Using vignettes to understand Latino/White[1] disparities in self-rated health by domain

Extended Abstract, Submitted for Population Association of America 2013 Annual Meeting

Sharon Bzostek, Rutgers University
Denise Duffy, University of Michigan
Noreen Goldman, Princeton University
Anne Pebley, University of California-Los Angeles
Narayan Sastry, University of Michigan
(authors are listed in alphabetical order)

Previous research documents that Hispanics report worse overall self-rated health than non-Hispanic whites in the United States (Arcia et al. 2001; Bzostek, Goldman & Pebley, 2007; Finch, Hummer, Reindl, & Vega, 2002; Ren & Amick, 1996). This is true despite evidence suggesting that for many other measures of health, Latinos fare better than would be expected (relative to non-Hispanic whites) given their low average levels of socioeconomic status (a phenomenon often referred to as the "Hispanic health paradox.") (Franzini, Ribble, & Keddie, 2001; Morales, Lara, Kington, Valdez, & Escarce, 2002; Williams, 2001).

Past studies documenting Hispanic-white disparities in overall self-rated health and trying to understand the reason for Hispanics' worse self-rated health have suggested a number of potential reasons for this finding. For example, there may be differences by nativity status and level of acculturation in social norms regarding ways of thinking about and describing one's health. For example, more traditionally-oriented Hispanic immigrants may find it less socially acceptable to rate their health highly, or to "boast" (Angel & Angel 1992). There is also some evidence that at least some of the difference in Hispanics' and whites' self-ratings may be due to differences in the meaning of the various response categories (particularly "fair" in English and "*regular*" in Spanish) in the two languages (Angel & Guanarccia 1989; Bzostek et al. 2007;

---

[1] The terms "Hispanic" and "Latino" are used interchangeably in this paper. The term "white" refers to non-Hispanic white.

Franzini & Fernandez-Esquer, 2004; Phillips et al. 2005). Alternatively, there may be differences in the weighting of particular domains of health across groups, leading to differential self-rated health. In particular, this argument suggests that Hispanics may feel less comfortable expressing emotional distress, and may be more likely to "somatize," or to express emotional distress through physical symptoms. In this way, mental illness or depression may have more of an impact on Hispanics' overall self-rated health (Angel & Guarnaccia 1989; Finch, Kolody, & Vega, 2000; Finch, Hummer, Kolody, & Vega, 2001).

Although previous studies have tried to evaluate these explanations, standard analytical techniques are problematic for answering this question. Using standard techniques, it is difficult to determine how much of the difference between Hispanic and whites' self-rated health is due to "true" differences in health, rather than the fact that these groups (or subgroups within these broad groups) may rate their health differently. For example, the groups may place differential weights on particular dimensions of health status, or have systematically different rating styles that could result in misleading conclusions about differences among the groups' self-ratings.

In this paper, we use new anchoring vignette data from the second wave of the Los Angeles Families and Neighborhood Study (L.A.FANS-2, conducted in 2006-2008) to try to identify and adjust for systematic differences in rating styles in Hispanic and non-Hispanic whites' self-rated health by domain. Adjusting self-reports by using anchoring vignettes is a method designed to help standardize self-assessment ratings across individuals. The method is based on the idea that in any self-rating, differences could be due to either true differences in what researchers are seeking to measure or to different rating styles (often called response category differential item functioning, or DIF) (King et al. 2004, King & Wand 2007). Some people or groups, for example, may simply be more optimistic raters than others. This method

asks respondents to rate a hypothetical individual's health based on a given scenario. These hypothetical vignette ratings can then be used to adjust respondents' self-ratings for their general rating style. After this adjustment is made, any remaining differences across respondents' ratings should be due to actual differences in the respondents' underlying health rather than differential rating styles. This method relies on two key assumptions: that the hypothetical individual's health does not vary across survey respondents like the respondents' own actual health would (known as "vignette equivalence") and that respondents use the same response selection process for vignettes and self-reports (known as "response consistency.")

Methodologically, traditional ordered probit models or ordered logit models are the most commonly-used modeling strategies for studying self-rated health. Although these models assume constant cutpoints across individuals, in actuality, these cutpoints may vary across groups. More optimistic raters, for example, may have lower thresholds than other raters for reporting fair versus poor or very good versus good health. For this reason, two individuals with an identical level of "true" health might report a different level of health solely due to differences in reporting styles. Vignette ratings allow the researcher to relax the assumption of constant cutpoints across individuals by modeling and adjusting for different thresholds across respondents.

The anchoring vignettes method has been used for a wide range of substantive topics. A handful of studies to date have used this method to address rating differences in domains of self-rated health, focusing specifically on topics like adjusting self-ratings of work-related disabilities for reporting differences across countries (Kapteyn, Smoth & Van Soest 2007), differences across subdomain ratings for older individuals based on age, sex and race/ethnicity (Dowd & Todd 2011), and differences in overall self-rated health between men and women (Grol-

Prokopcyzk, Freese & Hauser 2011). Grol-Prokopcyzk et al. (2011) found, for example, that although unadjusted models suggested better self-rated health among women, this difference disappeared after the authors adjusted for women's greater "health-optimism" in self-rated health. Dowd and Todd (2011) found evidence of considerable differences in reporting styles across demographic groups, with great variation by the particular health subdomain under consideration.

To date, no study (to our knowledge) has used anchoring vignettes to study Hispanic/white differences in self-rated health in-depth. Although Dowd and Todd's 2011 study of older individuals did compare reports between Hispanics and non-Hispanic whites (as well as non-Hispanic blacks), their analysis focused on larger questions of reporting differences across groups, and did not provide a detailed analysis of Hispanic/white differences.

In this study, we build upon this growing body of research that suggests the importance of identifying and adjusting for systematic reporting differences across groups. In particular, we use the new anchoring vignettes from L.A.FANS-2 to understand differences in subdomains of self-rated health between non-Hispanic whites and Hispanics. Table 1 presents a summary of the vignettes and the corresponding self-rating questions included in L.A.FANS-2, representing six subdomains of health drawn from the SF-36 health assessment (physical functioning, role limitations, bodily pain, vitality, social functioning, and mental health). Our primary analyses compare results obtained using standard ordered logit models with those from hierarchical ordered logit models, which allow thresholds to vary across individuals and by explanatory variables, using the data drawn from the vignettes for identification (King et al., 2004; King & Wand 2007). In particular, we address the following three research questions:

1) Do Latinos consistently rate their health as worse (or better) than whites across

   subdomains of health? Does this difference in rating vary based on national origin,

   nativity status, and age at immigration?

2) Is there evidence of DIF among these groups (systematic differences in reporting styles)?

3) Do our conclusions regarding self-ratings of health in the subdomains by

   ethnicity/nativity status change once we have adjusted for any DIF?

Preliminary Findings

Evidence of Differential Item Functioning

Results from ordered logit models (not shown here) predicting respondents' ratings of the

severity of the vignettes across the subdomains indicate that, after controlling for a series of

socio-demographic and health-related characteristics, Latino groups tend to rate the vignettes *less*

severely than whites. Those interviewed in Spanish also appear to rate the vignettes less severely

than those interviewed in English. This provides some initial evidence that DIF may be an

important problem to consider when comparing self-ratings of health across these groups.

Comparing racial/ethnic differences in self-rated health before and after DIF adjustment

Table 2 compares results from standard ordered logistic regression models predicting

self-rated health across the subdomains of health with hierarchical ordered logistic regression

models that adjust for DIF across groups. Our substantive conclusions regarding the magnitude

and direction of differences in self-rated health between the Latino and white subgroups in our

sample change in the case of emotional problems and energy levels after adjusting for DIF. In

both cases, most of the Latino subgroups have significantly worse self-ratings (relative to native-

born whites) in the models that were adjusted for DIF, but not in the unadjusted models.

Understanding systematic differences in reporting styles

Figure 1 displays mean values for four separate threshold levels that together define the five response categories in the adjusted models for the ratings of energy levels. These thresholds measure the level of (lack of) energy required for respondents in particular groups to report a higher (more severe) rating of energy problems. As the figure demonstrates, non-Hispanic whites appear to consistently have the lowest thresholds for moving to the next, more severe rating category. In other words, in the case of energy levels, the non-Hispanic white respondents in the sample (net of other characteristics) appear to have a more pessimistic rating style than the various Latino subgroups. As we saw in Table 2, adjusting for this difference in reporting style shows us that the Latino groups tend to have worse self-ratings of energy level.
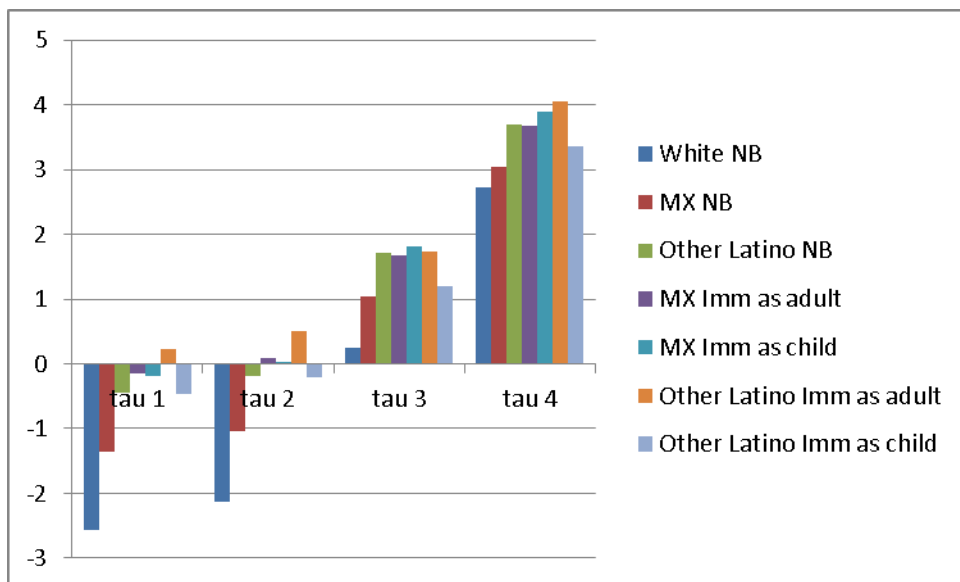
Figure 1. Mean threshold levels by race/ethnicity/immigration status for each of the four thresholds, for ratings of low energy

Analysis Plan and Expected Results

Additional analyses to be conducted will compare among the Latino subgroups and consider the role of potentially important factors like age at immigration and linguistic and social

acculturation.We will also test the sensitivity of our findings to different modeling specifications. Specifically, we will test the robustness of the results across models using different groups of variables as threshold predictors and test alternative classification schemes for our race/ethnicity, nativity status, and age at immigration variables.

The results of this study will provide useful new information about differences in the conceptualization and reporting of self-rated health by Hispanics and non-Hispanic whites, as well as among different Hispanic subgroups (such as those defined by nativity status and level of acculturation). By assessing if---and how---failing to account for potential differences in reporting style leads to misleading comparisons among these groups, our findings could have important implications for our substantive understanding of disparities in self-rated health by race/ethnicity.

Table 1. Summary of L.A.FANS-2 Self-Reported Health Status and Vignettes

1. Physical functioning	During the past 4 weeks, how much did health problems limit your physical activities (such as walking or climbing stairs)? Would you say not at all, very little, somewhat, a lot, or severely?"	[NAME] goes walking every day for half an hour, about one mile. [NAME] does not do any strenuous sports because she/he feels out of breath when he/she walks very quickly or runs. How much did health problems limit [NAME]'s physical activities? Would you say not at all, very little, somewhat, a lot, or severely?
2. Role limitations	During the past 4 weeks, how much difficulty did you have doing work, both at home and on the job, because of health or emotional problems? Would you say none at all, very little, some, a lot, or severe?	[NAME] suffers from allergies every month. Because of the symptoms, he/she is unable to go to work for one or two days but has no problem catching up with his/her tasks. How much difficulty did [NAME] have doing work because of his/her health? Would you say none at all, very little, some, a lot, or severe?
3. Social functioning	During the past 4 weeks, how much did health or emotional problems limit your social activities with family or friends? Would you say not at all, very little, somewhat, a lot, or severely?	[NAME] is usually an outgoing and cheerful person who has many friends and enjoys going out. Three or four days a month, he/she feels sad all day so tends to avoid people. How much did health or emotional problems limit [NAME]'s usual social activities with family or friends? Would you say not at all, very little, somewhat, a lot, or severely?
4. Pain	How much physical pain did you have during the past 4 weeks? Would you say none, very mild, mild, moderate, or severe?	[NAME] has a headache once a month that gets better if he/she takes a pill. When he/she has a headache, he/she can continue to do her/his normal activities. How much physical pain did [NAME] have? Would you say none, very mild, mild, moderate, or severe?
		[NAME] has pain that radiates down his/her right arm and wrist when he/she is working on a computer at work. It is slightly better in the evenings when he/she does not use a computer. How much physical pain did [NAME] have? Would you say none, very mild, mild, moderate, or severe?
5. Vitality	During the past 4 weeks, how much energy did you have? Would you say none, a little, some, a lot, or very much?	[NAME] is not a physically active person but enjoys a walk around the neighborhood most weekends. Whenever he/she walks a mile or more, he/she feels tired afterwards and needs to rest for an hour or so. How much energy did [NAME] have? Would you say none, a little, some, a lot, or very much?
		[NAME] feels tired every afternoon, which makes any task that he/she does a great effort. Whenever he/she does the dishes, tidies the house, or prepares a meal for more than 10 minutes he/she needs to sit down and rest. How much energy did [NAME] have? Would you say none, a little, some, a lot, or very much?
6. Emotional problems	During the past 4 weeks, how much have you been bothered by emotional problems (such as feeling anxious, depressed or irritable)? Would you say not at all, very mildly, mildly, moderately, or severely?	[NAME] feels nervous and anxious. He/She worries and thinks negatively about the future, but feels better when he/she is not alone or when doing something that really interests him/her. When he/she is alone he/she tends to feel useless and empty. How much was [NAME] bothered by emotional problems? Would you say not at all, very mildly, mildly, moderately, or severely?

Table 2. Coefficients from ordered logits predicting self-ratings of health across domains, with and without vignette adjustment

	Emotional problems		Role limitations		Social limitations		Phys. Act. Lims.		Physical pain		Energy level	
	Unadj.	Adj.	Unadj.	Adj.	Unadj.	Adj.	Unadj.	Adj.	Unadj.	Adj.	Unadj.	Adj.
Race/ethnicity and nativity status												
MX native-born	0.01	0.00	-0.11	0.25	-0.15	-0.38	0.04	-0.01	-0.20	-0.30	0.06	0.45 ^
Non-MX native-born	0.18	0.85 ^	0.32	1.17 *	-0.07	0.26	0.60 ^	1.17 *	0.09	0.57	-0.06	0.94 *
MX, Imm as adult	0.28	1.12 *	0.33	0.58	-0.05	0.02	0.27	0.29	-0.09	0.01	0.26	0.51
MX, Imm as child	0.25	1.22 **	0.37	0.52	0.08	0.34	0.43	0.35	0.04	0.30	0.19	0.77 *
Non-MX, Imm as adult	0.18	1.17 *	0.18	0.47	-0.23	-0.01	0.04	0.21	-0.37	-0.43	0.24	0.67
Non-MX, Imm as child	-0.02	1.04 ^	0.52	1.19 *	0.05	0.56	0.27	0.34	-0.61	-0.63	0.25	0.78
Interviewed in Spanish	-0.07	0.29	-0.19	0.23	-0.27	-0.47	-0.13	-0.20	-0.02	0.85 **	-0.27	0.11
Linguistic acculturation	0.11	0.06	0.13	0.07	-0.14	-0.39	0.18	0.03	0.07	0.02	0.12	-0.16
Social acculturation	0.02	0.01	0.11	0.18 ^	0.06	0.10	0.26 **	0.33 **	0.16 **	0.31 **	-0.04	-0.12
Vignette 1		4.54 **		5.02 **		4.48 **		4.80 **		4.74 **		2.82 **
Vignette 2										3.71 **		3.75 **

Note: All models also include controls for an extensive set of covariates in predicting the outcomes, and race/ethnicity/nativity status, language of interview, and linguisting and social acculturation in predicting the thresholds.

References

Arcia, E., Skinner, M., Bailey, D., & Correa, V. (2001). Models of acculturation and health behaviors among Latino immigrants to the US. *Social Science & Medicine, 53*(1), 41-53.

Angel, J. L., & Angel, R. J. (1992). Age at migration, social connections, and well-being among elderly Hispanics. *Journal of Aging and Health, 4*(4), 480-499.

Angel, R., & Guarnaccia, P. J. (1989). Mind, body, and culture: Somatization among Hispanics. *Social Science & Medicine, 28*(12), 1229-1238.

Bzostek, S., Goldman, N. & Pebley, A. (2007). Why do Hispanics in the USA report poor   health? *Social Science and Medicine, 65*, 990 – 1003.

Dowd, J.B., & M. Todd. (2011). Does self-reported health bias the measurement of health inequalities in U.S. adults? Evidence using anchoring vignettes from the health and retirement study. *The Journals of Gerontology, Series B: Psychological Sciences and   Social Sciences, 66(4)*, 478–489,

Finch, B. K., Hummer, R. A., Kolody, B. & Vega, W. A. (2001). The role of discrimination and acculturative stress in the physical health of Mexican-origin adults. *Hispanic Journal of Behavioral Sciences, 23*(4), 399-429.

Finch, B. K., Hummer, R. A., Reindl, M., & Vega, W. A. (2002). Validity of self-rated health among Latino(a)s. *American Journal of Epidemiology, 155*(8), 755-759.

Finch, B. K., Kolody, B., & Vega, W. A. (2000). Perceived discrimination and depression among Mexican-origin adults in California. *Journal of Health and Social Behavior, 41*(3), 295-313.

Franzini, L., & Fernandez-Esquer, M. E. (2004). Socioeconomic, cultural, and personal influences on health outcomes in low income Mexican-origin individuals in Texas. *Social Science & Medicine, 59*(8), 1629-1646.

Franzini, L., Ribble, J. C., & Keddie, A. M. (2001). Understanding the Hispanic paradox. *Ethnicity and Disease, 11*(3), 496-518.

Grol-Prokopczyk H., J. Freese, & R. Hauser. (2011). Using anchoring vignettes to assess group differences in general self-rated health. *Journal of Health and Social Behavior, 52*, 246.

King, G. & J. Wand. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis, 15*, 46 – 66.

King, G., C. Murray, J. Salomon, & A. Tandon. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review, 98*, 191–205.

Kapteyn, A., J. Smith, A. Van Soest. (2007). Vignettes and self-reports of work disability in the   United States and the Netherlands. *The American Economic Review, 97*(1), 461 – 473.

Morales, L. S., Lara, M., Kington, R. S., Valdez, R. O., & Escarce, J. J. (2002). Socioeconomic, cultural, and behavioral factors affecting Hispanic health outcomes. *Journal of Health Care for the Poor and Underserved, 13*(4), 477-503.

Phillips, L. J., Hammock, R. L., & Blanton, J. M. (2005). Predictors of self-rated health status among Texas residents. *Preventing Chronic Disease: Public Health Research, Practice and Policy, 2*(4).

Ren, X. S., & Amick, B. C. (1996). Race and self assessed health status: The role of socioeconomic factors in the USA. *Journal of Epidemiology and Community Health, 50*, 269-273.

Williams, D. R. (2001). Racial variations in adult health status: Patterns, paradoxes, and prospects. In N. J. Smelser, W. J. Wilson & F. Mitchell (Eds.), *America becoming: Racial trends and their consequences* (Volume 2). National Research Council, Commission on Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.