



# Forget gamification; try writing a humanized survey

| By Annie Pettit

Despite the plethora of qualitative and quantitative methodologies, surveys remain the No. 1 source of data for marketing researchers. They offer us the opportunity to collect opinions about predetermined topics using standardized questions and answers, thereby avoiding creating biases that impede the generalization of research results.

Myself included, most professional survey writers have gotten stuck in a rut. In our quest to be as accurate and precise as possible, we have chosen to write surveys that are so clear and specific that they are also ultra-boring and difficult to read. Who doesn't recognize the choice and style of words here:

Based on any advertising, such as TV commercials, roadside billboards, banners on Internet Web sites, or other similar forms of advertising, please list in the text box below the products or types of products that you yourself have seen publicly advertised between January 1, 2013 and December 31, 2013. (Please type your answer in the box below.)

Surveys with such formal and impersonal wording don't have a particularly positive impact on responders, nor do they excite and encourage someone to happily complete a survey. Surely our failure to engage responders with interesting questions means that we are at least partly to blame for the ongoing decline in response rates and data quality.

There are many ways to encourage more engagement with surveys. Gamification is a trendy one, though researchers without extra funds can't partake in this solution. It isn't cheap to translate a research objective and the associated survey ques-

tions into a fully-gamified, point-collecting, badge-earning, hide-and-seek game.

A less-extreme option would be to incorporate a series of fun tasks within a traditional survey, perhaps adding little written and visual features such as word-finds, puzzles or jokes in between questions. But must we go to such creative steps? Must we add to and change the appearance of a survey to generate engagement?

Why can't we simply write better surveys? And by better surveys, I don't mean even more grammatically-correct and precise. I mean why can't we use language that is simpler and more reflective of who we are today, language that reflects the current trends and styles we have of communicating with our fellow human beings? Sure, this language includes theoretically incorrect grammar and slang, but that's, like, how real people talk to each other every day, you know? Researchers could easily stop using business language and start using human language if they so desired.

## Beneficial on all fronts

This article intends to demonstrate that writing surveys in a humanized fashion is beneficial on all fronts. First, we aim to show that the data quality of humanized surveys does not suffer; that speeding, straightlining and red-herring rates are similar to or better than those of traditional surveys.

Second, we aim to demonstrate that responders enjoy participating in humanized surveys, more than they do traditional

surveys, and that this will be evidenced via survey behaviors and overt responder ratings.

Third, and finally, we know that the humanized survey will generate different raw scores than the traditional survey, as does any test/retest of an identical survey with identical responders. We know that results differ when an answer option is added or removed; a word is added or removed or changed; the formatting is altered; or a picture is added or removed or changed. That is not under dispute here. However, we will demonstrate that the action outcomes generated from the humanized survey are the same or similar to those generated by the traditional survey.

## Method

Respondents were sourced from the Peanut Labs online research panel, a non-probability incentivized source of prescreened adults who have agreed to participate in online surveys. Test (n = 530) and control (n = 532) groups were matched on U.S. census-rep demographics (age, gender, ethnicity, Hispanic, region).

The research was conducted with a non-branded survey about computer- and video-gaming opinions and behaviors. The survey was first written using traditional survey wording and style. It was then copied into a second version where about 90 percent of the questions were rewritten in a more casual and humanized style. Some of the revisions were very slight, incorporating only the addition of a few words, whereas other revisions were complete rewrites. Some rewrites ended up in questions or answer options that were much longer or shorter than the original version. Regardless of whether the question and answer at hand was revised, responders in the test group would have been exposed to all of the previous humanized rewrites. Table 1 provides several examples of matching questions from both versions of the survey.

## Results

### Data quality

Survey completes that include poor quality data have no effect on survey results as those identified data points are removed and replaced with good-quality data. However, for every survey complete that fails the data quality requirement, another

complete must be purchased at additional cost. The real problem with poor-quality data, therefore, is financial and this is an important problem to factor in.

Five data quality measures were designed to identify data that did not conform to what was expected.

**Straightlining:** Failed if data demonstrated straightlining to a half-positive, half-negative eight-question grid.

**Open-ends:** Failed if less than 10 characters were provided in response to a request for “Three reasons why...”

**Speeding:** Failed if the survey was completed in the slowest 5 percent of completion times.

**Red herrings:** Failed if two fake computer games were selected.

**Acquiescing:** Failed if 14 or more items in a list of 16 were selected

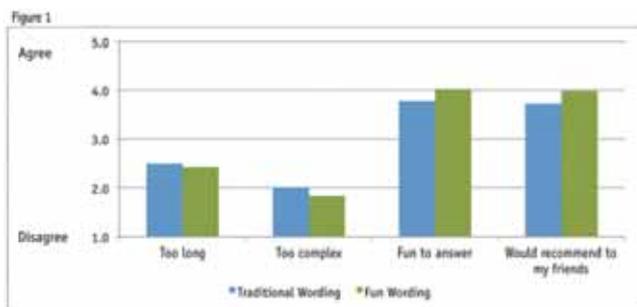


Figure 1 shows that the humanized survey generated slightly more favorable data quality results for the straightlining, open-ends and speeding measures. Additionally, the humanized survey generated a lower overall failure rate. These results suggest that the humanized survey did not create data quality issues but may have in fact improved data quality.

After the data quality analysis, data from those who failed three or more of the measures were ignored. Anyone generating two or fewer failures was deemed a good-faith responder and their data were retained for the subsequent analyses.

### Responder engagement

Surprisingly, the humanized survey generated partial-complete rates that were relatively low but still 68 percent higher than those of the traditional survey (10.9 percent vs. 6.9 percent). Thus, among people who started the survey – and therefore learned the topic and viewed the question styles – more people in the humanized survey ended up not completing the survey. Perhaps not every responder appreciated the style of words and humor used and perhaps some responders felt that the humor meant the survey was not a serious endeavor. Or this was simply a random-chance result.

Open-ended questions are another opportunity for responders to demonstrate survey engagement. The humanized survey averaged slightly longer open-ends than the traditional survey (25.7 vs. 22.9 characters).

Finally, for those responders who did complete the full survey and subsequently also completed the “responder perception of survey” section, the humanized survey fared slightly more positively. It generated more positive scores for fun, recommending and complexity. Responders perceived the length of the two surveys to be the same.

Table 1: Matched Questions From the Traditional and Humanized Surveys

Traditional Survey	Humanized Survey
Thank you for participating in this survey. Your careful and honest opinions help improve existing products and create brand new products.	Thanks a bunch for answering this survey! May the survey force be with you.
Which of the following gaming systems... 1) ...do you currently own? 2) ...have you used in the past 3 months? 3) ...have you used in the past week?	Which systems do you... 1) ...own? 2) ...use? 3) ...use a lot?
What is your favorite video/computer game?	If you could only play one video game for the rest of your life, what game would you choose?
Have you ever lined up at a store so you could be among the first to purchase a new video/computer game?	Are you one of those people who lines up at stores to be the first to buy a new game?
How much do you agree or disagree with these statements about gaming?	Love 'em or hate 'em, what do you think about video games?
What is your opinion about each of the following television shows?	Awesome or terrible, how do you rate these TV shows?
Is there any other information you would like to share with us regarding your opinions and use of video/computer games?	Sharing is caring! Is there anything else you'd like to share about gaming?

Table 2: Verbatim Comments About the Survey Experience

Traditional Survey	Humanized Survey
It was short and nice:)	nice short survey not enough points though
I found this to be an interesting survey. Thank you.	I love surveys and quizzes and this one was very interesting, because i like sharing things about myself.
the survey should be worth more points considering how long the survey was	very original method of conducting a survey?
NA	This was one of the Best surveys I have done!!!
No	loved it.. funny and quick
Good	"May the survey FORCE be with you." Was a nice hilariously awesome thing to see. I laughed a lot. o3o
This survey was enjoyable.	this is an awesome survey. I thoroughly enjoyed myself while taking it.
I don't know.	good job, sir! If I had a fedora, i'd tip it in your direction.

Finally, when it came to the sentiment of verbatim comments, the humanized survey easily outperformed the traditional survey. While both surveys generated a similar number of responses, the sentiment of responses towards the humanized survey outright praised the survey designer for providing a unique and enjoyable experience (Table 2).

### Action outcomes

As researchers, we know that if the same survey is launched to the same group of people twice on the same day, the results will be different. As such, results from surveys are always shared in combination with other data so that readers can make their own informed judgements about any differences they see. It comes as no surprise, therefore, that the raw results from the traditional and humanized surveys are different. Given that each data set has several hundred datapoints, the laws of statistics demand that a certain percentage of the datapoints be different, simply to account for chance.

However, it is not the raw datapoints that necessarily matter but rather the end conclusions and the resulting action outcomes. Regardless of the raw frequencies and scores discovered, for researchers to feel confident in taking liberties with question design, the final conclusions from both surveys must lead to the same or similar action outcomes.

Four questions having actionable outcomes were selected for analysis.

### Actionable result No. 1: likelihood to purchase

#### Traditional question

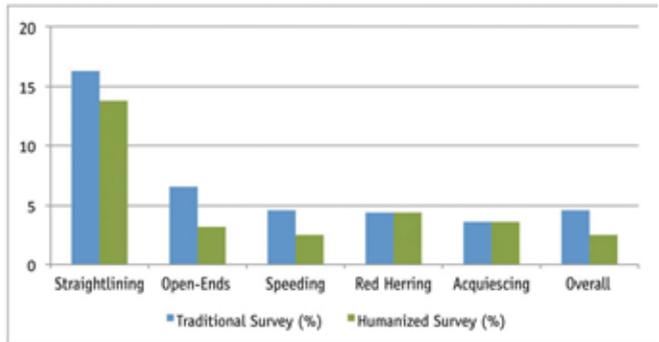
Once you've purchased a video/computer game, how likely are you to also purchase other releases from the same series (e.g., Halo 1 and Halo 2)? (Please select only one) Not likely at all, Somewhat likely, Very likely, Extremely likely

#### Humanized question

Are you likely to buy other releases of a game you've already bought? (e.g., Halo 1 and Halo 2) (Please select only one) Not likely at all, One is good enough for me!, Somewhat likely, Very likely, Extremely likely, I need them all!

There was little to no difference in the actionable outcomes from the two survey versions. The raw scores were expected to be and were different. However, the magnitude of each pair of numbers was very similar and the top- and bottom-ranked

Figure 2



answer options were the same. The only difference was the order of the two mid-ranked answers. The final conclusion drawn from these data, regardless of which survey version used, was that about 17 percent of gamers are very likely or 42 percent of gamers are very or extremely likely to purchase more than one release of a game.

### Actionable result No. 2: fair prices

#### Traditional question

What is a fair price to pay for a video/computer game? (Please select only one)" [Sole difference in response options] \$0

#### Humanized question

What is a fair price to pay for a video game? (Please select only one) [Sole difference in response options] \$0 They should be free!

Results for this question also suggest that little to no difference in the actionable outcome would result. Again, though there are expected differences in the raw data, the relative magnitudes were similar, the four most-frequent choices shared the same rank orders and the lowest-ranked choice remained the same. The end decision for both surveys was that a fair price for games is in the range of \$20 to \$29.

### Actionable result No. 3: advertising approval

#### Traditional question

How much do you approve or disapprove of these kinds of advertising in video/computer games? (Please select one in each row), Somewhat disapprove, Neutral, Somewhat approve, Strongly approve

#### Humanized question

Love 'em or hate 'em, ads are everywhere. Do you approve or disapprove of these? (Please select one in each row), Somewhat disapprove, Neutral, Somewhat approve, Strongly approve

Again, the two surveys generated very similar action outcomes. Among seven questions about ad locations, six generated raw scores that were within three points of each other. The remaining item differed by seven points. Because five of the items generated top-box scores that were all within confidence intervals of each other (i.e., the race was too close to call), rank orders could be reliably differentiated within or between groups. In other words, both survey versions came to the same conclusion.

### Actionable Result No. 4: game preference

#### Traditional question

Which of the following two video/computer games would you prefer

to play? (Please select one.) [Repeated six times]

### *Humanized question*

If you could play only one of these two games for the rest of your life, which one would you choose? (Please select one)

5 more... What if you had to pick between these two games? (Please select one)

4 more... And between these two? (Please select one)

3 more... And between these two? (Please select one)

2 more... And between these two? (Please select one)

And lastly, between these two? (Please select one)

In all six choices, the preferred game remained the same though the raw scores between the two survey versions differed within a range of 1.6 to 7.6 points, comparable to expected error rates. In the end, the rank order of the preferred game remained stable except for a swap between the games ordered two and three. We must conclude that the humanized survey did not generate different results from the traditional survey.

### Many reasons

As researchers, we can find many reasons to keep on writing surveys in the traditional fashion: The grammar is more correct. The wording is more precise. There is less chance of misinterpreting what is required of the responder. We are used to this style of writing. The last 11 surveys for our monthly tracker were written exactly the same way.

But there are many reasons to completely rethink our processes and write surveys for human beings. Surveys have to compete more often than ever with social networks like Facebook, Twitter and YouTube and online games like Grand Theft Auto, FarmVille and Angry Birds. These online activities have been treating our research responders like human beings since Day 1.

Survey response rates and survey engagement continue to decline while we continue to moan about the time and

expense required to completely reformulate a boring survey into some kind of stunningly exciting 3D mash-up of Pokémon and Call of Duty.

The results of this research show that writing a humanized survey – a more fun, casual and current survey – is a good thing. Data quality does not have to suffer and may even improve. Survey engagement does not suffer and may even be better. Research results may not be strikingly different and the action outcomes can be very similar.

Making a better survey doesn't have to take more time from even more programmers and cost even more money. These results have shown that the cost of writing a better survey may simply be the cost of injecting a little bit of creativity and personality into your word choices. If you are ready to do something, anything, to improve survey engagement and satisfaction, do this. Humanize your surveys. For free.

And for those who still question the differences in the action outcomes, and the leniency and vagueness introduced by easing up on question specificity, I offer this: Who's to say what the "right" question is? Would the question have generated more valid results if the answer anchor was "Very" instead of "Extremely"? And what if the question said "Do you own" instead of "Do you currently own?"

We know that changing a single word generates different results; that sending the same survey to a different group of people at the same time generates different results; and that resurveying people just hours apart generates different results. Let's understand the intent of our statistical and methodological techniques instead of focusing so blindly on raw data.

Let's humanize our surveys. 

Annie Pettit is chief research officer at San Francisco research firm Peanut Labs. She can be reached at [annie@peanutlabs.com](mailto:annie@peanutlabs.com) or 416-273-9395.