

To double-code or single-code?

In quantitative marketing research, we typically ask open-ended questions in the surveys we conduct, and record what the respondent says verbatim. Then what?

We do not report what is said exactly (the verbatim), because that would generate a very, very long report. What happens next is that the verbatim is coded. Coders group similar responses, and create summaries of what the consumers say.

For example, suppose we are conducting an interview about air fresheners. The respondent is asked "What do you like about it?" A random selection of verbatims from the question are examined, and code numbers are assigned based on the similarity of various statements. For example, one person says, "I like the smell, because it is sweet.", and another says, "I like the sweet smell, and the color." Both would get a "sweet smell" code, and the second one would also receive a "like the color" code.

These coded responses are then assigned to categories. In this example, the "sweet smell" comments may be assigned to a Smell/Odor category, and the color comment would likely be assigned to a category called Appearance.

One person can make two comments that fall into one category. Say what? For example, if a respondent said, "I like the smell because it is sweet and subtle," the code for "sweet smell" would be assigned, as would a code for "subtle smell". Since we want to count people, not comments, the category net is created that tells us how many people made a Smell/Odor comment. The person who said "I like the smell because it is sweet and subtle." would be counted once in "sweet smell", once in "subtle smell", and only once in the

Comments	Total Respondents	
	#	%
Base: total (250)		
<u>Smell/Odor (Net)</u>	<u>65</u>	<u>26</u>
Sweet smell	55	22
Subtle/quiet smell	32	13

category net line, because it counts people, not comments. The net line will always be equal to or less than the

Editor's note: Stephen J. Hellebusch is president of Hellebusch Research & Consulting, Inc., Cincinnati, OH. He can be reached at 800-871-6922 or at info@hellrc.com

sum of the number making the comments in the category. For example, the two comments, as shown in the box, add to 35 percent, but were given by 26 percent of the 250 respondents. If the net is not less than or equal to the sum of the comments, something is wrong.

One challenge, handled in different ways by different researchers, is a comment that logically fits into two different categories. Suppose, with our air freshener, someone says they like it because, "The size makes the placement easier." There is Placement net which includes comments like "would place in family room" or "would put in bedroom". There is also an Ease of Use category that includes codes like,

Double -coded		
Comments	Total Respondents	
	#	%
Base: total (250)		
<u>Ease of Use (Net)</u>	<u>63</u>	<u>25</u>
Just plug it in	25	10
Size makes placement easier	20	8
Easy to remove from package	18	7
<u>Placement (Net)</u>	<u>60</u>	<u>24</u>
Put in family room	25	10
Size makes placement easier	20	8
Put in bathroom	12	5
Put in bedroom	11	4

"easy to remove from package" and "just plug it in".

Some coders "double-code" -- that is, put the comment both places. The 20 people who said "Size makes

If the research user is thinking of the categories as mutually exclusive with respect to the comments

Single -coded		
Comments	Total Respondents	
	#	%
Base: total (250)		
<u>Ease of Use (Net)</u>	<u>63</u>	<u>25</u>
Just plug it in	25	10
Size makes placement easier	20	8
Easy to remove from package	18	7
<u>Placement (Net)</u>	<u>40</u>	<u>16</u>
Put in family room	25	10
Put in bathroom	12	5
Put in bedroom	11	4

placement easier" are the same 20 people in both the Ease of Use and the Placement categories.

Others will single code -- that is, only include the comment once, in the category judged to be the best fit. In this case, the comment fits better in Ease of Use, since the size is what is making things easier, and all of the Placement comments are referring to a place.

So, what percent made Placement comments? It depends. If you double-code, you would say 24%. If you single code, you would say 16%. Neither is right; neither is wrong, technically. If many comments are double-coded, results will look very different than if only single-coding is used.

(usually an implicit assumption), then that user will take an incorrect understanding from the double-coded question. It is likely that this will occur when the user is given a presentation that summarizes the research where only the category net lines from the open-ends are shown. There is no way to know from just the net lines that the same comment is in two category nets.

If you like to understand what you are looking at, it may be helpful to learn whether the open-end responses are double-coded, triple-coded, or (perhaps) single-coded. It may make a big difference in how results from an open-ended question are understood. | Q