Last August and September 2011, we published two-part survey articles on common analyses in business surveys. And last year, we published four-part articles on executing successful business surveys. In this article and a subsequent article next month we continue our discussion of business surveys with focus on a critical part of any survey study; namely the reliability and validity of the survey.

In this article, we first explain the concepts of reliability and validity of a survey study. We then explain the different types of reliability tools. Our coverage of reliability relied to a large extent on five primary references.

**RELIABILITY**

Reliability refers to the degree of stability, consistency or repeatability demonstrated when a survey is repeated under identical conditions. It is particularly important if the survey is to be used on an on-going basis to detect change over time. In other words; a survey is reliable when it does not result in one of the following two errors:

- indicating a change when there is no change.
- indicating no change when there is a change.

In general, creating a perfectly reliable survey is impossible since there are many, difficult to control, intervening factors that could change a person's response from one question to the next or from one survey to the next. These factors include:

- Variations in mood.
- Fluctuations of human memory.
- Unpredictable fluctuations in attention or accuracy.
- Health and fatigue.

However, there are other factors that are relatively easy to control. These include:

- Inaccuracies in scoring.
- Motivation of respondent.

There are different types of reliability that will be detailed later in this article. But in general, the assessment of reliability requires examining two kinds of consistencies in the questionnaire. The first is consistency within the questionnaire. This is to assess the extent to which responses to similar questions within the questionnaire are similar. The second is consistency over time, which assesses the extent to which responses to the same questionnaire at different time slots are consistent.

**VALIDITY**

Validity refers to the degree that a survey questionnaire actually measures what it is designed or intended to measure. Establishing the validity of a survey questionnaire is especially important for most surveys since it strengthens the data generated from the data collection process, which results in greater confidence in the interpretation of the survey results.

External validity allows for extending the interpretation of the survey results across populations, settings and time. Suggestions on improving external validity include:

- Use of a random sample.
- Use of a representative sample (of different segments of population, time, and settings).
- Repetition of the study.

Internal validity, on the other hand, is related to the extent to which the survey is accurately answering/addressing the question/purpose of the study. A major cause for lack of internal validity is the presence of confounding variables (also referred to as extraneous variables or artifacts); the sources of which are many; including the presence of intervening variable(s) in a pre- and post-intervention survey; where the change in opinion is due to the intervening variable(s) rather than the intervention itself; e.g. maturation of the subjects, biases in the
sample selection, tendencies of the respondents to be biased by researcher’s prejudice, tendencies of the respondents not to be truthful when being observed.

Validity is considered more essential than reliability since there is no point of measuring the consistency of the survey questionnaire without first ensuring that the questionnaire is actually measuring what it is designed to measure. However, validity is difficult to measure and there is no single measure that is considered an absolute proof of validity. Therefore, it is recommended to use more than one tool to ensure the highest possible validity. Some of these tools are explained in more detail in a future article.

**RELIABILITY TYPES**

Reliability consists of three types: stability, internal consistency, and equivalency. A forth type is exaggeration which is closely related to equivalency.

**Stability**

The first type of reliability is stability (also known as test-retest reliability), which tests whether repeating the survey under the same conditions produces the same results. Basically, survey respondents are asked to complete the survey at two different points in time, and the results of the two survey responses are compared by a correlation coefficient. If this correlation coefficient is greater than +0.70 (maximum is +1.00), then the survey is considered reliable.

Sometimes however, respondents could answer the survey the second time based on their memory, which would inflate reliability. Another limitation of this tools is that the concept or construct being measured may change over time; thus lowering the reliability estimate.

In general, measuring consistency of survey responses over time is a minor concern in business surveys, since most business surveys deal with attitudes, opinions, and behaviors that are expected to change over time. This reliability type is therefore best used for establishing a survey questionnaire to be used in certain applications or research purposes, and consequently, is not considered generally practical for business applications.

**Internal Consistency**

Internal consistency reliability is applied to groups of items that are thought to measure different aspects of the same concept. It is a reflection of how well the different items complement each other in their measurement of different aspects of the same concept or quality. It is quantified using two methods: the split-half test and Cronbach’s coefficient alpha.

In split-half test, the items of the questionnaire (say 30 items) are split into two halves. A high reliability is attained when the total score on the first half of the test correlates highly (at least 0.70) with the score on the second half of the test. The split-half test is relatively easy to implement, and considered generally practical when the items of the questionnaire measure different aspects of the same concept.

The Cronbach's coefficient alpha method uses the analysis of variance to assess internal consistency. Cronbach's coefficient alpha reliability is considered time consuming and requires excess data collection until the final survey questionnaire is reached; therefore, it is not considered generally practical for use in business surveys.

**Equivalency**

Equivalency reliability (also referred to as alternate-form and cross-test reliabilities) refers to the use of differently worded forms to measure the same items on a survey or even using two parallel forms of a survey. Unlike stability and internal consistency reliabilities, alternate-form reliability is used to detect inconsistency or falsehood in survey responses and not in the survey questionnaire structure itself. Alternate-form reliability could be achieved through simply changing the order of the response alternatives, changing the wording of the response alternatives without changing the meaning, stating some questions in a positive form and others in a negative form, or changing the wording of the question while maintaining the level of difficulty.

**Exaggeration**

Another tool used to measure the reliability of survey responses that is related to but not directly under equivalency reliability is exaggeration, where fake elements are added to the survey to determine how truthful the respondent is in answering the survey.
Reliability and Validity of a Survey Study – Part One

Equivalency reliability and exaggeration are simple to use and help detect falsehood in survey responses. They are not time consuming, do not require excess data gathering and they help eliminate survey responses that exhibit a certain level of inconsistency so as not to negatively affect survey analysis and conclusions. Figure (1) Example of equivalency exaggeration reliabilities.

CONCLUSION

We have introduced in this article the concepts of reliability and validity in survey studies, and discussed the common tools of reliabilities. Table (1) lists the discussed reliability tools along with their evaluation as judged for being generally practical in business surveys.

### Table 1. General business survey practicability of reliability tools.

<table>
<thead>
<tr>
<th>Reliability</th>
<th>Stability (test-retest)</th>
<th>Split-half test</th>
<th>Cronbach’s alpha</th>
<th>Equivalency (alternate-form, cross-test)</th>
<th>Exaggeration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practicability</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Figure (1) Example of equivalency exaggeration reliabilities.

Sources


Tariq A. Aldowaisan has a Ph.D. from Arizona State University in 1990. He is currently an associate professor of industrial and management systems engineering at Kuwait University, and a general manager of Gulf Lead Consultants. He has professional certifications from reputable organizations; they are Certified Safety Professional (CSP) from the Board of Certified Safety Professionals (BCSP), USA; and Certified Quality Engineer (CQE), Certified Quality Auditor (CQA), and Certified Quality Manager/ Organizational Excellence (CQM/OE) from the American Society for Quality (ASQ). He has over 20 years consulting and training experience in quality, organizational excellence, management systems design, and occupational safety.

Aliaa M. Alnaggar is an Industrial and Management Systems Engineer at Gulf Lead Consultants, Kuwait. She received her bachelor degree in industrial and management systems engineering from Kuwait University. She has experience in ISO 9001 and management system development, conducting business surveys, and human resource management.
• Thank you for reading this article; hope you found it both enjoyable and useful. I look forward to receiving your opinion and to always maintain communication with you at tariq@glc-im.com.

• If you wish your name to be removed from our database for future articles, please let me know. I also welcome your suggestion for new friends to be added to our database to receive future articles.

Upcoming GLC and Partners Events

• Benchmarking for Performance and Best Practice (4-5 Dec 2011) http://kuwaitbenchmarking.com/

Quote of the Month

"One of the deepest needs of the human soul is to be understood. Once that need is met, the personal focus can shift to interdependent problem solving." – Stephen Covey