

An Analysis of the New Vocabulary Items on the 2008 General Social Survey

Tom W. Smith

NORC/University of Chicago

GSS Methodological Report No. 115

October, 2009

The General Social Survey (GSS) has regularly included a 10-item vocabulary battery since 1974 (Davis, Smith, and Marsden, 2009). This scale was based on items developed by Edward Thorndike and George Gallup in the 1940s (Malhorta, Krosnick, and Haertel, 2007). Respondents are shown 10 words and for each word a list of five possible matching words and asked to select which one comes closest in meaning (see Appendix A).

Analysis by Malhorta, Krosnick, and Haertel (2007) indicated that the 10 words tended to cluster into two groups, hard and easy words with no words of medium difficulty. Based on previous work by Thorndike and Gallup and their own analysis, they identified four words that were expected to fall in the middle range. In 2008 these four new words were administered to a random half of the 2006 panel when they were reinterviewed. The new items appeared immediately after the standard 10 words and to respondents were just a continuation of the vocabulary battery.

As Table 1 indicates these words did prove to be medium in difficulty. Their levels were also close to that predicted (Malhorta, Krosnick, and Haertel, 2007). An analysis of the psychometric properties of new items and how they relate to the existing 10 items will be carried. Table 2 shows the correlation of these new items amongst themselves and with the existing items. This preliminary comparison indicates that should be suitable for use in an augmented vocabulary battery.

Table 1

% Correct on 2008 GSS

	2006 Panel	2008 Panel (reinterview)
WORDA	79.2	77.0
WORDB	86.4	82.2
WORDC	20.5	16.5
WORDD	90.2	90.7
WORDE	74.4	74.7
WORDF	75.5	76.8
WORDG	32.0	30.0
WORDH	32.8	26.7
WORDI	71.1	70.7
WORKJ	24.3	23.6
WORDK	61.0	----
WORDL	53.9	----
WORDM	50.5	----
WORDN	50.9	----

Source: GSS2008 merged file

Table 2

Inter-item Correlations of New Vocabulary Items  
(Pearson's r)

	WORDK	WORDL	WORDM	WORDN
WORDA	.259	.227	.274	.300
WRODB	.288	.320	.204	.311
WORDC	.227	.277	.220	.189
WORDD	.321	.294	.211	.317
WORDE	.370	.320	.239	.342
WORDF	.422	.358	.354	.430
WORDG	.243	.249	.253	.273
WORDH	.275	.265	.254	.266
WORDI	.291	.299	.219	.266
WORDJ	.275	.269	.285	.361
WORDK	----	.432	.345	.333
WORDL	.432	----	.278	.385
WORDM	.345	.278	----	.355
WORDN	.333	.385	.355	----

All statistically significant at .000 level.

Source: 2006 panel of 2008 GSS

## References

Davis, James A.; Smith, Tom W.; and Marsden, Peter V., General Social Survey Cumulative Codebook, 1972-2008. Chicago: NORC, 2009.

Malhotra, Neil; Krosnick, Jon A.; and Haertel, Edward, "The Psychometric Properties of the GSS Wordsum Vocabulary Test," GSS Methodological Report No. 111. Chicago: NORC, 2007.

## Appendix A: Question Wording

We would like to know something about how people go about guessing words they do not know. On this card are listed some words – you may know some of them, and you may not know quite a few of them.

### HAND CARD

On each line the first word is in capital letters – like BEAST. Then there are five other words. Tell me the number that comes **closest** to the meaning of the word in capital letters. For example, if the word in capital letters is BEAST, you would say “4” since “animal” comes closer to BEAST than any of the other words.

If you wish, I will read the words to you. These words are difficult for almost everyone – just give me your best guess if you are not sure of the answer.

### EXAMPLE

BEAST            1. afraid            2. words            3. large            4. animal            5. separate