The annual ATE survey is around the corner! This year, the survey will be open from February 19 through March 18. Download a review copy of the survey here: evalu-ate.org/annual_survey. We’re working on confirming contact and grant information now.

Welcome to the EvaluATE team, Emma Perk! Emma is working on outreach, material development, and project management.

Thanks to everyone who has responded so far to the EvaluATE Stakeholder Survey. Our external evaluator, Dr. Lana Rucks, is going to keep the survey open a few more weeks to gather more responses. Please take a moment to participate if you haven’t already.

Happy 2013!
The ATE program turns 20 this year!

---

**Aligning Survey Questions with Evaluation Use**


To enhance evaluation use, decisions about methods—like those about an evaluation’s focus—should be guided and informed by the intended use of intended users. This process is informed by having intended users actively engage in making decisions about methods. This turns out to be a controversial approach.

Evaluators have mostly come to accept that use can be enhanced by actively involving intended users during the conceptual phase of an evaluation. Where my colleagues and I often part company, is in the role to be played by intended users in making measurement and design decisions. They see evaluators as experts in methods and statistics who ought to be left with full responsibility for operationalizing program goals and determining data collection procedures. Utilization-focused evaluation takes a different path.

Designing an evaluation to generate credible and useful findings typically involves making difficult decisions about what to do, given time and resource limitations. Since those who will use the findings have a major stake in the evaluation, their involvement in deciding what is most appropriate for the situation is critical for credibility and utility.

Let’s consider the odd-even question: Should survey response scales be even-numbered (e.g., 4 or 6 choices) or odd-numbered (e.g., 3 or 5 choices)? While it doesn’t seem like such a big deal, I’ve seen evaluators on both sides of the question go at each other with the vehemence of Marxists versus capitalists, osteopaths versus chiropractors, or cat lovers versus dog lovers. What’s all the ruckus about? It’s about the value and validity of a midpoint on questionnaire items. In conducting workshops on evaluation, one of the most common questions I get is: “Should we give people a midpoint?”

Even-numbered scales force respondents to lean in one direction or the other (although a few will circle the two middle responses, creating their own midpoint). Odd-numbered scales allow the respondent to hedge, to be undecided, to avoid making a decision one way or the other, or to be genuinely in the middle.

Which approach is best? Having carefully considered the arguments on both sides of the issue, having analyzed large numbers of questionnaires with both kinds of items, and having meditated on the problem at great length, I find that I’m forced to come down firmly and unwaveringly right smack in the middle: It depends. Sometimes odd-numbered scales are best and sometimes even-numbered scales are best. How to decide?

The issue is really not technical, statistical, or methodological. The issue is one of utility. What do intended users want to find out? Will the findings be more useful if respondents are forced to lean in one direction or the other? Or is it more useful to find out how many people are undecided or “don’t know?” The evaluator helps the primary intended users determine the value and implications of offering a midpoint. Do they believe that “down deep inside” everyone really has an opinion of the matter, or do they believe that some people are genuinely in the middle on the issue and they want to know how many have no opinion?

Not only can nonresearchers make this choice, they often enjoy doing so, and engaging them in thinking about such alternatives and their implications teaches evaluative thinking.
### Levels of Measurement

<table>
<thead>
<tr>
<th>Levels of Measurement</th>
<th>REAL ANSWERS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nominal</strong> data makes distinctions by name. Examples include gender and academic discipline. Analysis of nominal data is typically limited to counts and percentages. Nominal data can aid in analyzing other types of data, e.g., if you wanted to investigate if completion rates vary by discipline. Unlike nominal data, there is an inherent order to <strong>ordinal</strong> data. Examples include educational attainment (e.g., associate’s degree, bachelor’s degree, and so on) and extent of agreement (strongly disagree to strongly agree). However, the “distance” between the points on the scale is not necessarily equal—e.g., “strongly agree” is not twice the amount of agreement of “agree.” Because of this, it is not appropriate to average ordinal data. <strong>Interval</strong> and <strong>ratio</strong> data have equal distances between points on a scale. With interval data, there is no “true zero.” An example is GPA—the distance between a 2.0 and 3.0 is the same as 3.0 and 4.0, but it doesn’t make sense to interpret a 4.0 as “twice as high” as a 2.0. There is a true zero with <strong>ratio data</strong>. Examples include age and credit hours. A student with 60 credit hours has twice the amount as a student with 30. Both interval and ratio data can be averaged. To learn more, see the Research Knowledge Base at <a href="http://www.socialresearchmethods.net">www.socialresearchmethods.net</a> &gt; Measurement &gt; Levels of Measurement.</td>
<td></td>
</tr>
</tbody>
</table>

### REAL QUESTIONS

If you need to know what groups you are reaching with your activities or whether your project is impacting some groups differently than others, then it’s important to ask demographic questions. However, if you can’t articulate a strong rationale for why you need to ask for personal information, don’t do it (see Michael Patton’s comments on p. 1 about aligning survey items with evaluation purposes). Typical demographic questions include those about age, income, educational attainment, gender, and race/ethnicity. Based on my personal experience, I can’t think of a good reason to ask age (in lieu of grade/education level) when evaluating STEM education programs. Income data may be relevant if you are tracking impact on employment and wages. NSF’s interest in broader impacts—one aspect of which is increasing participation of underrepresented groups in STEM—suggests that all grantees should strive to learn how they are addressing this national need. That may mean having to ask about gender and race/ethnicity on surveys. Note that there is a trend toward asking gender and race/ethnicity questions in more sensitive and inclusive ways.

With regard to **gender**, a common way to ask about gender has been to pose the question as “What is your gender?” or “What is your sex?” with the options of **Male** and **Female**. There has been an increasing awareness that these binary categories are not reflective of the real world and could even be offensive to individuals who do not identify with either option. A more inclusive way to ask this question is “What is your gender identity?” with the response options of **Male, Female, Transgender** or **Other _____**.

Not answering should always be an option, and make sure your data management plan has provisions for how the data will be protected, used, and reported. (For more information about how to ask about gender on surveys, go to the Human Rights Campaign website at [www.hrc.org](http://www.hrc.org), select Resources > Transgender.)

Regarding **race and ethnicity**, it makes sense to align your response options with those used typically used in national datasets such as the Census and National Center for Education Statistics: White, Black or African American, Asian or Asian American, Hispanic or Latino, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and Other. Using these standard categories will allow you to compare your outcomes with national trends. But like gender, not everyone can or should be forced into a single box. Instead of asking “What is your race/ethnicity?” you can ask “How do you describe yourself?” Additionally, you can allow respondents to check multiple boxes—or for a more straightforward dataset—offer **Multiracial** and **Other** options. (For more information on race and ethnicity categories—as well as gender, language, and disability, go to the U.S. Department of Health and Human Services Office of Minority Health’s website at [minorityhealth.hhs.gov](http://minorityhealth.hhs.gov), select Data/Statistics > ACA Section 4302.)

**HAVE OTHER QUESTIONS ABOUT SURVEY ITEMS?**

Be sure to catch our next webinar: Developing Questions for Effective Surveys on January 16. We will discuss the **Dos and Don’ts** of survey item construction (see p. 4).
At the conclusion of each EvaluATE webinar, we ask participants to complete a brief survey. Since 2010, we have consistently asked the same two closed-ended questions to gauge perceptions of the webinar. For one of these, participants are asked to indicate the extent to which they agree or disagree with the statement, “I will use the content of this webinar in my work.” Because nearly all respondents (94% or more) across all webinars either agree or strongly agree with this statement, we pay closest attention to variations in the percentage that reflects the most favorable response of “strongly agree.” According to this measure, our best webinars—those with 60% or more respondents strongly agree that they will use the content—are those that have been most closely aligned with ATE PI needs: webinars on proposal development (July 2010, July 2011, August 2012) and orienting new PIs to evaluation (November 2010, September 2011, September 2012).

Evaluator competency frameworks specify the knowledge, skills, abilities, and dispositions that professional evaluators should possess. Various frameworks have been put forth by individuals and organizations around the world. The most established set in North America is the Competencies for Canadian Evaluation Practice, which support the Canadian Evaluation Society’s credentialing program. There are five domains in the Canadian competencies, which overlap significantly with other frameworks: 1) Reflective Practice, 2) Technical Practice, 3) Situational Practice, 4) Management Practice, and 5) Interpersonal Practice. A quick glance reveals the diverse skills required for competent evaluation practice.

When using such a framework to guide the selection of an evaluator, keep in mind that not every evaluator can or should be expected to be fully competent in every evaluation skill and context. Look for an evaluator whose experience and skill fit reasonably well with your evaluation needs and who is open to adding the right person to the evaluation team to fill a skill or knowledge gap.

Practicing evaluators can use an evaluator competency framework to reflect on and guide their own professional growth. Short-term professional development opportunities (from webinars to multiday workshops) are offered through the American Evaluation Association (www.eval.org), EvaluATE (www.eval-ate.org), and the Evaluators’ Institute (tei.gwu.edu). A list of degree programs is available from AEA’s website at www.eval.org/Training/university_programs.asp.

At EvaluATE, we are consulting multiple evaluator competency frameworks as we develop our webinar evaluation curriculum to support evaluation in the NSF-ATE context, relying heavily on Canada’s (see evaluationcanada.ca > Professional Designation) and New Zealand’s (see www.anzea.org.nz > anzea Evaluation Competencies > Evaluation Competencies Project).
Upcoming Webinars

Developing Questions for Effective Surveys

January 16, 2013 | 1-2:30 p.m. ET

The use of surveys for data collection is ubiquitous in evaluation, so writing good survey questions is an essential skill for any evaluator. Grounded in the work of Don Dillman, in this webinar, we’ll cover the essential DOs and DON’Ts of writing survey items. Crafting good questions is an art as well as a science and requires careful attention to context, including respondents, evaluation purposes, and intended use of results. We’ll dissect examples of good and bad question phrasing and response options, explore the implications of various question-and-answer formats for data analysis, and offer strategies to ensure that surveys will yield meaningful and useful data for your evaluation.

From Valuing to Visualization: Data Interpretation and Reporting

March 20, 2013 | 1-2:30 p.m. ET

For evaluation reports to be meaningful and useful, clients and other stakeholders need to be able to see the forest through the trees. In other words, they need to convey clear answers to big questions drawing upon a variety of data sources. To accomplish this, evaluators need to interpret, synthesize, and report results in meaningful and efficient ways. What do all these data points mean? Is the project on track? Even if it is on track, how well is the project doing—is it having an impact? In this webinar we’ll share strategies for transforming data into genuinely evaluative conclusions and presenting results in visually compelling ways to aid in understanding and using results.