*Original Article*

# Comprehension Context and Sponsor Effects in a Hospital Mental Health Study

## C. Ben Gibson[1] and Timothy B. Mayhall[2]

## Abstract

Although a wealth of literature exists studying the effect of sponsor characteristics on self-reports of mental health, little work assesses a related but potentially powerful effect: a context comprehension effect, that is, a change in the respondent's interpretation of a survey question, given the concept elicited by the interviewer. Further, most studies of sponsor effects assess mean differences in responses according to sponsor type; no work has studied how these effects can affect the reliability of a psychometric instrument, which are of general importance to sociological studies of health. Here, using a study of 169 recently discharged hospital patients, we find that both effects are likely to have affected responses and, together, negated the reliability of a psychometric instrument. Future studies should consider the effect of not only sponsor characteristics upon responses but also consider how questions might be interpreted given in an interview context.

## Keywords

[1] University of California, Irvine, Irvine, CA, USA
[2] Southest Alabama Medical Center, AL, USA

**Corresponding Author:**
C. Ben Gibson, University of California, Irvine, 4100 Calit2 Building, Irvine, CA 92617, USA.
Email: cbgibson@uci.edu

Survey respondents' attitudes toward a survey sponsor—the institution or organization responsible for conducting the survey—have been shown to affect responses (Edwards, Dillman, and Smyth 2014). For example, sponsors such as the government and universities have shown to have higher response rates (e.g., Boulianne, Klofstad, and Basson 2010; Doob, Freedman, and Carlsmith 1973; Fox, Crask, and Kim 1988). Another important effect, related to but distinct from sponsor effects, is the effect of context comprehension upon responses (Sudman, Bradburn, and Schwarz 1996). While sponsor effects primarily reveal the presence of social desirability bias, the interpretation of interview questions by the respondent according to the context in which the survey (and question) takes place is also important. Respondents typically interpret survey questions by assessing the intention of the interviewer. Assuming that the interviewer is not asking meaningless or unimportant questions, the respondent relies on available concepts to form meaning from questions. This study explores comprehension effects by introducing multiple competing concepts for the interpretation of ambiguous survey questions via altering the religiosity of the interviewer.

Although most studies of this sort (rightfully) compare means, we propose that comparing measures of variation is also important. Here, we test the sponsor effects and context comprehension hypotheses using both a comparison of means and comparison of variances, as well as factor analytic methods. Our intention, in addition to testing these hypotheses, is to provide a straightforward framework for researchers interested in verifying the clarity of questions when constructing a survey.

## Interview Effects in Survey Research

### The Challenge of Parsing Religious-themed Sponsorship Effects

To our knowledge, no research has systematically observed the effect of a religious sponsor on survey responses. However, some literature on interviewer effects has observed religiosity of the interviewer upon responses. Blaydes and Gillum (2013) examined differences in survey responses in an Egyptian sample when the interviewer was wearing a *hijab* versus not. Muslims reported more piety when asked in the former case, while Christians reported less piety. They suggest that social desirability had distinct effects on Muslims and Christians—Muslim respondents may have desired to appear more pious to someone they considered to be pious, while Christian respondents may have desired to appear less at odds with the dominant religion in the area.

Although distinct from sponsor effects, the interviewer effects found by Blaydes and Gillum suggest possible complications in studying sponsor effect on responses. When religiosity of the interviewer was experimentally altered by Blaydes and Gillus in an interview context, differences in the personal religious affinity of the respondent invoked varied responses. In many cases, invoking a strong symbol of religiosity may result in varied responses by respondents who are more religious, resulting in higher variance on that question than if no religious symbol is invoked. Although most interview effect studies focus on mean differences in respondents, a study focusing on the religiosity of the interviewer or sponsor may benefit from testing differences in the extent to which responses vary, rather than differences in the mean response. Here, we test a theoretical distinction between a similar problem, sponsor effects, that would affect the average response versus one that affects the variation of the response. The implications of such a test could be applied to any systematic bias in survey responses due to sponsor, interviewer, or other effect.

## Measuring Sponsor- or Interviewer-related Sources of Distortion

Although most studies focus on mean differences, between-group differences in variation may also be informative, especially when the average sponsor (or other) effect does not have a clear theoretical direction for all respondents. Indeed, the examination of variation can be particularly informative in survey research in general (Heise 2010). Of particular concern is the reliability of a psychometric instrument that contains several items. In this context, several sources of variation may pose a problem for the measurement and comparison of an underlying concept. In this section, we describe a set of measures that sponsor characteristics could affect and the consequences for each effect.

*Mean differences, full instrument.* As what has been shown in the literature, a sponsor could distort an entire psychometric instrument, where an average response will be significantly different when altering the stated sponsor (Norton et al. 2013). An underlying concept could still be measured using the instrument; indeed, only the interpretation of the measurement would change (e.g., the survey respondent reports more depressive symptoms when interviewed by a psychiatric organization, which affect the full instrument measuring depression). In a similar case involving interviewer effects, where racist attitudes are moderated by heterophily between the interviewer and respondent, the interpretation would be shifted from an ostensibly "objective" measurement of racist attitudes versus one where the respondent is being observed by someone of a different race.

*Between*-item variation. In a psychometric instrument, a variety of questions attempt to measure an underlying concept, such as depression. To assess the reliability of an instrument, it is conventional to apply factor analysis to the responses to ensure that similar questions generate similar responses from the same respondent (e.g., Norton et al. 2013). Expanding upon the depression example, the idea is that respondents are likely to indicate depression on one question are likely to answer similarly on another question.

A sponsor effect could distort mean responses for some instrument items but not others. For example, an instrument on health behaviors administered by a medical institution could see distortion in fast-food intake but not how much time per day the respondent sits down at a desk. Although both affect health, a question on fast-food intake could be more socially undesirable. A systematic distortion in one instrument item but not another would produce an inflation in *between-item variation*, possibly reducing the ability to explain the data using a set of clear factors.

*Between*-respondent variation. A characteristic of a sponsor could also affect some respondents but not others. Depending upon the characteristics of the respondents, for example, a religious sponsor could elicit heterogeneous responses. In this case, the between-respondent variation for the instrument should vary between groups.

## Summary

Sponsor effects can be easily parsed when there is a clear theoretical direction of the effect but becomes more complicated with the effect may be varied by respondent (Krosnick 1991; Krosnick et al. 2002). Beyond observing differences in means, it is useful to examine the difference in variances as well, which is not often undertaken. Below, we review another possible source of error in survey research that could produce a lack of reliability: comprehension context.

# Comprehension Context in Survey Research

When a respondent responds to a survey, they tend to follow conversational norms to interpret the meaning of questions asked. Grice, Cole, and Morgan (1975) described the conversational norm as maximizing four dimensions: quality (veracity), relevance, informativeness, and concision. Respondents interpret survey questions as representing some degree of all these dimensions and will by default assume that meaning is in every question, even if the

questions are obscure or meaningless (Bishop, Oldendick, and Tuckfarber 1986; Schuman and Presser, 1981).

In order to provide a useful response, respondents will impose meaning upon a question by making assumptions about what the interviewer is asking. This process involves retrieving concepts immediately accessible in memory, particularly if a question is ambiguous, that is, unclear or with double meanings (Sudman et al. 1996). Respondents were shown to impose meaning on ambiguous questions by interpreting the question in terms of the preceding ones.

When assessing the reliability of a psychometric instrument, the error from question ambiguity could be attenuated by the presence of other, clearer questions asked prior to the ambiguous question. If an instrument measuring depression has 10 questions, and the 10th is ambiguous, the respondent would likely have inferred that the interviewer is intending on measuring sadness or depression and will answer the ambiguous question accordingly. However, the presence of competing concepts could exacerbate the problem. Competing concepts that alter the meaning of a question could increase the error in conveying the intended meaning of the question, rendering it difficult to correct. With no clear indication of which concept the respondent retrieved to interpret the question, the answers to that question could be useless for any substantive analysis.

This is problematic for a psychometric measurement, as at least one concept is always being elicited by the very nature of such an instrument attempting to measure an underlying concept. In the context of a multi-item instrument, asking an ambiguous question could (1) break symmetry in the respondent's interpretation of the full instrument and (2) inflate the variance of the responses to the question. The former may affect the convergence of measuring a factor, particularly on the questions asked after the ambiguous question; the latter will affect the ability of the ambiguous question to measure the underlying concept.

## Hypotheses

The text above suggests several testable hypotheses when parsing the effects of religious concept elicitation upon a psychometric instrument.

### Comprehension Context Hypothesis

When a religious concept is elicited, the interpretation of the full or partial instrument will vary. When measuring satisfaction with life, for instance, the interpretation of "satisfaction" might vary. According to the religious background of the respondent, satisfaction could mean satisfaction with the social

group from a church, the total justice of civilization, satisfaction with the God-concept intervention in their own life, and so on. Eliciting a religious concept could put the focus of "life satisfaction" upon a community concept, a religious philosophical concept, or other contaminating concept apart from what the researcher intended to measure. When the instrument contains multiple items, the religious concept is competing with items not contaminated with the religious concept. If questions contaminated by comprehension context exist alongside those that are not contaminated, the comprehension context hypothesis predicts

1. Within-respondent variance across instrument items will increase, and questions with contaminated meanings will substantially vary from items with uncontaminated meanings.
2. Factor analysis will neatly fit the data in the group where a religion is not mentioned, but not so in the alternative.

## Sponsorship Effects Hypothesis

The hypothesized sponsorship effects differ on two dimensions: the effect on the full instrument versus individual questions within the instrument; and secondly, in regard to its effect on respondents, a homogeneous versus heterogeneous effect.

*Full instrument, homogeneous response.* This hypothesis states that the sponsorship effect will, on average, have a similar effect for instrument items and individuals. Assuming an underlying phenomenon can be appropriately measured (i.e., the context comprehension hypothesis is false), the average response will differ according to the sponsorship context. The respondent pool in the group where a religious concept is mentioned will have similar cross-individual and cross-item variance as the group without a religious concept mention but significantly different average response across all items.

*Full instrument, heterogeneous response.* This hypothesis states that averaging the full instrument, the sponsorship effect will differ according to individual but will (on average) maintain the same effect across items. Assuming an underlying phenomenon can be appropriately measured (i.e., the context comprehension hypothesis is false), the average response will differ according to individual. Between-respondent variance will be higher among the group that was exposed to a religious concept.

*Partial instrument, homogeneous response.* This hypothesis states that the average response for some items will differ according to religious sponsorship exposure but not for the instrument in total. Cases where this will distinguish itself from the full instrument hypothesis are instances where there are conflicting average responses for items (i.e., one question is distorted toward a depression diagnosis while another question is distorted away from depression diagnosis) or not enough items are affected to move the full instrument mean. This hypothesis predicts equal variances across all items.

*Partial instrument, heterogeneous response.* This hypothesis states that the cross-respondent variance for some items will differ according to religious sponsorship exposure but not for the instrument in total. Cases where this will distinguish itself from the full instrument hypothesis are instances where there is both a reduction and inflation in variance in the religious concept group for separate items or not enough items are affected to move the full instrument variance.

## Data and Method

Data come from a chaplaincy-commissioned psychometric study in a large hospital in the Deep South. All respondents have chronic diseases and had a recent release from a hospital stay. Twelve questions concerning feelings of hope using Herth (1992) Hope Index were asked to 169 respondents through telephone interviews. The study was commissioned originally to assess the effectiveness of mental health and chaplaincy services at the hospital.

Social workers employed by the hospital administered the survey. Thirty-six respondents were mistakenly informed of the chaplaincy's involvement in the study. Once this was discovered, the social workers stopped mentioning the chaplaincy and instead informed the respondent only of the hospital's involvement in the survey.

The Herth (1992) Hope Index () was developed by Kaye A. Herth to assess hope in patients in clinical settings. Twelve questions (listed in Table 1) include questions that assess temporality as well as connectedness with a social environment.

### Data-specific Hypotheses

The *sponsorship effects* hypothesis would expect one or more items in the survey to vary either in mean or variance across the two groups: one group consisting of respondents that had the chaplaincy mentioned to them (i.e., *the*

**Table 1.** Herth (1992) Hope Index Questions.

| |
| --- |
| Q1. I have a positive outlook toward life |
| Q2. I have short and/or long range goals |
| Q3. I can see possibilities in the midst of difficulties |
| Q4. I feel scared about my future |
| Q5. I can recall happy/joyful times |
| Q6. I have deep inner strength |
| Q7. I am able to give and receive caring/love |
| Q8. I believe that each day has potential |
| Q9. I feel my life has value and worth |
| Q10. I feel all alone |
| Q11. I have a faith that gives me comfort |
| Q12. I have a sense of direction |

*religious elicitation context*) and the other having just the hospital involvement mentioned (i.e., *the nonreligious elicitation context*). The *comprehension context hypothesis* would suggest that ambiguity in the survey would be interpreted with greater error in the religious elicitation context due to the competing concepts of a religious figure and the question order context. For example, one possible source of ambiguity in this survey that is obvious a priori is question 11: "I have a faith that gives me comfort." An argument from the comprehension context hypothesis is that the interpretation of this question could be split between two concepts: the context of the 10 questions prior to it and the mention of the chaplaincy in the initial survey set. Due to this ambiguity, the comprehension context hypothesis would predict increased variance in this question in the context of mentioning the chaplaincy.

## Methods

Here, I use three primary methods: the *t*-test of equality of means, the *F*-test of equality of variances, and factor analytic methods. These methods will be applied between and within samples interviewed with and without a religious elicitation. Where applicable, we will perform bootstrap analyses to control for the differences in sample size between groups.

## Results

### Full Instrument Comparisons

We begin with comparing Herth Hope scores across the two groups—the religious elicitation group and the nonreligious elicitation group. If a
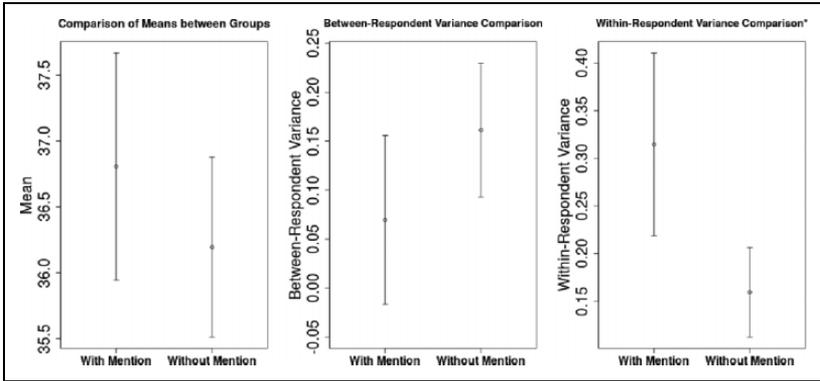
**Figure 1.** Full instrument comparisons.

religious sponsorship effect occurs homogeneously across respondents and items on average (i.e., *full instrument*, *homogeneous response*), we would expect a significant mean difference in scores. Panel 1 in Figure 1 shows that there is not, in fact, a significant difference between the two means.

Another type of sponsorship effect could be that the effects are homogeneous across items but heterogeneous across respondents. In this case, the first test one would do would be an $F$-test of between-respondent variance between the two groups. If significantly higher in the religious elicitation group, this would be necessary but not sufficient evidence that the religious elicitation had a differential effect on individuals according to an unobserved individual characteristic. The second panel shows the result from this test; there is no significant difference across groups for between-respondent variance.

The final type of full instrument comparison corresponds to the comprehension context hypothesis. This test determines the consistency of individuals in their answering of questions. This is a comparison of average within-respondent variation between groups. If significantly higher in the religious elicitation group, this is evidence that the underlying concept was less successfully measured, and respondents developed a response of the item that differed from what was intended by the scale. This test shows that there is a significantly greater within-respondent variation in the group with the religious concept elicitation.

These initial results suggest little support for a homogeneous item response by group. Indeed, the results so far suggest that respondents in the religious elicitation group were less likely to answer the survey in a way that
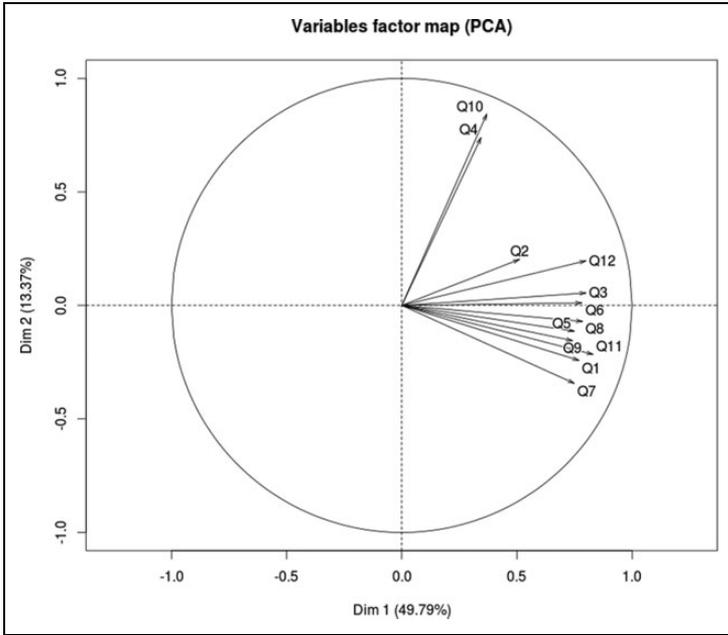
**Figure 2.** Principal Component Analysis without mention of the chaplaincy.

measured an underlying concept. To further explore this hypothesis in the next section, we employ factor analytic methods.

## Factor Analytic Comparisons

A key element of measuring reliability in psychometric instruments is factor analysis. Factor analysis reduces a correlation matrix into a predetermined number of dimensions, allowing the researcher to view a number of closely related "factors" that explain the data.

For the Herth Hope Assessment, a reasonable factor structure a priori would be the division of the data into two factors—a hope factor and a lack of hope factor. Questions 4 ("I feel scared about my future") and 10 ("I feel all alone") should be aligned along one factor, while the rest of the questions should be aligned upon another.

When imposing two factors upon the data, the two groups show remarkable differences in factor structure and model fit. Figures 2 and 3 compare a principal component analysis applied to both groups. In the group without a
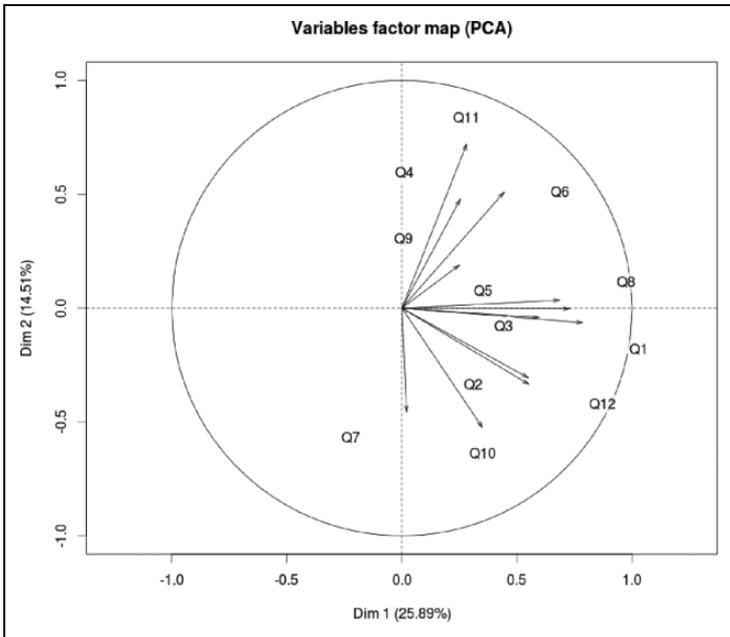
**Figure 3.** Principal Component Analysis with mention of the chaplaincy.

religion mention, the questions cluster nicely around separate factors, with questions 4 and 10 separated from the rest. In the group with a religious mention, the principal component is scattered and disorganized.

We compare the loadings of a factor analysis with varimax rotation in Table 2. With few exceptions, the loadings for all variables are lower in the case of mentioning the chaplaincy versus not.

Another method for comparing the results of a factor analysis is to measure the agreement, or congruence, between the factors calculated from separate analyses. The congruence coefficient between the two principal factors is .46, much lower than the .95 suggested value for agreement between factors. This suggests that the factors are not in agreement between situations where a chaplaincy was mentioned versus not, even though the questions are identical!

Finally, we compare the factor analysis model fit between groups. Model fit in factor analysis is simply the degree to which the factors correlate with the correlation matrix between input variables. The model fit in the case with no mention of chaplaincy is .56; in the case without the chaplaincy, the model is a much higher .92.

**Table 2.** Principal Factor Loadings with and without a Mention of Chaplaincy.

| Question | No Mention | With Mention |
|---|---|---|
| Q1. I have a positive outlook toward life | .76 | .77 |
| Q2. I have short and/or long range goals | .38 | .39 |
| Q3. I can see possibilities in the midst of difficulties. | .71 | .51 |
| Q4. I feel scared about my future | .14 | .19 |
| Q5. I can recall happy/joyful times | .73 | .54 |
| Q6. I have deep inner strength | .71 | .19 |
| Q7. I am able to give and receive caring/love | .77 | .06 |
| Q8. I believe that each day has potential | .73 | .70 |
| Q9. I feel my life has value and worth | .71 | .18 |
| Q10. I feel all alone | .06 | .41 |
| Q11. I have a faith that gives me comfort | .85 | .12 |
| Q12. I have a sense of direction | .69 | .55 |
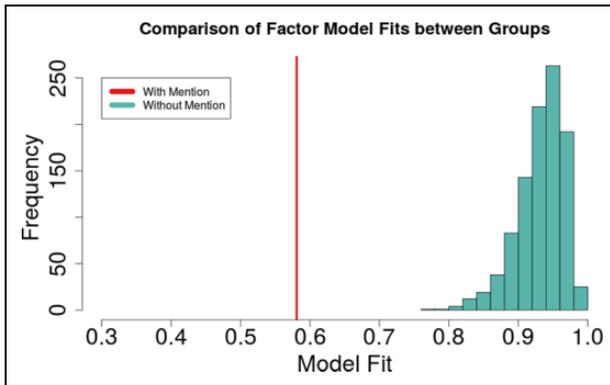


**Figure 4.** Comparison of factor model fits between groups ("without mention" group permuted).

There is the possibility that the much higher fit in the case without a chaplaincy mention is due to the much large sample size (133 vs. 36). To test this, we took 1,000 random samples of size 36 from the 133 persons in the nonreligious elicitation set. The results are shown in Figure 4: Of the 1,000 samples, none fit as poorly as the religious elicitation set, suggesting that the difference in model fit is not due to sample size.

The results above suggest that the ability to measure an underlying phenomenon using the Herth Hope Index is substantially reduced simply by mentioning that the chaplaincy was involved in commissioning the survey. Factor analysis neatly fits the data when the chaplaincy is not mentioned, but not so when the chaplaincy is mentioned.

## Partial Instrument Comparisons

This section tests the last part of both the comprehension context and sponsorship effects hypotheses. The results so far observe the full instrument, while ignoring heterogeneity across items; in this section, we decompose the responses by item, rather than pool the full instrument results. First, we compare the mean differences for each item between the religious versus nonreligious elicitation groups. Then, we compare variances. Finally, we compare uniquenesses for each question regarding the fitted factor analyses.

*Comparison of means.* Three questions show significant mean differences between groups. In each question, the mean for the group with a mention of chaplaincy is significantly higher than in the group without such mention. This suggests that sponsorship effects could be at play in these data. One theme between the three questions with significantly higher means (questions 5, 6, and 9) is that they do not relate to elements of future temporality (such as questions 1, 2, 4, 8, and 12) and instead relate to the past (question 5), or present, introspective elements of hope (questions 6 and 9). Also, note that the differences in these questions are uniformly positive; that is, those with a mention of chaplaincy typically report more "hope" than those without a mention.

*Comparison of variances.* A significant difference in variances between groups could mean one of the two possibilities: one, an increased variation in the group with mention of a chaplaincy could mean that the chaplaincy mention offers a competing concept that alters the interpretation of the question for some but not others. If this were the case, question 11 ("I have a faith that gives me comfort") would be most affected. Second, it could mean that a sponsorship effect could be affecting answers heterogeneously across respondents. Table 3 shows these differences. The greatest differences between groups show via question 11, the question most likely affected by a comprehension context effect. When a chaplaincy is mentioned, it resulted in five times the variation in responses than when the chaplaincy was not mentioned.

**Table 3.** Comparison of Variances between Respondents with and without a Mention of Chaplaincy.

|  | Ratio of Variances ($p$) |
|---|---|
| Q1. I have a positive outlook toward life | 1.13 (.716) |
| Q2. I have short and/or long range goals | 1.09 (.798) |
| Q3. I can see possibilities in the midst of difficulties | 1.76 (.098) |
| Q4. I feel scared about my future | 1.44 (.289) |
| Q5. I can recall happy/joyful times | 0.36 (.003)** |
| Q6. I have deep inner strength | 0.35 (.003)** |
| Q7. I am able to give and receive caring/love | 0.94 (.855) |
| Q8. I believe that each day has potential | 0.68 (.257) |
| Q9. I feel my life has value and worth | 0.2 (0)*** |
| Q10. I feel all alone | 1.4 (.329) |
| Q11. I have a faith that gives me comfort | 5.48 (0)*** |
| Q12. I have a sense of direction | 1.02 (.963) |

*$p$ < .05, **$p$ < .01, ***$p$ < .001

**Table 4.** Comparison of Means between Respondents with and without a Mention of Chaplaincy.

|  | Mean Difference ($p$) |
|---|---|
| Q1. I have a positive outlook toward life | .14 (.108) |
| Q2. I have short and/or long range goals | −.15 (.284) |
| Q3. I can see possibilities in the midst of difficulties | −.06 (.58) |
| Q4. I feel scared about my future | −.22 (.224) |
| Q5. I can recall happy/joyful times | .14 (.022)* |
| Q6. I have deep inner strength | .28 (.001)*** |
| Q7. I am able to give and receive caring/love | .1 (.272) |
| Q8. I believe that each day has potential | .13 (.105) |
| Q9. I feel my life has value and worth | .24 (0)*** |
| Q10. I feel all alone | .06 (.679) |
| Q11. I have a faith that gives me comfort | −.22 (.222) |
| Q12. I have a sense of direction | .01 (.945) |

Three other questions also had significant differences in variation between groups; these questions correspond exactly to the questions that had a mean difference in Table 4. What is likely occurring here is that those who have had a chaplaincy mentioned to them are altering their answers to be monolithically positive, which decreases the variation to the left of the mean. The distributions of these variables (Figures 5–7) show this to be the case:
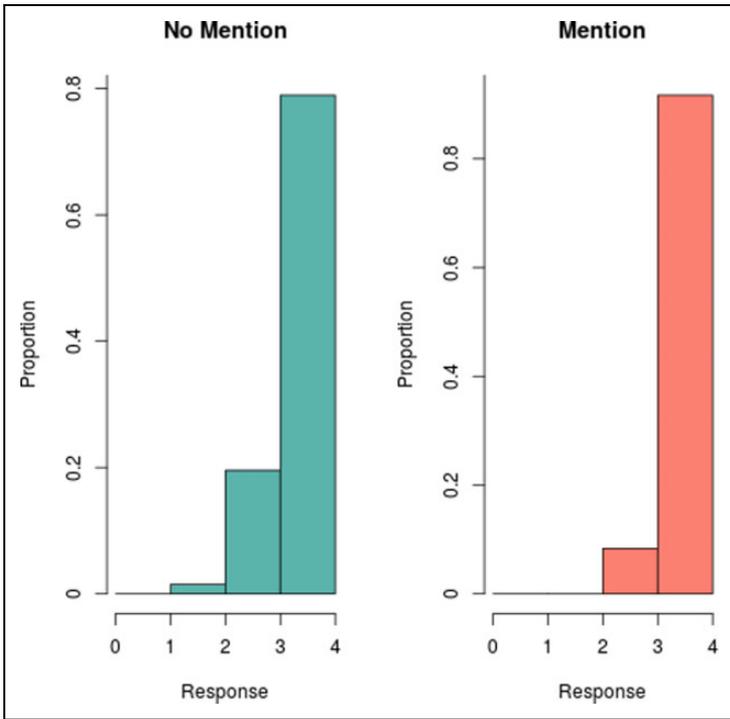
**Figure 5.** Distribution of the faith question (question 11) between groups with and without a mention of chaplaincy.

When the chaplaincy was mentioned, no one scored below 3 on questions 5, 6, or 9 (Figure 8), decreasing the variance. These three questions were thus likely affected by a sponsorship effect, which affected both the reported means and variances.

Observing the distribution of question 11, we see inflated values toward the 0 end of the scale, indicating an inflated number of respondents who answer that they do not have a faith that gives them comfort. In the condition where the chaplaincy is not mentioned, the answers decrease monotonically from 4, suggesting an anchor on 4 that is indicative of the rest of the questions. This strongly suggests that respondents are retrieving the chaplaincy mention in their interpretation of the question.

Lastly, we look at the uniqueness of each question, measured by the amount of variance not explained by the factors returned by the factor
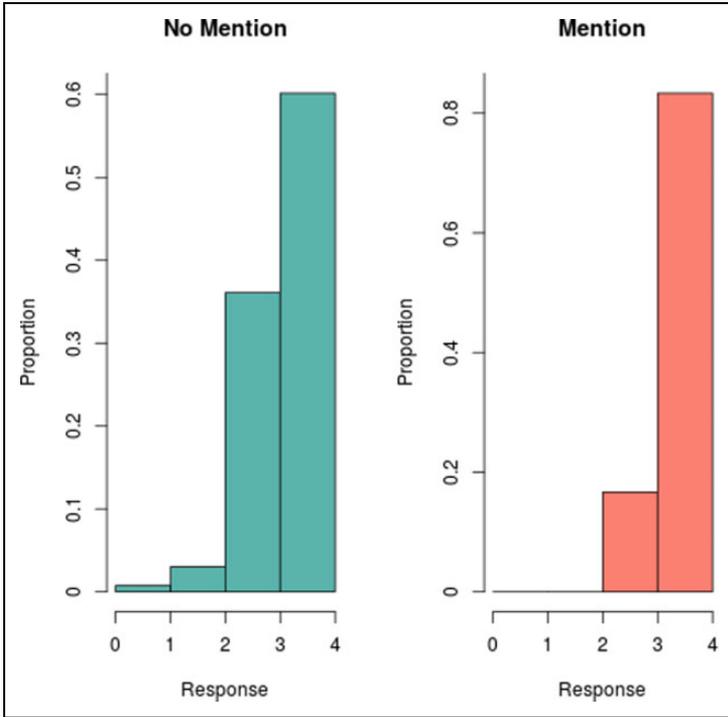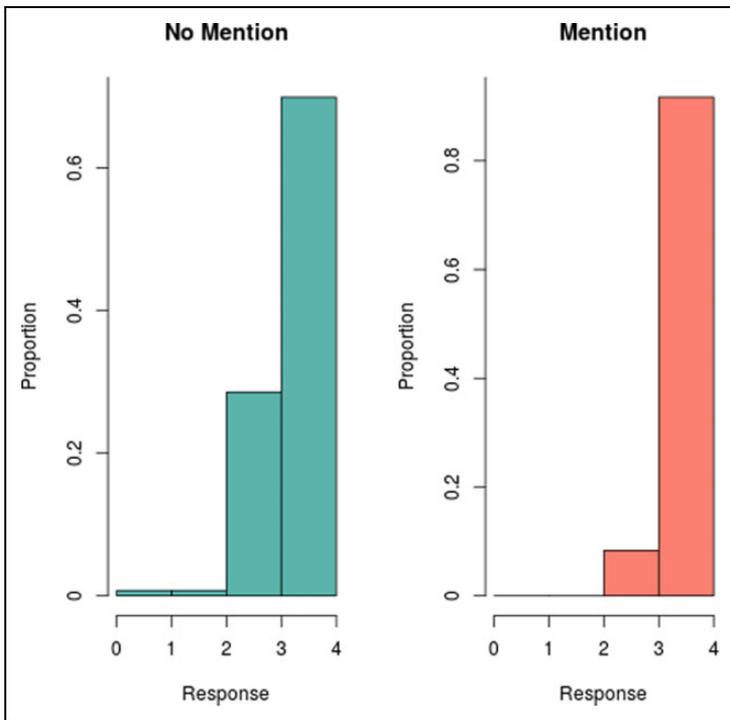
**Figure 6.** Distribution of question 5 between groups with and without a mention of chaplaincy.

analysis. Although questions 3, 11, and 12 all gain in uniqueness, question 11 has the largest gain, jumping from 27% variance unexplained to 55%. What this shows is that question 11 systematically varies from the other responses, for example, even when a respondent shows positive hope in other questions, it does not mean that they will respond similarly to this question (Table 5).

## Discussion

Overall, strong support for the comprehension context hypothesis suggests that a religious concept elicitation can substantially hinder the reliability of the Herth Hope Index. This is due in part to the ambiguity of a question on "faith," which in most contexts is interpreted in terms of the question ordering, that is, in the context of the rest of the instrument versus a religious

**Figure 7.** Distribution of question 6 between groups with and without a mention of chaplaincy.

context. With a religious concept elicitation, respondents instead interpreted this question in terms of a religious faith.

Support for the homogeneous sponsorship effect (partial instrument) also had support. Although the full instrument's mean did not significantly differ between groups, certain items did differ, and in a single direction. Respondents reported significantly greater "inner strength," reported that their lives had more value and worth, and that they agreed more with the statement, "I can recall happy/joyful times." Although the exact cause of this is unknown, one possible mechanism is that a religious concept imposes a feeling of gratitude in respondents as has been shown before (Tsang, Schulwitz, and Carlisle 2012). Respondents may be less inclined to report distress due to the obligation to be grateful for their circumstances handed to them by their "creator." Interestingly, while this effect occurs for questions regarding to
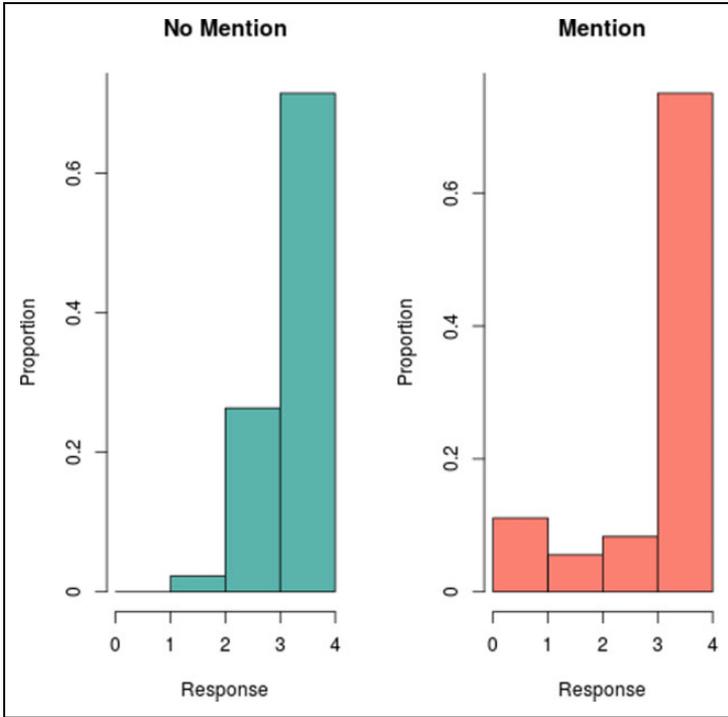
**Figure 8.** Distribution of question 9 between groups with and without a mention of chaplaincy.

present and past hope, questions on temporality and future hope did not systematically show any sponsorship effect.

   Although sponsorship effects usually depend upon mean differences, it is also prudent to compare measures of variation, which can contain a wealth of information in itself (Heise 2010). In this case, what could be construed as a "sponsorship effect" turned out to be a comprehension context effect, which is a closely related but distinct effect. This analysis suggests that researchers, when looking for sponsorship effects, also take into account the effects of sponsor characteristics on the interpretation of questions. This article provides a simple blueprint for how to conduct such an analysis.

   Finally, researchers should be careful to consider alternative interpretations of questions they construct and strive to reduce ambiguity in sample items. By simply introducing a religious concept, a well-verified instrument

**Table 5.** Comparison of Uniqueness between Groups with and without a Mention of Chaplaincy.

|  | No Mention | With Mention |
|---|---|---|
| Q1. I have a positive outlook toward life | .42 | .28 |
| Q2. I have short and/or long range goals | .78 | .00 |
| Q3. I can see possibilities in the midst of difficulties | .42 | .51 |
| Q4. I feel scared about my future | .57 | .00 |
| Q5. I can recall happy/joyful times | .44 | .15 |
| Q6. I have deep inner strength | .46 | .00 |
| Q7. I am able to give and receive caring/love | .41 | .00 |
| Q8. I believe that each day has potential | .42 | .39 |
| Q9. I feel my life has value and worth | .48 | .35 |
| Q10. I feel all alone | .00 | .18 |
| Q11. I have a faith that gives me comfort | .27 | .55 |
| Q12. I have a sense of direction | .39 | .60 |

cited 500 times lost any resemblance of reliability. Much of the reliability of instruments come from the respondent, understanding the intent of the questions being asked; when giving respondents another concept to consider, this study shows that the reliability enjoyed by many psychometric studies could suffer substantially.

## Conclusion

Little empirical work has observed (1) the effect of religious elicitation upon responses, (2) the effect of context comprehension hypothesis when introducing multiple concepts, and (3) measures of variability in extracting information on these effects. If a researcher is worried that a question has multiple meanings, a systematic study of this type determining whether introducing competing concepts systematically altering the variation and reliability of responses may be helpful. In some cases, it may reduce error by eliminating any ambiguity in respondents' conceptual understanding of a question.

### Declaration of Conflicting Interests

### Funding

## References

Bishop, George F., Alfred J. Tuchfarber, Robertand W. Oldendick . 1986. "Opinions on Fictitious Issues: The Pressure to Answer Survey Questions." *Public Opinion Quarterly* 50:240-50.

Blaydes, Lisa and Rachel M. Gillum. 2013. "Religiosity-of-interviewer Effects: Assessing the Impact of Veiled Enumerators on Survey Response in Egypt." *Politics and Religion* 6:459-82.

Boulianne, Shelley, Casey A. Klofstad, and Danna Basson. 2010. Sponsor Prominence and Responses Patterns to an Online Survey. *International Journal of Public Opinion Research* 23:7987.

Doob, Anthony N., Jonathan L. Freedman, and J. Merrill Carlsmith. 1973. "Effects of Sponsor and Prepayment on Compliance with a Mailed Request." *Journal of Applied Psychology* 57:346-7.

Edwards, Michelle L., Don A. Dillman, and Jolene D. Smyth. 2014. "An Experimental Test of the Effects of Survey Sponsorship on Internet and Mail Survey Response." *Public Opinion Quarterly* 78:734-50.

Fox, Richard J., Melvin R. Crask, and Jonghoon Kim. 1988. "Mail Survey Response Rate: A Meta-analysis of Selected Techniques for Inducing Response." *Public Opinion Quarterly* 52:467-91.

Grice, H. Paul, Peter Cole, and Jerry L. Morgan. 1975. "Syntax and Semantics." *Logic and Conversation* 3:41-58.

Heise, David R. 2010. *Surveying Cultures: Discovering Shared Conceptions and Sentiments*. Hoboken, NJ: Wiley.

Herth, K. 1992. "Abbreviated Instrument to Measure Hope: Development and Psychometric Evaluation." *Journal of Advanced Nursing* 17:1251-9.

Krosnick, J. A. 1991. "Responses Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5: 213-36.

Krosnick, J., A. L. Holbrook, M. K. Berent, R. T. Carson, W. M. Hanemann, R. J. Kopp, R. C. Mitchell, S. Presser, P. A. Ruud, V. Kerry Smith, W. R. Moody, M. C. Green, and M. Conaway. 2002. "The Impact of 'No Opinion' Response Options on Data Quality: Non-attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly* 66:371-403.

Norton, Sam, Theodore Cosco, Frank Doyle, John Done, and Amanda Sacker. 2013. "The Hospital Anxiety and Depression Scale: A Meta Confirmatory Factor Analysis." *Journal of Psychosomatic Research* 74:74-81.

Schuman, Howard and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. San Diego: Academic Press

Sudman, Seymour, Norman M. Bradburn, and Norbert Schwarz. 1996. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, CA: Jossey-Bass.

Tsang, Jo-Ann, Ashleigh Schulwitz, and Robert D. Carlisle. 2012. "An Experimental Test of the Relationship between Religion and Gratitude." *Psychology of Religion and Spirituality* 4:40.

## Author Biographies

**C. Ben Gibson** is a PhD candidate in the Department of Sociology at University of California, Irvine.

**Timothy B. Mayhall** is Director of Spiritual Care at Southest Alabama Medical Center.