An Analysis of Response Patterns to the Ten-Point Scalometer

Tom W. Smith
NORC 1155 East 60th St. Chicago, Il 60637

Introduction

Survey researchers have been concerned for decades about various measurement effects associated with response scales. Among the issues that have occupied methodologists are 1) the number of categories, including a) how many response options are optimum (Miller, 1956; Green and Rao, 1970; Jacoby and Matell, 1970; Lehmann and Hulbert, 1972; Ramsey, 1973; Masters, 1974; Hulbert, 1975; Lissitz and Green, 1975; Andrews and Withey, 1976; Cox, 1980; Sheatsley, 1983; Andrews, 1984; Upshaw, 1984; Peterson, 1985; Smith and Peterson, 1985; Kidder and Judd, 1986; Alwin and Krosnick, 1991)), b) the advantages of an odd vs. even number of categories or whether there should or should not be an explicit mid-point (Masters, 1975; Schuman and Presser, 1982; Sudman and Bradburn, 1982; Sheatsley, 1983; Andrews, 1984; Schwarz and Hippler, 1991; and Alwin and Krosnick, 1992); and c) whether there should be an explicit Don't Know option (Schuman and Presser, 1981; Andrews, 1981; Smith, 1984; and Alwin and Krosnick, 1992), 2) the labelling of scale points (Dawes and Smith, 1985), including a) the use of labels for endpoints vs. all points (Andrews, 1984; Krosnick and Berent, 1990) and b) the use of numbers vs. terms to denote categories (Wildt and Mazis, 1978; Gallup, 1979; McCroskey, Prichard, and Arnold, 1968; Hensler and Stipak, 1979; and Alwin and Krosnick, 1992), and 3) response styles, including a) acquiescence bias (Rorer, 1965; Rundquist, 1966; O'Neill, 1967; Becker and Myers, 1970; Kolson and Green, 1970; Carr, 1971; Arndt and Crane, 1975; Wright, 1975; Schuman and Presser, 1981; Dillman and Tarnai, 1991) and b) extremity bias (Kidd and Judd, 1986; Dillman and Tarnai, 1991; Greenleaf, 1992).

Recently there has been renewed interest in the issue of response scales in general and in particular about how people distribute their responses amongst the offered categories (Schwarz, et al., 1991; Alwin and Krosnick, 1991; and Greenleaf, 1992). Of particular interest has been the work of Schwarz, et al. (1991) which shows that people respond to 11-point, numerical scales differently according to the numbering convention used. We wondered how people responded to questions using a ten-point scalometer as the response scale (a scale similar to the 11-point, numeric scales used by Schwarz, et al, 1991.) and whether any of the response effects noted in the literature might appear for this response format.

10-Point Scalometer

The ten-point scalometer was devised by Jan Stapel of the Netherlands Institute of Public Opinion and was first used in the United States by the Gallup Organization in March, 1953. Since then it has been used by Gallup, NORC, and other survey research organizations over 1,000 times in national surveys.

The ten-point scalometer asks people to express their like/dislike of an object on a scale ranging from +5 to -5. While the wording has varied slightly, the item typically reads as follows:

You will notice that the boxes on this card go from the highest position of "plus 5" for a [country, person, etc.] which you like very much, to the lowest position of "minus 5" for a [country, person, etc.] you dislike very much. How far up the scale or how far down the scale would you rate the following [countries, persons, etc.]?

The boxes on the card also vary somewhat, but typically show ten vertically arranged boxes labelled from +5 to +1 and then from -1 to -5. In addition, an unlisted, off-scale response of Don't Know is usually also coded.

Data

To examine how people respond to the scalometer we used two sources of data. First, we collected response distributions from a large number of uses of the scalometer. Second, we looked at the characteristics of individuals who selected the various response options. The aggregate collection of distributions was used mainly to identify scale dependent response patterns and the analysis of individual characteristics was primarily used to test hypotheses about the causes of the scale dependent response patterns.

First, to examine how people distribute their responses on the scalometer, we collected the distributions to 188 items. 78 came from 11 of NORC's General Social Surveys (GSS) conducted between 1974 and 1991 and 110 came from 9 Gallup polls carried out from 1953 to 1973. The items inquire about 26 leaders, 18 countries, 12 voluntary associations, 6 government agencies, and 5 other objects. The distributions to these 188 items are the cases in our analysis. Our goal was to locate and explain response patterns or tendencies that were related to the scalometer itself rather than the distribution of responses to the objects being evaluated. This is inherently a difficult task since we have no ready way of separating the substantive distribution of a given item from its scale dependent distribution. However, if there are general response patterns linked to the scale itself, they should turn up across multiple uses of the scale and therefore should be detectable. Second, to augment the information from the analysis of the scalometer distributions we also analyzed the individual-level data from the GSSs. The GSSs are a series of full-probability samples of adults living in households in the United States. For full technical details see Davis and Smith, 1991. The scalometer questions on the GSS asked about eight countries (Brazil, Canada, China, Egypt, England, Israel, Japan, and Russia) in 1974, 1975, 1977, 1982, 1983, and 1985 and about six countries (minus Brazil and England) in 1986 and 1988-1991. For those countries asked about in all years the sample size was 14,570 and for Brazil and England the sample size was 8,353.

Scale Dependent Response Patterns

A perusal of the average distribution across the 188 scalometer questions (Table 1) indicated several distinctive patterns that might result (at least in part) from scale dependent responses. The first was a surfeit of +1 responses. The second was an uptick in extreme responses (+5 and -5).

Choosing +1

The modal response across all scalometer questions was +1. What is distinctive is that +1's are almost twice as common -1's (15.2% vs. 7.8%). A surplus of +1's over -1's does not only show up in the averaged results, but appears for 93% of the individual distributions. While one might well have expected a bunching of responses near the middle of the scale, this pattern is distinctive because the two mid-points (+1 and -1) vary so greatly in their attraction.[1]

Furthermore, if we calculate an expected value for +1's as the average of +2's and -1's,[2] we find that +1's exceeded their expected value by an average of 6.8 percentage points. In 95% of the cases the observed number of +1's exceeded the expected value, in 1% of the cases it equaled it, and in 4% of the cases it fell below it. All of these few exceptions are cases with extreme negative skews (68-94% selecting -5) and less than the average number of Don't Knows (1-9%). Even when examining cases when +1 was not the modal category, +1's exceeded the expected by 3.7 percentage points. That is, except for cases with extreme negative skews and less than an average number of Don't Knows, +1's exceeds the expected number and usually by a large margin.

Table 2 shows that cases that have a high number of Don't Knows also tend to have responses concentrated in the middle categories in general and in +1's in particular. With Don't Knows included we see that more Don't Knows are associated with more selections of +2, +1, -1, and -2 and that the association with +1 is especially strong. (If Don't Knows were not associated with other response choices, we would expect weak and similar negative relations when Don't Knows are included, since choosing Don't Know means that one of the 10 other scale points was not chosen. With Don't Knows excluded we would expect no significant relationships with the 10 scale points.) While this could be interpreted to mean that items with high Don't Know levels were also items on which respondents tended to hold moderate positions, we believe instead that this pattern results from people using Don't Knows and middle scale position for the same purpose - to indicate non-attitudes towards the object in question. To many people who don't realize that they can volunteer an off-scale Don't Know response or who want to avoid the "embarrassment" of a Don't Know response, a middle or "neutral" response provides a place to put themselves. It is a safe haven if they do not know what they feel about the object or are not even sure who or what the object is.[3]

In addition, the edge of +1's over -1's is greater when the % giving Don't Knows is higher ($r=.46$; prob. $< .01$). We interpret this to mean that when non-attitudes are greater even a higher proportion of the +1 response are substitutes for Don't Knows rather than middle-of-the-road evaluations and that this increases the edge of +1's over -1's.

The edge of +1's over -1's is also greater the larger the proportion of the distribution is found in the middle rather than at the ends of distribution ($r=.67$; prob. $< .01$). More cases in the middle is probably a function of more non-attitudes and more non-attitudes inflates the edge of +1's over -1's.

The link between Don't Knows and middle positions is further illustrated by a large number of examples in which there are two or more "humps" in the distribution, one at +1 due to non-attitudes and one or more other peaks reflecting the "true" distribution. For example, in 1966 ratings of President Johnson peaked at +5 and -5 and again at +1 and in every year from 1974 to 1991 Japan shows peaks at +3 and +1.

The link between +1's and Don't Knows can also be shown by looking at the linkage between giving Don't Knows across questions in the GSS. Table 3D shows the mean number of Don't Knows given to 18 other questions on the GSS. There is a strong connection between giving Don't Knows on the scalometer items and the other items. Those saying Don't Know to the country questions were much more likely to have said Don't Know to other items. Don't Knows then tended to be highest for either the +5 or -5 categories for reasons that are explained below. Next, Don't Knows were usually highest for the +1 category. This suggests that respondents with responses equivalent or similar to Don't Knows tend to concentrate in the +1 category.

While people apparently select a mid-point as an alternative to giving a Don't Know response, that does not explain why +1's are heavily favored over -1's. We have two possible explanations. First, +1 is the fifth point on the scale (starting with 1 for +5). People may incorrectly believe that 5 is the mid-point on a 10-point scale. Since the typical showcard does not actually number the scale points in this manner, this effect may be minor. However, numbering may not be needed to produce this effect. A scale of social standing that consisted to 10 stacked boxes with the highest labelled "Top" and the lowest "Bottom" that was used in eight countries produced generally normal distributions with a peak in the 5th box from the top, the equivalent of the +1 point on the scalometer. In each country the proportion in box 5 substantially exceeded the number in box 6.

Moreover, in the few cases when such numbers are provided on showcards, there is some indications that +1 (5) attracted more responses than usual. When the numbers 1-10 were listed, +1 exceeded -1 by 10.3 percentage points and when the numbers 1-10 were omitted, the difference was 8.1 percentage points (difference significant at .046 level).

Second, there is a general positivity bias in American society and when a lack of a single, mid-point forces people to tilt their leaning one way or another, this societal bias may move them to select the positive over the negative point on the scale (Smith, 1979). This explanation is supported by fact that when Don't Knows are greater the edge to +1's over -1's is larger. This suggests that when people are unsure, they go with the positive mid-point rather than the negative mid-point.

In brief, +1 are disproportionately favored on a scalometer because the absence of an explicit Don't Know response encourages people to select a mid-point as an alternative to giving a Don't Know response and +1 is probably chosen over -1 because of a positivity bias.

Endpoints

Second, the endpoints (+5's and -5's) outnumber their adjoining scaler points (+4's and -4's). The edge at the positive end is fairly modest (11.9% +5's vs. 8.7% +4's) and a surplus occurs in only 60% of the cases, but at the negative pole the difference is substantial (10.9% -5's vs. 2.8% -4's) and -5's exceed -4's in 96% of the cases. Given that most attitudinal distributions find more people in towards the middle than towards the extremes, this pattern (especially at the negative pole) is noteworthy.

The attraction of the endpoints over their

adjoining categories is shown if we control for the over all distribution. When the mode is at the positive endpoint, +5 exceeds +4 by 17.4 percentage points. When the mode is at some other point, the proportions selecting +5 and +4 do not differ. When the mode is at the negative endpoint, -5 exceeds -4 by 36.6 percentage points, when the mode is at from -4 to +4, -5 is greater than -4 by 3.1 percentage points and even when the mode is at +5, -5 bests -4 by 2.1 percentage points. Thus, the proportion +5 is not less than the proportion +4 and the proportion -5 is still greater than the proportion -4 even when the mode is not at the endpoint.

The selection of the endpoints over the near endpoints may result from the fact that of the 10-points on the visual scale only +5 and -5 are mentioned to respondents. This emphasis may focus attention on these endpoints and draw responses away from the unmentioned +4 and -4 categories. The mentioned +5/-5 categories may attract respondents who do not understand the numerical scale or who do not want to make the effort to come up with a fine-grain expression of their attitude. For some rather that representing extreme liking and disliking +5 and -5 may instead represent only general liking or disliking.

To examine whether the + and - endpoints might attract those less cognitively engaged respondents who might misunderstand the numerical scale, fixate on the verbally mentioned categories, or be unwilling or unable to differentiate their attitude among the + and - categories, we looked at respondents' mean years of schooling, mean vocabulary score, comprehension rating by interviewers, and mean number of Don't Knows given to unrelated questions. Of the 32 comparisons between +5's and +4's, +5's scored cognitively lower than +4's (i.e. less educated, lower vocabulary score, lower comprehension, and more DKs) in 30 instances. There were no differences in the remaining two instances. At the negative end, -5's scored lower cognitively than -4's in 27 of 32 instances. There were 5 reversal, all involving categories with very small numbers of respondents (see Canada and England in Table 3). However, in a number of these 64 comparisons there appears to be a general association between middling liking scores and higher cognitive scores (e.g. see years of schooling for England in Table 3) or another underlying pattern that prevents one from establishing that the lower scores at the endpoints are a result of a methodological effect. But in about 16 comparisons there is no general relationship between cognitive level and liking countries that could explain the patterns near the endpoints. For example, years of schooling and liking Canada are positively associated and this monatomic relationship only reverses between the +4's and +5's. Similarly, mentioning DKs and liking Japan shows little overall association, but the mean number of DKs rises sharply from -4's to -5's. In these 16 cases cognitive levels are always lower at the endpoints. The endpoints thus seems to be particularly favored by those with low cognitive levels.

We also looked at the association between cognitive ability and the selecting of endpoints on the eight GSS country questions by examining how many times people selected +5 or -5. Giving more +5 responses was not related to education or verbal ability and was weakly associated with better comprehension (r=-.035; prob. < .05). However, giving more -5 responses was related to less education (r=-.153; prob. < .01), lower verbal ability (r=-.134; prob. < .01); and less comprehension (r=.062; prob. < .01). These associations are at least in part substantive since isolationists tended to have lower cognitive abilities. We attempted to take this into consideration by controlling for internationalism. Internationalism was measured by a three-item scale measuring support for the United States taking an active role in world affairs (USINTL), wanting the United States to belong to the UN (USUN), and favoring foreign aid (NATAID). Since internationalism is related to liking countries on an additive scale of the eight countries (r=.17; prob. > .01) and to higher cognitive ability (r's=.17 to .22 with education, verbal ability, and comprehension), controlling for internationalism should help to eliminate a substantive association between liking countries and higher cognitive ability. In fact, it has little impact. With internationalism controlled for mentioning more +5's becomes marginally related to less education (r=-.025; prob.=.04) and is not related to verbal ability or comprehension. Mentioning more -5's remains related to less education (r=.125; prob. > .01), lower verbal ability (r=-.096; prob. > .01), and less comprehension (r=.066; prob. > .01). This indicates that attitudes towards isolationism/ internationalism do not explain the association between extremity and cognitive ability.

Other evidence for a methodological explanation for the distributions at and near the endpoints comes from comparing the inter-relationship amongst the eight GSS country items. All of the 28 inter-item Pearson correlations are positive and average .19. Given this overall positive association between countries, we would expect that selection of +5 and -5 responses would be negatively related. When we made two scale that counted the number of +5 and -5 responses to these eight items, we found that they were weakly positively associated (r=.05; prob. > .01). This probably result from people tending to give opposing endpoints because of an extremity bias.

It is uncertain why the extreme effect is greater at the negative end than at the positive end. It is possible that people may be inclined to make more precise judgments about degree of liking than they do about disliking. People oriented towards the positive end may make more of an effort to distinguish between the five positive response options, while those leaning towards the negative end may be prone to select -5 as a convenient point for merely expressing dislike. The substantive association of education with disliking countries may also reenforce the extremity effect at the negative pole.

Summary and Conclusion

There is reason to believe that at least two response effects influence how people answer the widely used scalometer. First, respondents with non-attitudes are attracted to the +1 category as a substitute for a DK response. Both they and those with true middle-of-the-road attitudes are drawn to +1 (rather than -1) as the preferred mid-point.

Second, endpoints in general and the negative endpoint in particular disproportionately attracts responses. Those with less cognitive ability are prone to select the endpoints. This is probably because they are unable or unwilling to fully utilize the 10-point scale.

Such response effects might be reduced if the scalometer was revised. First, providing a clear mid-point (e.g. making it an 11-point scale with a 0 response in the middle) would give the ambivalents a clear category in which to place

themselves and would reduce the problem created by people who apparently seek such a mid-point either by randomly choosing between +1 and -1, or, for the reasons outlined above, favoring +1 as the "mid-point." However, including a mid-point may draw in additional non-attitude holders who probably more appropriately belong in Don't Know. This tendency might be countered by adding an explicit DK option either on the showcard and/or explicitly mentioning a DK option in the question wording (Schuman and Presser, 1981; Smith, 1984).[4]

Second, verbally mentioning all points on the scale and possibly pointing to all 10 (or 11) points on the scale and explaining at greater length that a larger positive or negative number means more liking or disliking might reduce the over-selection of the extreme endpoints.

With such changes the scalometer should cause fewer response effects and collect ratings that are more accurate reflections of true scores. Appropriate split-ballot experiments could test whether such changes actual improve measurement.

Endnotes

1. Since there are an even number of scale points, there is no single mid-point. Respondents wishing to place themselves at the mid-point (5.5 on a scale from 1 to 10) must settle for either +1 (point 5) or -1 (point 6).

2. This is approximately what we would expect with most distributions with the mode being other that +1.

3. On the uses of middle positions and Don't Knows to handle non-attitudes see Smith, 1984.

4.On the 1990 and 1991 GSS experiments were conducted to see if offering a Don't Know response option on the showcard with the 10-point scale would increase Don't Know responses. In each case listing Don't Know on the showcard increased these responses. While the increase was statistically significantly only for Egypt, it was significant overall across the six counties. Likewise when we compare the average % say Don't Know across whether a DK option was listed on the scale or not, we find that listing is associated with more Don't Knows being mentioned (r=.16; prob. > .05).

References

Alwin, Duane F. and Krosnick, Jon A., "The Reliability of Survey Attitude Measurement," Sociological Methods and Research, 20 (August, 1991), 139-181.

Andrews, Frank M., "Construct Validity and Error Components of Survey Measures: A Structural Modelling Approach," Public Opinion Quarterly, 48 (Summer, 1984), 409-442.

Andrews, Frank M. and Withey, Stephen, Social Indicators of Well-Being: Americans' Perceptions of Life Quality. New York: Plenum Press, 1976.

Arndt, Johan and Crane, Edgar, "Response Bias, Yea-Saying, and the Double Negative," Journal of Marketing Research, 12 (May, 1975), 218-220.

Becker, Boris W. and Myers, John G., "Yeasaying Response Style," Journal of Advertising Research, 10 (Dec., 1970), 31-37.

Carr, Leslie G., "The Srole Items and Acquiescence," American Sociological Review, 36 (April, 1971), 287-293.

Cox, Eli P. III, "The Optimal Number of Response Alternative for a Scale: A Review," Journal of Marketing Research, 17 (November, 1980), 407-422.

Davis, James A. and Smith, Tom W., General Social Surveys, 1972-1991: Cumulative Codebook. Chicago: NORC, 1991.

Dawes, Robyn M. and Smith, Tom L., "Attitude and Opinion Measurement," in Handbook of Social Psychology, edited by Gardner Lindzey and Elliot Aronson. 3rd edition. New York: Random House, 1985.

Dillman, Don A. and Tarnai, John, "Mode Effects of Cognitively Designed Recall Questions: A Comparison of Answers to Telephone and Mail Surveys," in Measurement Error in Surveys, edited by Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman. New York: John Wiley & Sons, 1991.

Greenleaf, Eric A., "Measuring Extreme Response Style," Public Opinion Quarterly, 56 (Fall, 1992), 328-351.

Hippler, Hans-Juergen and Schwarz, Norbert, "Response Alternatives: The Impact of Their Choice and Presentation Order," in Measurement Error in Surveys, edited by Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman. New York: John Wiley & Sons, 1991.

Hulbert, James, "Information Processing Capacity and Attitude Measurement," Journal of Marketing Research, 12 (February, 1975), 104-106.

Kidder, Louise H. and Judd, Charles M., Research Methods in Social Relations. 5th edition. New York: Holt, Rinehart, and Winston, 1986.

Kolson, Kenneth L. and Green, Justin J., "Response Style and Political Socialization Research," Social Science Quarterly, 51 (December, 1970), 527-538.

Lissitz, Robert W. and Green, Samuel B., "Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach," Journal of Applied Psychology, 60 (1975), 10-13.

Masters, James R., "The Relationship between Number of Response Categories and Reliability of Likert-Type Questionnaires," Journal of Educational Measurement, 11 (Spring, 1974), 49-53.

Peterson, Bruce L., "Confidence: Categories and Confusion," GSS Methodological Report No. 31. Chicago: NORC, 1985.

Rorer, Leonard G., "The Great Response-Style Myth," Psychological Bulletin, 63 (March, 1965), 129-156.

Rundquist, Edward A., "Item and Response Characteristics in Attitude and Personality Measurement: A Reaction to L.G. Rorer's "The Great Response-Style Myth," Psychological Bulletin, 66 (Sept., 1966), 166-177.

Schuman, Howard and Presser, Stanley, Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context.

Schwarz, Norbert, et al., "Rating Scales: Numerical Values May Change the Meaning of Scale Labels," Public Opinion Quarterly, 55 (Winter, 1991), 570-582.

Sheatsley, Paul B., "Questionnaire Construction and Item Writing," in Handbook of Survey Research, edited by Peter H. Rossi, James D. Wright, and Andy B. Anderson. New York: Academic Press, 1983.

Smith, Tom W., "Happiness: Time Trends, Seasonal Variations, Inter-Survey Differences, and Other Mysteries," Social Psychology Quarterly, 42 (Spring, 1979), 18-30.

Smith, Tom W., "Nonattitudes: A Review and Evaluation," in Surveying Subjective Phenomena, edited by Charles F. Turner and Elizabeth Martin. New York: Russell Sage, 1985.

Smith, Tom W. and Peterson, Bruce L., "The Impact of the Number of Response Categories on Inter-Item Associations: Experimental and Simulated Results," Paper presented to the American Sociological Association, Washington, D.C., August, 1985.

Sudman, Seymour and Bradburn, Norman M., Asking Questions: A Practical Guide to Questionnaire Design. San Francisco: Jossey-Bass, 1983.

Upshaw, Harry S., "Output Processes in Judgment," in Handbook of Social Cognition, edited by Robert S. Wyer, Jr. and Thomas K. Srull. Volume 3. Hillsdale, New Jersey: Lawrence Erlbaum, 1984.

Wildt, Albert R. and Mazis, Michael B., "Determinants of Scale Response: Labels Versus Position," Journal of Marketing Research, 15 (May, 1978), 261-267.

Wright, James D., "Does Acquiescence Bias the 'Index of Political Efficacy'?" Public Opinion Quarterly, 39 (Summer, 1975), 219-226.

Table 1

Distribution of Responses to 188 Scalometer Questions

	Don't Knows Included	Don't Knows Excluded
+5	11.9%	12.9%
+4	8.7	9.5
+3	13.3	14.8
+2	10.9	12.2
+1	15.2	17.3
-1	7.8	8.8
-2	4.2	4.7
-3	4.4	4.9
-4	2.8	3.1
-5	10.9	11.9
Don't Knows	10.0	----
	(188)	(188)

Table 2

Relation between % Saying "Don't Know" and % Giving Response on the 10-Point Scale

(Pearson's r)

	+5	+4	+3	+2	+1	-1	-2	-3	-4	-5
Don't Knows Included	-.350	-.352	-.287	+.029	+.313	+.182	+.094	-.041	-.053	-.140
Don't Knows Excluded	-.312	-.282	-.125	+.285	+.567	+.394	+.274	+.103	+.056	+.107

n=188

Table 3
Levels of Education, Verbal Ability, Interviewer's Rating of Comprehension and Don't Knows by Response to Scalometer Questions

A. Mean Years of Schooling

Responses	Egypt	Israel	China	Russia	Japan	England	Brazil	Canada
+5	11.6	12.2	12.0	12.3	12.5	12.1	11.2	12.4
+4	12.2	12.5	12.4	12.5	13.0	12.4	11.8	12.7
+3	12.5	12.7	12.6	12.7	12.9	12.3	12.1	12.4
+2	12.7	12.8	12.8	13.1	12.7	12.2	12.3	12.1
+1	12.8	12.6	12.7	12.9	12.4	12.0	12.6	11.7
-1	12.7	12.5	12.4	12.6	12.0	11.5	12.5	11.6
-2	12.4	12.2	12.4	12.7	11.8	10.9	12.5	11.0
-3	12.2	12.1	12.1	12.6	11.7	11.2	12.0	10.5
-4	11.8	11.9	12.1	12.2	11.1	10.1	12.0	10.3
-5	11.5	11.5	11.5	11.4	10.7	10.4	10.7	9.8
DK	10.3	10.1	10.1	10.3	9.9	9.3	10.3	9.8

B. Mean Number of Words Known[a]

Responses	Egypt	Israel	China	Russia	Japan	England	Brazil	Canada
+5	5.2	5.9	5.7	5.9	5.6	5.9	5.4	6.1
+4	5.9	6.2	5.8	6.0	6.3	6.1	5.4	6.3
+3	6.0	6.4	6.0	6.1	6.2	6.2	6.0	6.0
+2	6.1	6.0	6.0	6.3	6.2	5.9	6.2	5.7
+1	6.2	6.1	6.1	6.2	6.0	6.0	6.3	5.6
-1	6.3	6.0	6.2	6.1	6.0	5.7	6.3	5.7
-2	6.2	6.1	6.1	6.2	5.0	5.7	6.0	4.9
-3	6.0	5.7	6.1	6.2	5.9	5.4	5.9	4.1
-4	5.9	5.7	6.1	5.1	5.8	5.0	6.0	4.0
-5	5.3	5.2	5.6	5.4	5.3	5.6	5.4	4.3
DK	5.1	5.0	5.1	5.0	4.8	4.8	5.4	4.9

C. Mean Interviewer Rating of Comprehension[b]

Responses	Egypt	Israel	China	Russia	Japan	England	Brazil	Canada
+5	1.27	1.22	1.27	1.34	1.25	1.19	1.31	1.19
+4	1.22	1.18	1.26	1.20	1.15	1.19	1.24	1.17
+3	1.20	1.16	1.18	1.16	1.14	1.19	1.18	1.20
+2	1.19	1.18	1.18	1.11	1.16	1.17	1.19	1.27
+1	1.15	1.16	1.16	1.16	1.19	1.24	1.16	1.29
-1	1.15	1.19	1.19	1.21	1.25	1.21	1.15	1.33
-2	1.21	1.23	1.20	1.21	1.24	1.30	1.17	1.54
-3	1.21	1.23	1.21	1.17	1.23	1.30	1.27	1.43
-4	1.23	1.23	1.19	1.23	1.35	1.36	1.19	1.58
-5	1.31	1.23	1.29	1.27	1.37	1.28	1.28	1.53
DK	1.58	1.62	1.62	1.65	1.69	1.72	1.49	1.72

D. Mean Number of DK Responses[c]

Responses	Egypt	Israel	China	Russia	Japan	England	Brazil	Canada
+5	0.39	0.35	0.42	0.50	0.36	0.34	0.38	0.34
+4	0.34	0.31	0.35	0.31	0.30	0.29	0.38	0.33
+3	0.30	0.35	0.30	0.31	0.32	0.34	0.31	0.35
+2	0.31	0.31	0.35	0.32	0.32	0.36	0.33	0.42
+1	0.34	0.35	0.36	0.39	0.40	0.40	0.29	0.51
-1	0.32	0.36	0.35	0.35	0.38	0.27	0.29	0.48
-2	0.27	0.31	0.30	0.34	0.35	0.32	0.24	0.71
-3	0.32	0.31	0.33	0.33	0.36	0.38	0.26	0.34
-4	0.33	0.31	0.28	0.35	0.34	0.32	0.24	0.33
-5	0.39	0.39	0.36	0.36	0.41	0.46	0.50	0.63
DK	1.32	1.42	1.43	1.52	1.61	1.57	1.09	1.66

[a]Score on ten item vocabulary test (WORDSUM).
[b]Comprehension rated as 1=good,2=fair,3=poor (COMPREND).
[c]Count of number of Don't Know response to CAPPUN, COURT, NATFARE, NATFAREY, SPKATH, COLCOM, LIBRAC, EQWLTH, HELPSICK, NATENVIR NATENVIY, USINTL, ABANY, PORNLAW, POSTLIFE, RACMAR, RACOPEN, and USUN.