

Extreme Response Style (ERS) among U.S. Hispanics vs. Non-Hispanics: An Experiment of Rating Scale Usage and Implications for Survey Design and Analysis

Martin Cerda and Ilgin Basar - Encuesta, Inc.

Background

It is well documented that U.S. Hispanics have a tendency to give extreme rating scale responses more so than white non-Hispanics.*

However, there is no consensus as to how to best measure extreme response style (ERS) among Latinos and control for its effects in the data analysis and interpretation.

Furthermore, there are a small number of documented cases and academic papers written on the subject, and the few related to U.S. Hispanics are very limited or flawed from a study design perspective.

To shed new light, Encuesta, Inc. carried out a small scale experiment within a nationally representative study conducted for its *pro bono* non-partisan American Poll series. The study was conducted in such a manner as to obtain proper representation of all U.S. Hispanics by language usage and acculturation level.

The findings will provide evidence of extreme response style and highlight differences and possible solutions when conducting research among Hispanics.

* Gerardo Marin, Raymond J. Gamba, Barbara V. Main, "Extreme Response Style and Acquiescence among Hispanics"

Objectives

Given these common survey research challenges and a review of findings from numerous sources, a study was designed that explored topics related to:

Measuring ERS using various methods,

Determining how Hispanic and non-Hispanic subjects respond to a rating scale, specifically how they map their somewhat "elastic" individually held subjective categories onto the response categories provided,

Demonstrating the use of the ipsative process to reduce response set bias* and illustrate the rescaling technique in comparing Hispanic and non-Hispanic results,

Providing the groundwork for more robust experiments that take into account more sophisticated means of measuring and correcting for ERS such as Greenleaf's ERS correction factor^, correlation analysis, or cultural norms, and

Stimulating thought on how to minimize ERS, acquiescence response style (ARS) and other "response set effects" by using proper question and rating scale design and the effective presentation and rotation of rating scales to minimize unwanted effects.

* William H. Cunningham, Isabella C.M. Cunningham, and Robert T. Green, "The Ipsative Process to Reduce Response Set Bias"

^ Eric A. Greenleaf, "Measuring Extreme Response Style"

Methodology

A nationally representative telephone survey was conducted with n=358 U.S. Hispanic and n=302 non-Hispanic (all other races and ethnicities in a representative mix) adults.

Field dates: November 6-27, 2009

A proprietary random probability landline (RDD and listed Hispanic last name, n=443) and cell phone (RDD, n=217) frame sampling approach was used.

Interviews were conducted in English or Spanish, according to respondent preference.

A total of 56% of all interviews conducted among Hispanics were in Spanish and 44% were in English.

Almost half (46%) of Hispanics interviewed were Spanish-Dominant, 39% were Bilingual, and 15% were English-Dominant according to the Encuesta, Inc./Marin Acculturation Scale.*

* The scale is based on a series of questions that determine language usage in different situations.

Description of the Experiment

The interviewing process had two phases, testing and retesting of six experimental "bedrock" questions (i.e. opinions should not shift over a brief period of time).

Each question represented a different construct (e.g. importance of family) with a unique single item rating scale used for each question (i.e. a mix construct-specific and commonly used response options).

For the test phase, questions related to several topics were asked with a mixture of 5-point Likert (with each point anchored plus a don't know/refused category which was not read) and 10-point end-point only anchored scales. This approach* would help reduce "response set effects" as the use of same response format was avoided.

An example of a test (Q.E) and retest question (Q.1) is presented below (see Appendix A for the exact questions broken out by cell and phase).

(READ: Now I am going to ask you some questions regarding different aspects of your life.)

(SCREENER) (ROTATE Q.E-Q.H)

E. How important is having a close relationship with your family? Would you say it is...? (ACCEPT ONE RESPONSE)

(ROTATE SERIES A & B)

(SCREENER) (ROTATE Q.1-Q.4)

1. How important is having a close relationship with your family? Using a scale from 1 to 10, where 1 means "not at all important" and 10 means "extremely important", please tell me how important having a close relationship with your family is.

1=NOT AT ALL IMPORTANT

10=EXTREMELY IMPORTANT

99= (DON'T KNOW/REFUSED)

(EXTREMELY IMPORTANT)

(VERY IMPORTANT)

(SOMEWHAT IMPORTANT)

(NOT VERY IMPORTANT)

(NOT AT ALL IMPORTANT)

(DON'T KNOW/REFUSED)

In the retest phase, half of the subjects were asked the six "bedrock" questions except with a modified question and different rating scale. Specifically, some of the retest questions were converted to 5-point and some to 10-point scales. The other half of the subjects were retested to assess the reliability of the six "bedrock" questions using the exact scales from the test phase.

* C. Harry Hui, Harry C. Triandis, "The Instability of Response Sets"

Sample Cell and Question Design

	Cell Type			
	Control		Experimental	
	Test (A)	Retest (B)	Test (C)	Retest (D)
FAMILY	5-PT	5-PT	5-PT	10-PT
BREAKFAST	5-PT	5-PT	5-PT	10-PT
UNIVERSE	5-PT	5-PT	5-PT	10-PT
PETS	5-PT	5-PT	5-PT	10-PT
HEALTH	5-PT	5-PT	5-PT	10-PT
FINANCIAL	10-PT	10-PT	10-PT	5-PT

	Total Sample	
	Hispanics	Non-Hispanics
Base: Total Respondents	n=358	n=302
Base: Control Cell Only	n=100	n=100
Base: Experimental Cell Only	n=146	n=102

Orange measures are based on 5-point rating scale data and green on 10-point rating scale data

A sub-sample of the total sample used for the study was randomly divided into two key cells, a control cell and an experimental cell, further divided by Hispanic and non-Hispanic subjects.

Both cells were administered the exact same six questions and ratings scale formats as part of a test phase during the course of the full-length interview.

Within three days of the initial interview, both cells were re-contacted and a brief interview was administered.

The control cell was given the exact same questions used as part of the test phase. The experimental cell was given a variation of the original question which included the use of a different rating scale.

Rating Scale Descriptives and Correlation Analysis

	Control				Experimental			
	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic
	Test (A)	Retest (B)	Test (C)	Retest (D)	Test (A)	Retest (B)	Test (C)	Retest (D)
	n=100	n=100	n=100	n=100	n=146	n=146	n=102	n=102
MEAN (RAW)								
FAMILY	4.70	4.54	4.58	4.75B	4.77	9.88D	4.73	9.72
BREAKFAST	4.13C	4.12	3.86	3.96	4.24C	8.87D	3.79	7.71
UNIVERSE	3.04	3.17	3.28	3.46	3.03	5.97	3.57A	6.40
PETS	4.29	4.04	4.44	4.55B	4.09	8.01	4.52A	8.21
HEALTH	3.50	3.42	3.68	3.61	3.08	7.69D	3.55A	7.26
FINANCIAL	5.62	5.60	5.85	5.92	4.90	2.24	5.40A	2.70B
STANDARD DEVIATION								
FAMILY	0.50	0.76	0.59	0.46	0.42	0.43	0.53	1.02
BREAKFAST	0.91	0.90	1.10	1.00	0.91	2.28	1.11	2.72
UNIVERSE	1.19	1.22	1.24	1.21	1.12	3.28	1.19	3.35
PETS	0.96	1.13	0.73	0.79	1.18	2.56	0.93	2.08
HEALTH	0.96	0.93	0.98	1.03	1.13	2.12	1.12	2.08
FINANCIAL	2.06	2.12	2.00	1.78	2.28	0.88	1.81	1.00
MEAN POINT SHIFT								
FAMILY	→ -0.16*	→ -0.17**	→ -0.17**	→ -0.17**	→ -0.17**	→ -0.17**	→ -0.17**	→ -0.17**
BREAKFAST	→ -0.01	→ -0.10	→ -0.10	→ -0.10	→ -0.10	→ -0.10	→ -0.10	→ -0.10
UNIVERSE	→ +0.13	→ +0.18	→ +0.18	→ +0.18	→ +0.18	→ +0.18	→ +0.18	→ +0.18
PETS	→ -0.25**	→ -0.11	→ -0.11	→ -0.11	→ -0.11	→ -0.11	→ -0.11	→ -0.11
HEALTH	→ -0.08	→ -0.07	→ -0.07	→ -0.07	→ -0.07	→ -0.07	→ -0.07	→ -0.07
FINANCIAL	→ -0.02	→ +0.07	→ +0.07	→ +0.07	→ +0.07	→ +0.07	→ +0.07	→ +0.07
CORRELATION (T vs. RT)								
FAMILY	→ .191*	→ .163*	→ .163*	→ .163*	→ .082	→ .276***	→ .276***	→ .276***
BREAKFAST	→ .388***	→ .433***	→ .433***	→ .433***	→ .466***	→ .590***	→ .590***	→ .590***
UNIVERSE	→ .545***	→ .507***	→ .507***	→ .507***	→ .557***	→ .650***	→ .650***	→ .650***
PETS	→ .542***	→ .416***	→ .416***	→ .416***	→ .615***	→ .481***	→ .481***	→ .481***
HEALTH	→ .451***	→ .740***	→ .740***	→ .740***	→ .570***	→ .601***	→ .601***	→ .601***
FINANCIAL	→ .545***	→ .501***	→ .501***	→ .501***	→ .367***	→ .471***	→ .471***	→ .471***

	Control				Experimental			
	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic
	Test (A)	Retest (B)	Test (C)	Retest (D)	Test (A)	Retest (B)	Test (C)	Retest (D)
	n=100	n=100	n=100	n=100	n=146	n=146	n=102	n=102
MEAN (RAW)								
FAMILY	4.70	4.54	4.58	4.75B	4.77	9.88D	4.73	9.72
BREAKFAST	4.13C	4.12	3.86	3.96	4.24C	8.87D	3.79	7.71
UNIVERSE	3.04	3.17	3.28	3.46	3.03	5.97	3.57A	6.40
PETS	4.29	4.04	4.44	4.55B	4.09	8.01	4.52A	8.21
HEALTH	3.50	3.42	3.68	3.61	3.08	7.69D	3.55A	7.26
FINANCIAL	5.62	5.60	5.85	5.92	4.90	2.24	5.40A	2.70B
STANDARD DEVIATION								
FAMILY	0.50	0.76	0.59	0.46	0.42	0.43	0.53	1.02
BREAKFAST	0.91	0.90	1.10	1.00	0.91	2.28	1.11	2.72
UNIVERSE	1.19	1.22	1.24	1.21	1.12	3.28	1.19	3.35
PETS	0.96	1.13	0.73	0.79	1.18	2.56	0.93	2.08
HEALTH	0.96	0.93	0.98	1.03	1.13	2.12	1.12	2.08
FINANCIAL	2.06	2.12	2.00	1.78	2.28	0.88	1.81	1.00

	Control				Experimental			
	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic
	Test (A)	Retest (B)	Test (C)	Retest (D)	Test (A)	Retest (B)	Test (C)	Retest (D)
	n=100	n=100	n=100	n=100	n=146	n=146	n=102	n=102
MEAN POINT SHIFT								
FAMILY	→ -0.16*	→ -0.17**	→ -0.17**	→ -0.17**	→ -0.17**	→ -0.17**	→ -0.17**	→ -0.17**
BREAKFAST	→ -0.01	→ -0.10	→ -0.10	→ -0.10	→ -0.10	→ -0.10	→ -0.10	→ -0.10
UNIVERSE	→ +0.13	→ +0.18	→ +0.18	→ +0.18	→ +0.18	→ +0.18	→ +0.18	→ +0.18
PETS	→ -0.25**	→ -0.11	→ -0.11	→ -0.11	→ -0.11	→ -0.11	→ -0.11	→ -0.11
HEALTH	→ -0.08	→ -0.07	→ -0.07	→ -0.07	→ -0.07	→ -0.07	→ -0.07	→ -0.07
FINANCIAL	→ -0.02	→ +0.07	→ +0.07	→ +0.07	→ +0.07	→ +0.07	→ +0.07	→ +0.07

	Control				Experimental			
	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic	Hispanic	Non-Hispanic
	Test (A)	Retest (B)	Test (C)	Retest (D)	Test (A)	Retest (B)	Test (C)	Retest (D)
	n=100	n=100	n=100	n=100	n=146	n=146	n=102	n=102
CORRELATION (T vs. RT)								
FAMILY	→ .191*	→ .163*	→ .163*	→ .163*	→ .082	→ .276***	→ .276***	→ .276***
BREAKFAST	→ .388***	→ .433***	→ .433***	→ .433***	→ .466***	→ .590***	→ .590***	→ .590***
UNIVERSE	→ .545***	→ .507***	→ .507***	→ .507***	→ .557***	→ .650***	→ .650***	→ .650***
PETS	→ .542***	→ .416***	→ .416***	→ .416***	→ .615***	→ .481***	→ .481***	→ .481***
HEALTH	→ .451***	→ .740***	→ .740***	→ .740***	→ .570***	→ .601***	→ .601***	→ .601***
FINANCIAL	→ .545***	→ .501***	→ .501***	→ .501***	→ .367***	→ .471***	→ .471***	→ .471***

Letters denote significant difference at a .100 level (2-tailed)

Paired sample correlation or t-test is significant at the . * .100 ** .050 *** .001 level (2-tailed)

^ Indicates not stat tested as rating scale was changed from 5-point to 10-point scale

Every effort was made to use a random assignment procedure that would yield equivalent groups of subjects per cell.

However, in some cases the Hispanic and non-Hispanic group measures from both the control and experimental cells were not identical, likely a product of the small sample sizes used per cell.

Measuring ERS Based On Extreme End-Point Use

	Hispanic		Non-Hispanic		Hispanic		Non-Hispanic	
	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
	Test (A)	Retest (B)	Test (C)	Retest (D)	Test (E)	Retest (F)	Test (G)	Retest (H)
	n=100	n=100	n=146	n=102	n=100	n=100	n=102	n=102
SCALE TYPE (ALL FIVE CASES)	5-PT	5-PT	5-PT	10-PT	5-PT	5-PT	5-PT	10-PT
4/5 CASES	1%	2%	1%	9% ^H	2%	1%	2%	3%
5/5 OR 5/5 CASES	11%	7%	12%	30% ^H	12%	18% ^B	16%	18%

Letters denote significant difference at a .100 level (2-tailed)

Note: This table is based to five cases (not six) as this was the number of times the exact scale type was repeated by cell.

Rating Scale Reliability and Impact of Changes

A complete cross-correlation analysis of the six constructs used as part of this small scale experiment indicated few notable inter-item (i.e. or construct in this case) correlations that might have yielded an unwanted effect (see Appendix B).

An analysis of the control cell indicated similar use of the rating scales and no strong evidence of ERS among Hispanics. Furthermore, it is evident that mixing the use of 5-point and 10-point rating scales helped minimize response set bias.

The control cell was used to assess the reliability of the test rating scale series and this was favorably indicated given the high degree of correlation between the test and retest means and distributions when using the exact same question and rating scale.

The use of a 10-point scale on the experimental cell as part of the retest phase indicated a higher degree of use of the higher points of the rating scale among Hispanics compared to non-Hispanics. This was not evident in the case of the retest phase use of the 5-point scale. In short, it was apparent that Hispanics were more prone to ERS when using the 10-point scale as indicated by the higher mean point shift between the test and retest phase.