

Effect of Language on Mental Health
Reporting Among Hispanics

Laura Flicker, Elizabeth Dean, Rosanna Quiroz, Mai Wickelgren
RTI International

Laura DiGrande, and Tina McVeigh
New York City Department of Health and Mental Hygiene

Paper presented at the 2006 annual conference of the American Association for Public Opinion
Research (AAPOR), Montreal, Canada, May 18-22

RTI International is a trade name of Research Triangle Institute
3040 Cornwallis Road
P.O. Box 12194
Research Triangle Park, North Carolina, USA 27709
Phone 919-316-3510 Fax 919-541-1261 e-mail lflicker@rti.org

Abstract

Our research examines mental health items in two surveys administered in Spanish and English. The first is the World Trade Center Health Registry (WTCHR), a population-based study of persons exposed to the World Trade Center attacks of September 11, 2001. The second is the National Survey on Drug Use and Health (NSDUH), an annual household survey that collects data on drug/alcohol use and mental health symptoms.

Both surveys employed the K-6, a mental health screener that asks respondents how often they feel sad, nervous, worthless, hopeless, restless, and that everything is an effort. The WTCHR results found that Hispanics in general and Spanish-speakers in particular had such feelings more often in the past 30 days compared to non-Hispanics. By extension, they were found to have higher rates of “psychological distress” according to the K-6 diagnostic. The K-6 items from the NSDUH asked about a longer reference period -- the worst month out of the past 12 months -- yet the differences between Hispanics and non-Hispanics were much smaller.

The current research analyzes each study’s data to explain how question translation and language preference may be related to mental health reporting by Hispanics. In addition to comparing point estimates for mental health status, this work models psychological distress (as measured by the K-6) using predictors for ethnicity, language of survey administration, and other demographic variables.

Keywords: Translation, mental health, Hispanics

Introduction

Questionnaires and symptom checklists have long been used to diagnose physical health or disability in clinical settings. Many instruments have been developed to assess mental health in lieu of clinical interviews. More recently, check lists and frequency measures have been adapted to population surveys to estimate prevalence of mental illness. Adapting measures for population surveys may introduce bias through ad hoc wording changes, changes to reference periods, translation to foreign languages, and differing cultural sensitivity to certain topics.

Epidemiologists refer to the Hispanic paradox, the fact that Hispanic immigrants have better health status than native-born Hispanics. The concept has been applied to physical rather than mental health. In contrast, a culturally defined syndrome of “nervios” has been identified by social scientists and clinicians in reference to Hispanic mental health (Guarnaccia, et al., 1989). Actual empirical research on Hispanic mental health has produced conflicting results. The National Co-Morbidity Survey (NCS) obtained higher rates of psychological distress among Hispanics than the general population (Kessler, et. al, 1994). Mexican Americans were found to have better mental health status despite economic disadvantages (Escobar, et al., 1998; Farley et al., 2005). Time spent in the US has been associated with higher rates of mental illness and drug and alcohol abuse among Mexican Americans (Vega, et al., 1998).

A surprising shortcoming in epidemiological literature about mental health is a systematic exploration of language and translation effects. For that matter, the aforementioned NCS was the first population survey designed to diagnostic interview to assess prevalence of DSM-III-R disorders in the general population, but was not administered in Spanish. A component of the NCS, the so-called K-6, was employed in the National Health Interview Survey (NHIS). A recent examination of NHIS data found that the increased likelihood of psychological distress among non-English speakers is readily reversed by introducing controls for education, employment and marital and status (Bratter and Eschbach, 2005).

Method

The K-6 is a six-item measure designed to measure the construct, “nonspecific psychological distress” in the general population. Its questions ask respondents about frequency for feeling nervous, worthless, sad, hopeless, restless, and that everything is an effort. The K-6 collects a summary score based on the self-reported frequency for each emotion: 0=None of the time; 1=A little of the time; 2=Some of the time; 3=Most of the time; 4=All the time. A cumulative score of 13 or higher for the six items produces “diagnostic” for psychological distress.

In 2004, RTI International administered the K-6 for two projects, the World Trade Center Health Registry (WTCHR) and the National Survey on Drug Use and Health (NSDUH). Both employed Spanish translations, but they differed in their survey designs and topics of interest. The K-6 translations differed slightly by virtue of different staff completing that work, but the measure retained its original scoring scheme in both the English and Spanish.

The World Trade Center Health Registry (WTCHR) was a collaborative effort by New York City Department of Health and Mental Hygiene and the Agency for Toxic Substances and Disease

Registry. It surveyed persons most exposed to events of 9/11 in New York City and included self-enrollment and list samples. The interviews were conducted primarily by telephone. They collected data about health conditions before/after 9/11. Data collection occurred between September, 2003 and November, 2004. Its K-6 items asked about feelings in the past 30 days and thus measured psychological distress in the past month. The dataset contained 10,676 non-proxy interviews for residents, and our analysis was based on 1,400 Hispanics completing interviews in Spanish or English.

The National Survey on Drug Use and Health (NSDUH) is an annual household survey sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA). It produces estimates of tobacco, drug, and alcohol use and mental health. Its design is an area probability survey that samples about 170,000 dwelling units and administers some 67,500 interviews annually via an audio computer-assisted self interview (ACASI). Its K-6 items asked about feelings in worst month in the past 12 months and thus diagnosed psychological distress in the past year. As such, it measured psychological distress in the past 12 months. Analyses accessed 2004 public use files, retained “long form” respondents,¹ and applied weights for the 2,605 Hispanic cases.

Before beginning analyses, an RTI language methodologist and co-author reviewed the K-6 translations for the WTCHR and NSDUH. Although the translations were not identical (having been translated at different times by different staff) most were quite similar. (See Table 1.) Two exceptions were questions asking about feelings of sadness and worthlessness. For sadness, the WTCHR and NSDUH retained the original K-6 wording, “...how often did you feel so sad or depressed that nothing could cheer you up?” However, the WTCHR translation used the Spanish cognate for “console” (*consolar*). That wording was thought to be more strongly worded than the corresponding NSDUH translation that used a word that conveyed the idea of “encourage” or “enliven” (*animar*). For worthlessness, the English version of NSDUH was the same as the original K-6, “...how often did you feel down on yourself, no good, or worthless?” The corresponding English question in the WTCHR asked only, “...how often did you feel worthless?” As such, the NSDUH was more strongly expressed in both English and Spanish translation.

These differences led to two hypotheses based on expectations for biased wording or translation. First, that Spanish-speaking Hispanics would less often endorse the sadness item in the WTCHR compared to the NSDUH. Second, that endorsement of the worthlessness item would be lower in NSDUH than WTCHR for both languages. The second hypothesis was more difficult to test, due to design and population differences between the two studies. However, it was thought that relative ratios for subgroups may differ for worthlessness in comparison to other K-6 items, thereby suggesting effects from the wording corresponding translation difference.

Apart from translation effects, we were interested in differences among Hispanics who responded in English versus Spanish. Those differences would be expressed both in the individual items comprising the K-6 and the summary score, or diagnostic, for psychological distress. The direction

¹ The 2004 NSDUH employed an experiment in which the K-6 was administered alone (“short form”) and in the context of other mental health questions (“long form”). Due to unusually high reporting on the short form, SAMHSA elected to retain only long form data in its 2004 reports of mental illness in the general population. A discussion of the psychometric properties discovered in the two forms is discussed in a methodological report by Aldworth, et. al. (2004).

of any such differences was not predicted beforehand. Item level nonresponse was also of interest, since higher rates in Spanish language could suggest translation problems.

Results

Of the 1,400 WTCHR non-proxy Hispanic interviews, 1,232 respondents completed the interview in English and 168 completed it in Spanish. NSDUH analysis retained 2,605 Hispanics who responded to the long form version of the K-6 in the 2004 data collection. Of those, 1,690 completed the interview in English and 915 in Spanish.

Table 2 describes the percentages of respondents answering “All of the time” or “Most of the Time” for each of the K-6 items, for each study. The hypothesis that nervousness and worthlessness lower in NSDUH than WTCHR was not specifically supported. All K-6 items obtained higher rates of endorsement in the WTCHR, not just these two. Moreover, their relative rates of endorsement (NSDUH/WTCHR) were not lower than the other four items. The second hypothesis, that sadness would be less often endorsed by Spanish-speakers on the WTCHR was contradicted. In fact, relatively more Spanish-speaking Hispanics endorsed sadness in WTCHR relative to NSDUH when one considers the relative ratio -- $32.7/21.8=1.5$. Relatively fewer English-speaking Hispanics endorsed sadness question in the WTCHR relative to NSDUH -- $18.3/13.8=1.3$.

Table 3 describes interesting patterns of item nonresponse (refusals, don't know) to the K-6. In WTCHR it was consistently higher among Spanish-speaking Hispanics than English-speaking Hispanics for all items except hopelessness, for which both rates were equal (0.6%). Some differences were very high, such as worthlessness (0.5% English, 5.4% Spanish), hopelessness (0.4% English, 7.7% Spanish), and effort (1.8% English, 7.7% Spanish). Compared to nonresponse for other items in the WTCHR interview, these are high rates indeed.

Conversely, K-6 item nonresponse on the NSDUH was higher for English-speaking Hispanics, with the exception of sadness (2.1% English, 2.4% Spanish). Elsewhere nonresponse differences between the languages did not attain a factor 5 to 10 as in the WTCHR, but only 2 to 5. Nonresponse was highest for effort (3.5% English, 1.6% Spanish) with similar results for hopelessness (3.3%, 1.4%) and restlessness (3.2% and 1.4%). Nonresponse for other NSDUH items was similar or even higher, as might be expected for the number of questions about illegal behavior (e.g., drug use).

Although the K-6 reference period for the NSDUH was 12 months and for the WTCHR 30 days, psychological distress for Hispanics was twice as high in the WTCHR (20.2%) as the NSDUH (10.2%). The rate for non-Hispanics was lower in the WTCHR (8.0%) than the NSDUH (9.6%). Differences by language of survey administration were especially interesting. English-speaking Hispanics in the WTCHR were more often diagnosed as distressed according to the K-6 than those in the NSDUH -- 18.1% versus 10.6%. The difference for prevalence among Spanish-speaking Hispanics was even more striking -- 35.7% in the WTCHR versus 9.8% in the NSDUH.

Psychological distress was more common among respondents reporting low income. This is widely expected from epidemiological studies of mental illness in the general population. However, low income was associated a much higher rate of distress among WTCHR respondents compared to

those in NSDUH (20.0% versus 15.4%) and high income a lower rate in the WTCHR than the NSDUH (5.4% versus 8.1%). An interesting contradiction between the studies was the fact that more WTCHR respondents under age 50 reported psychological distress than NSDUH respondents -- 8.2% versus 11.0%. Conversely, respondents over the age of 50 obtained virtually reversed rates in the two studies -- 11.7% versus 7.6%.

Working with two studies' data required a means of controlling for differences in design in order to isolate possible effects of language and ethnicity. For that reason, logistic models were developed to predict psychological distress (See Table 5.) The models relied on the larger datasets that included non-Hispanics and employed similar predictors: Hispanic, Spanish language, low income, and under age 50 and were run separately for each study.² In the WTCHR, Hispanics were twice as likely and low income respondents more than 3.5 times as likely to suffer psychological distress in the past 30 days ($p < .001$). Spanish speakers (almost all Hispanics) were half again as likely to have distress, whereas respondents under the age of 50 were about three-fourths as likely to have distress ($p < .05$). In the NSDUH, neither Hispanics nor Spanish speakers (again almost all Hispanic) were more or less likely to be distressed. Low income respondents were more than twice as likely to be distressed, and those under the age of 50 were half again as likely to be distressed ($p < .005$).

With uneven results produced by these studies, we sought validation for the K-6 diagnostic from other items from the WTCHR and NSDUH interviews. The NSDUH "long form" used in the analyses included many other questions relating to mental health. Two items in particular discussed mental health *treatment* in the past 12 months, the same period referenced in the NSDUH adaptation of the K-6. Among respondents who were classified as having psychological distress, fewer Spanish-speaking Hispanics (19.2%) had mental health treatment compared to English-speaking Hispanics (27.9%). However, this may reflect objective differences in health care access and/or utilization. More interesting was the perceived *unmet need* for mental health treatment. Among those found to have psychological distress, less than half the Spanish-speaking Hispanics reported an unmet need for mental health treatment in the past 12 months (12.3%) compared with of English-speaking Hispanics (27.2%).

The main focus for the WTCHR interview was reported physical conditions before and after 9/11. Only two mental health items outside of the K-6 were available to provide external support for that diagnostic. The first was reported lifetime feelings of depression. That measure was virtually identical for Spanish and English-speaking Hispanics, 33.8% and 33.3%, respectively. Among the subset who reported having lifetime feelings of depression, the percent reporting *increased* feelings of depression since 9/11 were extremely high for Spanish-speaking Hispanics (94.1%) and almost unanimous among English-speaking Hispanics (98.3%).

Conclusions

Before beginning this current work, the authors had no hypotheses for expected relationships between ethnicity and psychological distress, given conflicting evidence in mental health research.

² Dummy variables for population density were added to the NSDUH model to control for urban status and better explain differences with WTCHR findings. However those predictors did not change the parameters or resulting odds ratios, and they were subsequently dropped. Interaction terms for Hispanic x income and Hispanic x age were added to the NSDUH and WTCHR models, but they produced no effect and were subsequently dropped.

Our results found higher rates of psychological distress for Hispanics than non-Hispanics in both studies. But where the ethnic differences in the NSDUH with regards to prevalence was small (10.2% Hispanics versus 9.6% for non-Hispanics), it was quite large in the WTCHR (20.0% Hispanics versus 8.0% non-Hispanics). Essentially the prevalence rate for distress among Hispanics was twice as high in the New York study as the general population. For that matter, psychological distress was predicted by ethnicity, language, age, and income in WTCHR to statistical significance. Future analyses may introduce socioeconomic predictors employed in Bratter and Eschbach's analysis (2005) to try and eliminate the effects of ethnicity and language in the WTCHR. It is interesting that in the NSDUH, effects of ethnicity and language were not apparent at all; psychological distress was simply predicted by income and age.

The Community Health Survey (CHS) is another telephone survey administered in New York City by the NYC Department of Health and Mental Hygiene. It also employs the K-6. CHS obtained age-adjusted prevalence rates for Hispanics of 11.8% in 2002 and 10.3% in 2003 (McVeigh, et. al., 2006). Those rates were higher than for whites, blacks, Asians, and other respondents, but not by a large factor. Furthermore, those authors determined that residential proximity to the World Trade Center site did not affect reported psychological distress.

It seems likely that factors unique to WTCHR, including its design (recording exposures from the 9/11 attacks) and respondent pool (mostly self enrolled), helped produce higher prevalence rates than reported elsewhere for New York City or for the city's Hispanics. Unfortunately, the NSDUH public use file had no geographic variables to inform our discussion about this area. Adding NSDUH predictors for population density had no impact on the logistic model and they were dropped.

While our findings suggest language effects in WTCHR, those were not clearly linked to translation *per se*. Specifically, the Sadness question was more strongly translated in the WTCHR and should have produced lower endorsement among Spanish-speaking Hispanics. However, it was more often endorsed among those persons than among English-speaking Hispanics. Yet it was notable that in the WTCHR, Spanish item nonresponse on K-6 was much higher than English, and higher than other survey questions. Translations may not have been readily understood by Spanish speakers in the New York City sample.

It seems unlikely that the nonresponse reflects social desirability bias, given the otherwise high levels of agreement with the K-6 items. For that matter, it would be helpful for future work involving the K-6 or other psychological measures to collect measure of social desirability. This will help determine whether Hispanics are more willing to report mental health symptoms, in the absence of any perceived need for treatment.

Isolating the effect of language *per se* is difficult in applied settings. Spanish language preference can be a proxy measure for acculturation, which itself may affect actual mental health status. At a minimum, further research with K-6 and similar measures would help assess its performance in various subpopulations. This may include cognitive interviews to determine how translations of K-6 are understood or other qualitative analysis with small samples stratified by response to K-6. Finally, the K-6 may be less prone to mode effects, given the readiness of self report by telephone in the WTCHR.

References

Aldworth, J., Chromy, J. R., Foster, M., Heller, D., & Novak, S. (2005, October). Serious Psychological Distress Report. In 2004 National Survey on Drug Use and Health: Methodological resource book (prepared for the Substance Abuse and Mental Health Services Administration, Office of Applied Studies, under Contract No. 283-03-9028, RTI/0208726.187.013). Research Triangle Park, NC: RTI International. [Available as a PDF at <http://www.oas.samhsa.gov/nhsda/methods.cfm#2k4>]

Bratter, J., Eschbach, K. (2005). Race/Ethnic Differences in Nonspecific Psychological Distress: Evidence from the National Health Interview Survey. *Social Science Quarterly*, Sep; 86(3):620-644.

Escobar, J.I. (1998). Immigration and Mental Health: Why Are Immigrants Better Off? *Archives of General Psychiatry*. Sep; 55(9):781-2.

Farley, T., Galves, A., Dickenson, L.M., Perez, Mde J. (2005). Stress, Coping And Health: A Comparison of Mexican Immigrants, Mexican-Americans, and non-Hispanic Whites. *Journal of Immigrant Health*, 7(3):213-220.

Guarnaccia, P.J., Angel, R., and Worobey, J.L. (1989). The Multiple Meanings of Ataques de Nervios in the Latino Community. *Medical Anthropology*, May, 11(1):47-62.

Kessler, R.C., McGonagle, K.A., Zhao, S., Nelson, C.B., Hughes, M., and Eschleman, S., Wittchen, H.U., Kendler, K.S. (1994). Lifetime and 12-Month Prevalence of DSM-III-R Psychiatric Disorders in the United States: Results from the National Co-Morbidity Survey. *Archives of General Psychiatry*, 51(1):8-19.

McVeigh, K.H., Galea, S., Thorpe, L.E., Maulsby, C., Henning, K., and Sederer, L.I. (2006). The Epidemiology of Nonspecific Psychological Distress in New York City, 2002 and 2003. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, May; 83(3): 394 - 405

Vega WA, Kolody B, Aguilar-Gaxiola S, Alderete E, Catalano R, Caraveo-Anduaga J. (1998). Lifetime Prevalence of DSM-III-R Psychiatric Disorders Among Urban and Rural Mexican Americans in California. *Archives of General Psychiatry*. 1998 Sep; 55(9):771-8.

Table 1. Comparison of NSDUH and WTCHR Versions of K-6 Items

2004 NSDUH		WTCHR		Translation Remarks
English	Spanish (masculine form)	English	Spanish (masculine form)	
In that worst month, how often did you feel nervous?	Ese mismo mes en que sus emociones estaban de lo peor, ¿con qué frecuencia estaba nervioso?	During the past 30 days, that is, since [REFERENCE DATE] how often did you feel nervous?	Durante los últimos 30 días, es decir, desde [REFERENCE DATE] ¿con qué frecuencia se sintió nervioso?	Translations are similar. Estabar and sintarse have similar meanings.
In that worst month, how often did you feel so sad or depressed that nothing could cheer you up?	Ese mismo mes en que sus emociones estaban de lo peor, ¿con qué frecuencia se sintió tan triste o deprimido que no había nada que lo podía animar?	During the past 30 days, that is, since [REFERENCE DATE] how often did you feel so sad that nothing could cheer you up?	Durante los últimos 30 días, es decir, desde [REFERENCE DATE] ¿con qué frecuencia se sintió tan triste que nada lo podía consolar?	WTC translation is worded more strongly by saying “nothing could console you”.
In that worst month, how often did you feel hopeless?	Ese mismo mes en que sus emociones estaban de lo peor, ¿con qué frecuencia se sintió desolado?	During the past 30 days, that is, since [REFERENCE DATE] how often did you feel hopeless?	Durante los últimos 30 días, es decir, desde [REFERENCE DATE] ¿con qué frecuencia se sintió sin esperanzas?	The NSDUH question says “how often did you feel desolate?”, while the WTC question asks “how often did you feel you had no hope”.
In that worst month, how often did you feel restless or fidgety?	Ese mismo mes en que sus emociones estaban de lo peor, ¿con qué frecuencia se sintió agitado o inquieto?	During the past 30 days, that is, since [REFERENCE DATE] how often did you feel restless or fidgety?	Durante los últimos 30 días, es decir, desde [REFERENCE DATE] ¿con qué frecuencia se sintió inquieto o intranquilo?	Translations are similar. Agitado and intranquilo have similar meanings.
In that worst month, how often did you feel everything was an effort?	Ese mismo mes en que sus emociones estaban de lo peor, ¿con qué frecuencia sintió que para todo tenía que hacer un gran esfuerzo?	During the past 30 days, that is, since [REFERENCE DATE] how often did you feel that everything was an effort?	Durante los últimos 30 días, es decir, desde [REFERENCE DATE] ¿con qué frecuencia se sintió que todo le costaba trabajo?	Translations are similar. Esfuerzo and “trabajo” have similar meanings.
In that worst month, how often did you feel down on yourself, no good, or worthless?	Ese mismo mes en que sus emociones estaban de lo peor, ¿con qué frecuencia se sintió mal consigo mismo o que usted no valía nada?	During the past 30 days, that is, since [REFERENCE DATE] how often did you feel worthless?	Durante los últimos 30 días, es decir, desde [REFERENCE DATE] ¿con qué frecuencia se sintió sin ningún valor?	The added terms on NSDUH make it stronger than the WTCHR.

Table 2. Hispanic Responses to WTCHR and NSDUH K-6, by Language, Percent Answering “All of the Time” or “Most of the Time”

	WTCHR (N=1,400)		NSDUH (n=2,605)	
	English	Spanish	English	Spanish
Worthlessness	8.7	18.5	8.7	9.8
Sadness	18.3	32.7	13.8	21.8
Nervousness	19.6	32.1	9.5	11.3
Hopelessness	11.5	25.6	8.7	7.3
Restlessness	21.7	33.3	7.2	8.7
Effort	19.7	28.6	12.8	15.0

Table 3. Hispanic Non-Response to the K-6, by Language, Percent Answering “Don’t Know” or “Refused”

	WTCHR (N=1,400)		NSDUH (n=2,605)	
	English	Spanish	English	Spanish
Worthlessness	0.5	5.4	2.6	1.3
Sadness	0.2	2.4	2.1	2.4
Nervousness	0.0	1.2	3.2	0.9
Hopelessness	0.4	7.7	3.3	1.4
Restlessness	0.6	0.6	3.2	1.4
Effort	1.8	7.7	3.5	1.6

Table 4. Prevalence of Psychological Distress as Measured by K-6, by Group

Group	WTCHR	NSDUH
Non-Hispanics	8.0	9.6
Hispanics	20.0	10.2
English-speaking Hispanics	18.1	10.6
Spanish-speaking Hispanics	35.7	9.8
Low Income*	20.0	15.4
High Income	5.4	8.1
Under age 50	8.2	11.0
Over age 50	11.7	7.6

*NSDUH income was coded “low” when under \$20,000/year. WTCHR income was coded “low” when under \$25,000/year.

Table 5. Odds Ratios for Reported Psychological Distress

WTCHR – a telephone survey enrolling a targeted population affected by 9/11. Reference period – past 30 days.

- Hispanic = 1.9***
- Spanish language = 1.5*
- Low Income = 3.7***
- Under age 50 = 0.8*

***Statistically significant at $p < .001$

*Statistically significant at $p < .05$

NSDUH – a probability-based population survey administered in households. Reference period – worst month past 12 months.

- Hispanic = 1.0
- Spanish language = 0.8
- Low Income = 2.3***
- Under age 50 = 1.5***