, born in Puerto Rico

[c] Mexican ancestry: 0 = non-Mexican, 1 = Mexican

language category. These relationships are shown in Fig. 2, and we find that for respondents who are provided with the response category *regular,* their likelihood of reporting good health is 30 % as compared to 23 % when they are given *Mas o Menos* as the translation for fair health ($p = 0.031$; Fig. 1).

Conclusion

Scholars have struggled to explain outcomes of studies using 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 suggests that Latinos have health outcomes that are on a par with, 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 completed the survey in Spanish. By manipulating only the translation of the category "fair" health into Spanish, we find convincing evidence that

Fig. 1 Summary statistics for the distribution of self-rated health

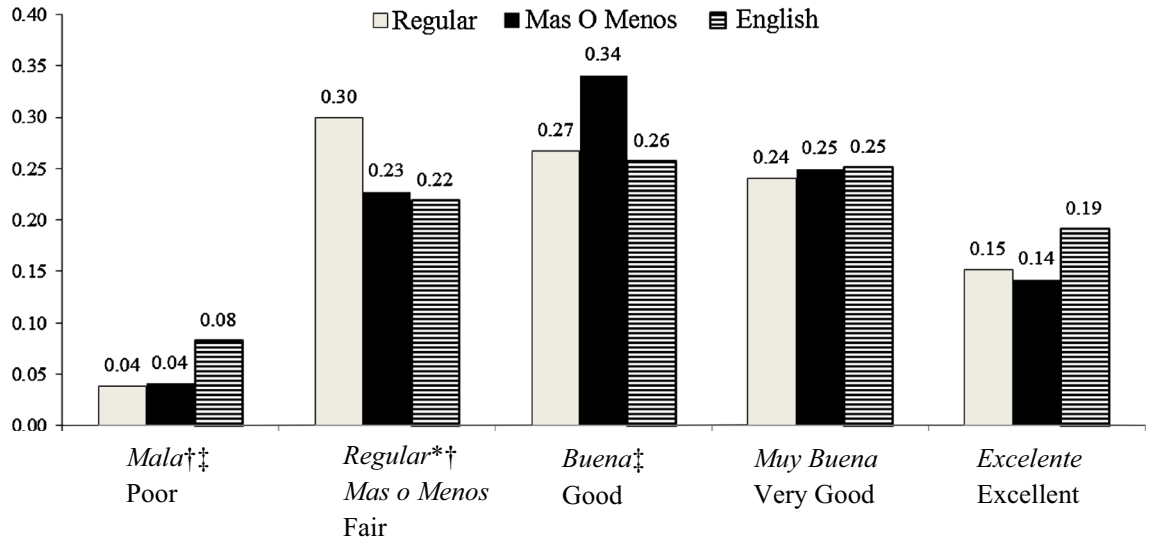

Fig. 2 Adjusted predicted probabilities of multinomial logistic regression model of self-rated health for Latinos by language of interview: Latino Decisions/ImpreMedia Survey 2011 ($n = 1086$). *Note*: Controlling for age, gender, education, income, citizenship, insurance coverage, Mexican origin, and 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, $^{\dagger}p < 0.05$ for the difference between English and Spanish-*regular* language interviewees in the same response category, $^{\ddagger}p < 0.05$ for the difference between English and Spanish-*Mas o Menos* language interviewees in the same response category

respondents provided with the term *regular* report poorer health when compared to those who were given the alternative translation of *Mas o Menos*, following a recommendation for future research suggested by Viruell-Fuentes et al. [12]. 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. Future research should examine "*Mas o*

Menos" versus "*regular*" using forward-and-back-translation methods that take into consideration the rich literature in cultural equivalency [13–15]. While we could not achieve that here, we believe our findings should motivate scholars to implement such experiments.

Acknowledgments Edward D. Vargas was supported by the NICHD Health Disparities Research Scholars Training Program Grant Number 5T32HD049302-09.

Compliance with ethical standards

Conflict of interest The authors report no conflict of interests.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Finch, B. K., & Vega, W. A. (2003). Acculturation stress, social support, and self-rated health among Latinos in California. *Journal of Immigrant Health, 5*(3), 109–117.
2. Finch, B. K., Hummer, R. A., Reindi, M., & Vega, W. A. (2002). Validity of self-rated health among Latino(a)s. *American Journal of Epidemiology, 155*(8), 755–759.
3. McGree, D. L., Liao, Y., Cao, G., et al. (1999). Self-reported health status and mortality in a multiethnic US cohort. *American Journal of Epidemiology, 149*, 41–46.
4. Shetterly, S. M., Baxter, J., Mason, L. D., et al. (1996). Self-rated health among Hispanics vs. non-Hispanic White adults: The San Luis Valley Health and Aging Study. *American Journal of Public Health, 86*, 1798–1801.
5. Berkanovic, E. (1980). The effect of inadequate language translation on Hispanics' responses to health surveys. *American Journal of Public Health, 70*(12), 1273–1276.
6. Carrasquillo, O., Orav, E. J., Brennan, T. A., & Burstin, H. R. (1999). Impact of language barriers on patient satisfaction in an emergency department. *Journal of General Internal Medicine, 14*(2), 82–87.
7. Hendricson, W. D., Russell, J., Prihoda, T. J., Jacobson, J. M., Rogan, A., & Bishop, G. D. (1989). An approach to developing a valid Spanish language translation of a health-status questionnaire. *Medical Care, 27*(10), 959–966.
8. Morales, L. S., Cunningham, W. E., Brown, J. A., Liu, H., & Hays, R. D. (1999). Are Latinos less satisfied with communication by health care providers? *Journal of General Internal Medicine, 14*(7), 409–417.
9. Shetterly, S. M., Baxter, J., Mason, L. D., & Hamman, R. F. (1996). Self-rated health among Hispanic vs non-Hispanic white adults: The San Luis Valley Health and Aging Study. *American Journal of Public Health, 86*(12), 1798–1801.
10. Bzostek, S., Goldman, N., & Pebley, A. (2007). Why do Hispanics in the USA report poor health? *Social Science Medicine, 65*(5), 990–1003.
11. Kandula, N. R., Lauderdale, D. S., & Baker, D. W. (2007). Differences in self-rated health among Asians, Latinos, and non-Hispanic Whites: the role of language and nativity. *Annual Epidemiology., 17*(3), 191–198.
12. Viruell-Fuentes, et al. (2011). Language of interview, self-rated health, and the other Latino health puzzle. *American Journal of Public Health, 101*(7), 1306–1313.
13. Herdman, M., Fox-Rushby, J., & Badia, X. (1997). 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research,6*(3), 237–247.
14. Bullinger, M., et al. (1993). Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Quality of Life Research, 2*(6), 451–459.
15. Anderson, R. T., Aaronson, N. K., & Wilkin, D. (1993). Critical review of the international assessments of health-related quality of life. *Quality of Life Research, 2*(6), 369–395.
16. Mutz, Diana C. (2011). *Population-based survey experiments*. Princeton: Princeton University Press.