

Fall/Winter 2009

A Web Panel Response Rate—It's More Than You Think!

KN takes the lead in developing and proposing an industry standard

By Charles DiSogra



Copyright Nopow - Dreamstime.com

The most misrepresented metric used when reporting Web survey results is the response rate. With Web surveys dominating commercial market research and being used much more frequently by academics and the government, you would think that this response rate metric would be nailed down. In fact, among organizations of survey research professionals worldwide, there is still no agreed upon standard about how to calculate a Web panel response rate. This ambiguity makes comparisons between Web surveys regarding their “response” subject to voodoo interpretation since you may not be able to tell how a rate was arrived at or if two published rates were calculated in the

same way. This confusion should not be the case, because the too-often maligned response rate persists as our initial window on quality when we are evaluating or comparing surveys. While non-response bias rightfully deserves our focus in assessing the validity of any survey, the response rate remains that lone single number we use at the outset of that evaluation. It does not tell the whole story on quality, but it is that all important opening paragraph and thus needs to be calculated correctly.

Knowledge Networks (KN), a pioneer in probability-based Web panel methods, recently published a paper in the journal, Public Opinion Quarterly, on this very topic.¹ KN proposed a method that builds on a calculation approach that we have developed over the years, and it has the attention of survey professionals. Before I go further and describe this method, you may be wondering what the fuss could possibly be about figuring out a response rate. Perhaps you are thinking, what could be so difficult about taking the total number of completed Web questionnaires and dividing it by the total number of panel members invited to take the survey?

¹ Callegaro, M. and DiSogra, C. “Computing Response Metrics for Online Panels.” Public Opinion Quarterly 2008 72(5):1008-1032.

Let's say you have a real nice sample of 1,000 panel members that you invite to your survey and 600 complete it. OK, that's a 60% response rate, right? Not exactly. What you did in that simple calculation is technically called a *completion* rate and not a *response* rate. The completion rate is important to know because it is one component used when calculating a panel-based response rate..., but I'll come back to that.

A rigorous response rate accounting

A response rate requires a sample to be subjected to a more rigorous accounting from beginning to end. This involves sorting out many things, such as the eligibility of non-respondents, bounced email invitations, and incomplete surveys. Another important piece to include in this very complete accounting specifically relates to how panel members got recruited onto the panel in the first place!

Professional organizations like the American Association for Public Opinion Research (AAPOR) publish detailed formulas on how to calculate response rates for random-digit dial (RDD) and mail surveys, but up to today there is no such standard for how to do this for Web panel surveys. So beware, when a Web panel survey reports an undefined "response rate," it is likely a simple completion rate you are looking at. *And if the Web panel is a volunteer "opt-in" panel, it can only be a completion rate.* This is because opt-in panel membership is composed of self-selected volunteers, not initially constructed through a probability-based sampling method. Opt-in panel members, on the basis of how they are recruited, do not have any known

selection probability. As large convenience samples, they are not representative of the national population. So, seeking any sampling accountability for every assigned opt-in panel member relative to the larger US population is a baseless exercise. If it is an opt-in panel, it can only be a completion rate not a response rate. Thus, you are only seeing a piece of the true picture and will forever be misled into believing that you have a better "response" than what may be the real response rate, which is an unobtainable calculation for an opt-in panel.

Probability-based Web panels such as KnowledgePanel® do have all the necessary data dimensions for calculating a response rate for a given study. A KnowledgePanel response rate is the fully transparent accounting that provides that initial window into the quality of your sample. As I will explain, while it is conceptually straightforward to understand, the computational elements can be complex. But it is this conceptual picture that is essentially the key to understanding what might otherwise be an overwhelming, enigmatic response rate formula.

Introducing the components – it is all about time

You have to view a probability-based Web panel response rate in the same way as an astronomer views the cosmos. The lights you see today in the night sky are a snapshot of events that took place at different times in the distant past. In the same way, the assigned panel members who did or did not complete your survey today carry a history that also has to be factored into the response

rate. The individual member's proverbial Big Bang is the point in time when they were first recruited onto the panel. So the first factor to be incorporated into the response rate is the *recruitment rate* from the national sample that the panel member belonged to when that member was first recruited. Whatever the frame/mode used for recruitment (e.g., RDD/telephone, address-based/mail), there are standard AAPOR formulas published to calculate these recruitment rates. As stated earlier, because the opt-in Web panels do not have initial recruitment samples to start with, this essential first-step recruitment rate cannot be calculated for them.

The next point in time in a panel member's history is their follow-through to provide the necessary background information that is their "profile" data. In KnowledgePanel's case, as in the case of some opt-in panels, recruited persons become panel members only after they complete their demographic profile data sheet or "core profile survey" as we call it at KN. Not all recruited persons complete this essential step, so thus at KN we calculate a *profile rate* for a given recruitment cohort. The profile rate is yet another component of that response rate we are working towards. The opt-in Web panels can calculate a profile rate since their potential members who click to join also may or may not follow through and provide their demographic profile data. However, the opt-in panel managers have no control over their initial "come one, come all" recruitment process. For example, some unknown number of persons may click the "join" button by mistake or just out of curiosity with no real intent to join and as a result inflate the denominator and undermine what would be the true

profile rate. Thus, the opt-in panel managers can at best only approximate their profile rate, which can never be as exact as that of a probability-based panel.

As in all membership groups, members leave. On Web panels this occurs when panel members stop participating despite the best non-coercive efforts to encourage or incent retention. This is a natural cycle of panel life that necessitates ongoing recruitment to maintain and refresh the membership. Thus, we must account for yet another component in our construction of a response rate – the *retention rate*.

Starting with a single probability-based recruitment sample, we have so far calculated that sample's recruitment rate, profile rate, and now its retention rate. Each member of that sample who remains active on the panel carries these three rates with them at the time they are selected for a given study's sample. When a sample for your research study is drawn from a Web panel, that sample is sent an invitation to take and complete your survey. And, voila! We have at last arrived at the final piece, your study's *completion rate*. Although this rate is further adjusted with other factors (such as accounting for eligibility, break-offs, email bounces/undeliverables, etc.), for all practical purposes, it is ultimately the number of completes divided by the number of eligible invitees in the sample.

Proposed standard – the cumulative response rate

The response rate for probability-based panels can now be calculated as a cumulative product of all the above

mentioned rates. After independently calculating each of the four rates (recruitment, profile, retention, completion), we would then multiply these together to produce your study's response rate. Although a more simplified option would exclude the retention rate (we call this lower option Cumulative Response Rate 1, or CURR1), we believe that the more complete picture would include the retention rate (called Cumulative Response Rate 2, or CURR2). These two options are shown in Figure A (*see page 5*). Table 1 summarizes which component rates and cumulative response rates can and cannot be calculated for probability-based Web panels and opt-in Web panels.

The mathematical consequence of a cumulative product is that it produces a lower and lower rate each time you add another component (such as, half of a half and so on). This is the world of Web panel response rates where the rate turns out to be a lower percent than, say, what you would get in a cross-sectional random-digit dial population survey. Take the example where you have a 40% recruitment rate, a 60% profile rate, an 80% retention rate and finally a 70% completion rate with your study's Web panel sample. The response rate would be 13.4% obtained by multiplying $0.4 \times 0.6 \times 0.8 \times 0.7 \times 100\%$. You can see right away why some might prefer to mislabel their completion rate as their response rate, since 70% sounds so much better than 13.4%. The reality is that 13.4% is the true response rate, and in this example, it is actually pretty good. It is well within the range of Web panel cumulative response rates.

At Knowledge Networks, historically, we have employed ongoing panel recruitment with multiple independent RDD national samples and more recently address-based national samples fielded throughout each of many years. Therefore, it would not be improbable in a single study sample to have panel members coming from some 200 different recruitment cohorts. In this case, we employ complex computer programs that factor in these 200 or more cohort-specific recruitment rates, profile rates and retention rates to arrive at a weighted Web-panel response rate for each unique study sample. This is the type of cutting edge statistical product that we use at KN and one which we think should be the industry standard for all studies in this rapidly growing and evolving field of Web-panel survey research. As the number of Web surveys proliferate, it becomes crucial that we have a standardized, reliable and transparent measure of response rate on our dashboard for assessing survey quality.

Dr. Charles DiSogra is Knowledge Networks' Chief Statistician and heads the statistics unit at the Menlo Park office. He brings over 20 years experience in survey research, sample design, data analysis, and administration. Dr. DiSogra has a masters degree in public health and a doctorate in nutritional epidemiology with an emphasis in biostatistics and policy analysis from the University of California at Berkeley.



Figure A. Cumulative response rate calculations for Web panels

<p>Option 1 – the simpler (less complete) version:</p> $CURR1 = \text{Recruitment Rate} \times \text{Profile Rate} \times \text{Completion Rate}$ <p>Option 2 – Knowledge Networks proposed standard:</p> $CURR2 = \text{Recruitment Rate} \times \text{Profile Rate} \times \text{Retention Rate} \times \text{Completion Rate}$

Table 1. Rates that can be calculated for two types of Web Panels

Rate Metric	Probability-based Panel	Opt-in Panel
Recruitment	Yes	No
Profile	Yes	Approximate
Completion	Yes	Yes
Response (Option 1)	Yes	No
Response (Option 2)	Yes	No