The RQC's 2009 Foundations of Quality Initiative:

- Produced powerful evidence of the need to employ tools and procedures to minimize panel overlap and respondent duplication.
- Identified a need for new processes and systems to improve the ability to measure and manage critical variablity in sample composition across a series of studies, particularly where a mix of panels is involved.
  - To that end, an industry Panel Quality Enhancement Process (QeP) was set forth and pilot testing of the reporting templates was undertaken and completed with the release of the QEP 2.0 Templates.
  - QEP workshops have been launched to facilitate communication between buyer and seller on quality control.

However, not all the issues identified in the RQC knowledge sharing program could be accommodated in the intial FOQ study. The RQC Research on Research Committee was charged to lay out a plan by which the progress would continue. A three-fold program was developed and reviewed by the RQC in April, 2010.

I. FOQ Q&A Framework

The centerpiece of the FOQ 2.0 plan developed by the committee is a simple, straightforward framework for online research that market researchers and others can use, and build upon, to inform their decision on choice of mode (e.g., telephone, face-to-face, online, mail).

The framework recognizes that researchers make trade-offs when deciding whether online research (or any other mode) is appropriate for the task at hand. Typically, as part of the decision-making process, they ask themselves (or their research agency) questions on sample availability, quality, consistency, etc. So the framework is in the form of questions and answers.

II. Systematic Review of Existing Empirical Evidence

Collect, evaluate and organize existing reports on empirical evidence on online research quality, consistent with the proposed FOQ framework. This is a necessary step in developing a prioritized agenda for the new research in FOQ 2.0.

It is the goal of the systematic review to make plain what we know, and don't know regarding the issues identified in the questions outlined in the proposed FOQ Framework. The review is also seen as a necessary step in developing a prioritized agenda for new research. A second step will involve the review, modification, and approval of the agenda by the ARF RQC Leadership.

The cost of a high quality systematic review, overseen by academic researchers at either Stanford University or University of Michigan, will be $10,000 for 4-5 weeks of full-time work.

III. Conduct Research to Evaluate Data Quality Solutions

Extend existing learning through a new Foundations of Quality (FOQ 2.0) initiative across panels and other sample sources to answer questions concerning the impact on data quality practices such as real-time sampling, routers, identity verification, survey design, sampling, and weighting.

FOQ Discussion Framework: Develop a simple, straightforward framework for online research that market researchers and others can use to inform their decision on choice of mode. Revise this framework over time.

11 November 2010

A Question-Based Framework for Evaluating Online Research Quality and Establishing a Prioritized Agenda for New Research

I. The Value and Trustworthiness of Online Research

Questions for Buyers to Consider (in order to reduce the risk of making a bad business decision)

A. Are sufficient numbers of your target population accessible through online research? In other words, is online research a sustainable solution over time given your needs?

B. Will the results you obtain through online research…

1. give you valuable insights into your target population's thoughts and feelings about your products and services, or whatever it is that may interest you (e.g., loyalty to your brand, likelihood to recommend)?
2. give you valuable insights into your target population's behaviors; notably, those associated with the outcome(s) of interest to you?
3. be reliable – that is, at the same (or a subsequent time), will you obtain approximately the same results with a similar sample (when there is no underlying change in what you are measuring), and, related, will you be able to detect a change in results when there has been a real change (due to advertising, seasonality, external events – e.g., Toyota recalls, recession)?
4. be as accurate as, or more accurate than, those that you obtain through different modes of interviewing? (e.g., telephone, paper-pencil, face-to-face, mobile devices)?
5. be as accurate as, or more accurate than, those that you obtain if you depended on different methods of generating the sample of respondents (e.g., probability-based, snow-ball recruitment, central location sampling, and so on)?
6. be as cost-effective as, or more cost-effective than, other approaches?
7. be as timely as, or more timely than, other approaches?
8. How important are each of the factors identified above when you are deciding whether to commission online research? Do you weight each of these factors? If so, how? And for what reasons?

II. The Relationship between Sample Sources and the Reliability and Accuracy of the Data: *Questions for Buyers to Ask Sellers*

What is the relationship between the method by which online respondents are recruited and the reliability and accuracy of the responses they give to your questions? How do you know?

A. Access Panels

1. Probability-based recruited (e.g., randomly selected from a larger population, often resulting from recruitment from random telephone samples)
2. Non-probability-based recruited (e.g. creating a panel from volunteers who have been recruited by means <u>other</u> than population-wide random sampling)

B. Behavioral Panels (e.g., Nielsen, ComScore, IRI)

1. Probability-based recruited (e.g., randomly selected from a larger population, often resulting from recruitment from random telephone samples)
2. Non-probability-based recruited (e.g. creating a panel from volunteers who have been recruited by means <u>other</u> than population-wide random sampling)

C. Non-panel, Web-based Recruitment

1. River sampling/Real-time sampling
2. Web-intercepts (e.g., for online website-specific customer satisfaction or website evaluation studies)

D. Lists (e.g., customers, association members, employees, direct marketing opt-ins, and so on)

E. Other (Individuals who post comments online on sites such as Twitter or Facebook or even the comments themselves, irrespective of the identity of the poster)

III. Multiple Modes and Sample Sources: *Questions for Buyers to Ask Sellers*

What steps do you take to assure reliability and accuracy when you depend on multiple modes, multiple sample sources, or both for particular projects? What steps do you take when you transition studies from one mode to another? Are these steps effective? How do you know?

A. Multiple Modes, Multiple Sample Sources

1. Multiple sample sources for the same project
    a. Are there “house” (aka, “source”) effects?
        (1) If there are “house” effects, can they be adjusted for? For instance, can they be mitigated by respondent verification methods or panel conditioning metrics?

(2) If there are "house" effects, are they consistent for all variables?

2. Multiple modes (e.g., how do you put together the data given possible systematic differences due to mode effects?)
3. Multiple sample sources <u>and</u> modes

B. Transitioning from One Mode to Another

1. Telephone to Online
2. Face-to-face to Online
3. Others

IV. Sample Selection and Weighting Approaches to Improve the Reliability and Accuracy of the Data Obtained: *Questions for Buyers to Ask Sellers*

What (individual survey) sample selection and post-field adjustment methods do you depend on to improve the reliability and accuracy of the data you obtain? How do you know that they work?

A. Sample Selection Methods

1. Exclusions - What selection rules do you have in place on such factors as invitation frequency, systematic exclusions, respondent uniqueness and verification (are they who they say they are)?
2. Sample selection criteria (demographics, likely purchasers, stored information from prior surveys)?
3. Screening criteria within surveys
4. Use of quotas within surveys
5. Use of survey routers
6. Other

B. Post-fielding Adjustment Methods:

1. *Post-hoc* sample design adjustments: weighting--demographic, propensity scores, etc.
2. *Post-hoc* sample eliminations
   a) Response non-differentiation
   b) Random responding
   c) Illogical or inconsistent responding
   d) Overuse of item non-response (e.g. "Don't Know")
   e) Speeding (too rapid survey completion)
   f) Duplications (ensuring unique respondents)
3. Other

V. Optimal Survey Design: *Questions for Buyers to Ask Sellers*

How do you design the survey instrument optimally (e.g., wording, scales, colors, length, and so on) to improve reliability and accuracy?

A. Instructions (Clear and understandable to respondents?)
B. Question wording (Clear and understandable?)
C. Question order and logic (Does an unfolding question structure make sense to respondents?)
D. Use of audio and visual stimuli (Are “extras” being used in productive ways?)
E. Responses categories (Adequate for the topic, understandable, ordered appropriately?)
F. Total survey length (Has respondent burden been minimized?)
G. Engagement (Has boredom been eliminated through reduction of redundancy?)
H. Appearance (Does the survey look professional, is it obvious to respondents on how to answer questions?)
I. Other design factors

VI. Respondent Motivation and Incentives: *Questions for Buyers to Ask Sellers*

How do you address respondents’ motivational needs? Do the incentives you offer increase the reliability and accuracy of the results you obtain? Do they create any perverse effectives? How do you know?

A. Incentives (Type, level, and frequency)
   1. Extrinsic Rewards
      a) Monetary
      b) Points
      c) Sweepstakes
      d) Others
   2. Appeal to altruistic motivations
   3. Appeal to consumer/topical needs
   4. Appeal to intellectual curiosity
   5. Appeal to social comparison
   6. Other?

B. Effects
   1. Short-term
   2. Over time

VII. Respondent Communication: *Questions for Buyers to Ask Sellers*

How does communication with the respondents influence their response attributes? (e.g., response rates, suspend rates, reliability, accuracy, panel membership--joining and continuing)? How do you know?

A. Communicating survey characteristics (e.g., subject lines in the survey invitation, information concerning topics, incentives, length of time required to complete interview)?

B. Other (e.g., initial communications upon joining the panel, post-survey instant results, newsletters, help desk)?