White Paper

Questionnaire Length and Fatigue Effects: The Latest Thinking and Practical Solutions

By Pete Cape, Director, Global Knowledge Management and Keith Phillips, Senior Methodologist

April 2015
ABOUT THE AUTHORS

Pete Cape, Director, Global Knowledge Management
With over 25 years of experience in market research and originally a specialist in international telephone research, Pete was a founder of TNS Interactive and has concentrated on online research ever since. A frequent speaker at industry events around the globe and a regular contributor to industry publications, Pete was recently one of five experts who worked on the ESOMAR/GRBN Guideline on Online Sample Quality.

Keith Phillips, Senior Methodologist
Keith's role at SSI includes conducting primary research projects, helping clients with the research issues they face on a day-to-day basis, training colleagues, and working to support companywide sampling initiatives. Prior to joining SSI in March of 2010, Keith was a Senior Research Manager in the Motion Picture Division of OTX Research, which he joined in 2004. Keith is a regular speaker for the American Marketing Association (AMA) webinar series. He has presented webinars for the ARF, AAPOR, ESOMAR, and Quirk's.
To obtain high quality research data we need people to pay attention throughout the questionnaire. But what if they are unable to do that? What if it is not possible for an average participant to pay attention throughout an entire questionnaire? How long is too long for today’s research questionnaires?

This paper examines the effects of questionnaire length on data quality, discusses the implications for questionnaire length in a mobile world, describes practical solutions for reducing questionnaire length, and shares tactics for increasing attention throughout the questionnaire.

UNDERSTANDING THE NATURE OF HUMAN ATTENTION SPANS

A number of studies have been done\(^1\) to gain insight into “TUTS” – Tasks Unrelated Thoughts – and to demonstrate the level of consistency between simultaneous thought and action.

Respondents to one such study were sent throughout the day pager messages asking two simple questions: What are you doing? and What are you thinking about? 30% of the time there was a mismatch between what people were doing and what they were thinking about. Most people can relate to this concept since in our daily lives we are often on “autopilot” doing one thing and thinking about something quite unrelated. Therefore these results may seem unsurprising.

Perhaps more enlightening on this topic is the SART or Sustained Attention to Response Task in psychology. In this study people are asked to pay attention and do a task. They are shown a number on a screen and have to click a key on their keyboard. However, if the number three is shown, they need to refrain from clicking at all. The test consists of about 225 trials. The average error rate of 3%-4% is quite surprising to some. Participants in the test are trying to concentrate on a task that is not very difficult, but is repetitive – perhaps something like answering grid questions in surveys – and yet they fail 3%-4% of the time. (Several versions of this test are available online to try, via searching on Sustained Attention to Response Task.)

The average attention span of an adult is estimated to be around 20 minutes\(^2\). This is the amount of time adults can sustain attention for a task that they freely choose to do. We hear a lot of talk about the reduction of attention spans because of the internet and it is certainly true that we choose to pay attention to things for less time than we used to perhaps. For example, we choose to watch 30 second video clips on YouTube rather than sit down for an hour and watch a TV show; but that is a free choice. The brain is still wired enough that we could pay attention to 20 minutes of something if we wanted to.

RELATING IDEAS ABOUT INATTENTION TO RESEARCH QUESTIONNAIRES

How do these ideas about attention from psychology relate to attention during survey taking? Which type of inattention is most relevant to the survey-taking experience?

One relevant type of inattention can be demonstrated by a simple video test online involving basketball players\(^3\). This experiment is well worth trying online with a friend or colleague. If you wish to try it yourself, do so before reading on further, since the “secret” to the test is revealed in the next paragraph.
In this inattention experiment the viewer is asked to look at a video in which some people dressed in black and some people dressed in white, pass a basketball around. You are asked to concentrate on the people dressed in white and count how many times they pass the ball. During the video, a figure in a gorilla suit wanders into the scene, stops, does a short dance for a few seconds then wanders “offstage.” Many people who view this video never see the figure in the gorilla suit even though it is there in plain view. The gorilla suit example illustrates a type of inattention relevant to survey research, in which in an attempt to concentrate on the task at hand a person can easily miss something that is obvious. This is a reminder of the importance and impact of questionnaire design.

Another type of inattention is the “daydreaming” of a student in a math class who is simply staring out of the window. This type of inattention is very relevant to long questionnaires, in which a respondent pays attention for a while, but at a certain point can no longer pay attention. SSI’s research has demonstrated this same type of inattention in survey questionnaires over more than a decade, and across many countries and cultures.

**THE SSI INATTENTION EXPERIMENT**

The experimental setup for this research is very straightforward. We tested two different surveys: one longer version and a shorter version of the same questionnaire. Each questionnaire consisted of four blocks of questions around different subject matter. These question blocks were rotated so that we could clearly see if there were differences between answering questions early in the survey experience vs. later in the survey experience.

The study has been done three times now, first in 2004, in UK, France and the Netherlands, with results shared in an award-winning paper at the 2005 ESOMAR Panels Conference. The experiment was repeated in 2009 for a presentation at the ARF Conference in New York, and again in 2015. The 2015 wave covered new countries: Singapore, China and India, to see if the same results founds in earlier tests would also be found in Asia-Pacific.

**RESULTS FROM THE SSI RESEARCH**

The first finding is that longer surveys do not lead to increased dropout rates. Below are the dropout rates achieved in the three studies for the long and short versions of the survey.

**2004, 2009 and 2015 drop rates, long and short surveys**
Dropout points, long survey, 2004, 2009 and 2015 studies, by country

If questionnaire length alone was the driver of drops the lines would be in the shape of a J, starting very flat, then steeply rising. In fact, the lines are straight overall, indicating that the level of dropouts is constant throughout the questionnaire. The reasons for this are many: the baby cries, the doorbell rings, the survey crashes, the PC crashes. These are random events and tend not to be related to the survey length. Most surveys do not encourage people to take a break and people generally do not take breaks: they continue to the end of the survey.

The chart below shows the percentage of people who are outliers in the 2015 test by time. If some people did take a break during the longer survey we might expect to see more outliers when we examine the time taken to complete the survey, for while the break is being taken the survey timer is still running. The data below clearly shows that we do not see differential rates of outliers. Whilst not proof positive that breaks are not being taken it does suggest that respondents tend to ‘press on’ with their survey.

Percentage of people taking breaks, 2015 survey by country, long and short surveys

<table>
<thead>
<tr>
<th>Country</th>
<th>Long Survey</th>
<th>Short Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singapore</td>
<td>6%</td>
<td>4%</td>
</tr>
<tr>
<td>China</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>India</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Average</td>
<td>3%</td>
<td>3%</td>
</tr>
</tbody>
</table>

The level is not dramatically higher in the long survey compared to the short survey, indicating that people are not giving themselves a break in a long survey. This is not surprising since people who join panels enjoy taking surveys and will keep going until they have completed the survey. They may get fatigued but the fatigue they experience is mental, not physical fatigue and is not something they are aware of as they take the survey.

In summary, our research shows that there has been almost no change over the past 11 years in these
behaviors. The differences we see over the years are due to changes in the online survey-taking environment. The most significant of these are advances in technology.

**IMPACT OF TECHNOLOGY ON SURVEY COMPLETION BEHAVIOR**

Advances in internet connection speed over the past decade have played a very important role in the length of time it takes people to answer survey questions online. When SSI’s research began in 2004 a large proportion of participants, perhaps about half, would have had a dialup connection. By 2009, a smaller percentage of individuals had a dialup connection and of course today that number is very small indeed.

The impact of faster Internet speeds is demonstrated in the chart below:

*Average minutes taken to complete long and short surveys, 2004, 2009 and 2015 studies*

In 2004, the long survey took about 33 minutes to complete and in 2009, it took around 24 minutes. Although the 2015 survey took somewhat longer than in 2009, the data is from a different set of countries, not allowing a direct comparison. However, if we did field this study in 2015 within the same countries we had previously, we might hypothesize that it would take slightly longer to complete, because there are more mobile survey takers online than there were previously. It takes longer to complete a survey on a mobile device compared to a PC because respondents are using a smaller screen, where it is difficult to select buttons and more difficult to type in answers to open end questions.

An additional factor is how well the experience of the survey matches the device being used to complete it. If the survey has not been optimized for that device, abandon rates will be higher and completion time will be even longer as the respondent struggles to manipulate the screen to answer the questions. In the charts below we see abandon rates by device and median length of interview from research done by SSI in 2013 in which a mobile unfriendly survey, mobile friendly survey and mobile optimized survey were compared.
The chart below shows that even when the survey is optimized for mobile, it takes longer to complete on a mobile device than on a PC or laptop.

**IMPACT OF QUESTIONNAIRE LENGTH ON THE DATA**

SSI’s 2004 research clearly showed that people speed up during a long survey. In the chart below each line represents the different question blocks and the length of time taken to complete them by their position in the long survey questionnaire:
In every case more time is spent on the questions when they are asked earlier in the questionnaire than when they are asked later. In contrast, people completed the short questionnaire at an even pace:

Interestingly, in the long survey we saw the increased speed happening as early as the end of the first block of questions, but this was not seen in the short survey. We would hypothesize this has to do with the length of the survey blocks themselves. At the end of each block participants are given a mental break and a moment to reset. Expectations are being set for the length of the next block as well. In 2009, we saw a much more even pace in terms of survey taking across the life of the long survey:
The main reason for this may be that the overall survey length has decreased because of the increase in connection speed. The short survey from 2009 resulted in the flattest line among all the charts. In 2009, the short survey was completed in a very even pace throughout:

**Length of time taken to complete question block in minutes by position of block in questionnaire, 2009, short survey**

In 2015 we see a little more variance in the data in the long survey:

**Length of time taken to complete question block in minutes by position of block in questionnaire, 2015, long survey**
One important point to make about the long survey in 2015 is that the average length of time taken to complete it is 23 minutes. This happens to be the average survey length across SSI surveys in 2015. In other words, what we considered a long survey in 2004 has now become the average survey length. Far from making questionnaires shorter following the results in 2004 demonstrating that quality suffers on long surveys, surveys in the industry just continue to get longer.

In 2015, the shorter survey continued to show the same pattern as the 2004 and 2009 studies, with a more consistent pace of survey completion relative to the longer survey:

**Length of time taken to complete question block in minutes by position of block in questionnaire, 2015, short survey**

![Graph showing length of time taken to complete question block in minutes by position of block in questionnaire, 2015, short survey](image)

**IMPACT OF SATISFICING ON DATA QUALITY DEMONSTRATED**

What these results show is satisficing behavior. Satisficing is the opposite of optimizing. Respondents are doing just enough to satisfy the task. The result is data which does not look very wrong, but is not optimal. SSI’s research clearly demonstrates the impact on the data from satisficing behavior via several measures tested.

One of the symptoms of satisficing tested by SSI was instrument non-use. We gave respondents a slider with a central start or “default” position indicated. Respondents were not required to move the lever on the slider and could continue to the next question without doing so. The chart below shows that, as the position of the sliders moves back later and later within the survey, more and more and more people hit the “next” button without touching the slider lever, deciding that “three” or the midpoint was a good enough answer for them:
Rates of slider non-use by year, long survey

The true answer can probably be seen in position one. In the short survey, the same behavior was seen – respondents continue at an even pace and over time start to show satisficing behavior:

Rates of slider non-use by year, short survey

Similar behavior was seen in open questions:

Number of characters entered in open questions by position of question in survey, by year, long survey
As the open question moves further and further back in the survey, fewer and fewer characters are entered. In 2015, very few characters were entered in general. The explanation for this is partly that the survey was being done in Southeast Asia. In China for example, with pictograms, the character count may be low even when there is substantial content in the answer. Nonetheless, the data still shows an overall downward trend in character quantity as the survey progresses, a trend which is not as pronounced in the short survey:

**Number of characters entered in open questions by position of question in survey, by year, short survey**

![Graph showing number of characters entered in open questions by position of question in survey, by year, short survey.](image)

The next test reveals behavior which goes beyond satisficing into “cheating”, through respondents disqualifying themselves and therefore being able to skip question sections. We included a section in the questionnaire on short break holidays and asked: “Have you ever taken a short break holiday?” This is a simple yes/no question and should have been utterly consistent across all positions in the survey. The chart below shows that it was not consistent, declining quite dramatically as the survey progresses:

**Rate of having taken a short break holiday by position in questionnaire by year, long survey**

![Graph showing rate of having taken a short break holiday by position in questionnaire by year, long survey.](image)

Depending on the position of the question, the reported rate of participation in short break holidays in Europe varies from about 60% to below 50%: quite a substantial difference in marketing terms. This demonstrates that respondents are behaving differently as the survey continues in a long survey situation. Over time in the questionnaire participants have learned that saying “yes” to a question will cause them to receive more questions. In a short survey, in contrast, we see only a very slight decline:
RESEARCH CONCLUSIONS

The original research, done by SSI’s Sandra Rathod, concluded that 20 minutes was the maximum ideal questionnaire length, and we should try to make surveys shorter than that. In 2009, we concluded that the long survey was still too long. Even though its physical length was shorter, the long survey was still producing poor quality data near the end. The 2009 long survey was only 23 minutes long. It should be within the attention span but is not. Again, 23 minutes is the average survey length seen by SSI in 2015.

Our conclusion, looking back at this data in hindsight, is that cognitive length is an issue, not just physical length. Even though improved technology allows respondents to get through a survey faster than they could in 2004, they are still being asked too many questions.

PRACTICAL WAYS TO REDUCE COGNITIVE – AND PHYSICAL – QUESTIONNAIRE LENGTHS

The reality of the research business is that our client often wants 40 minutes worth of data; they have 40 minutes worth of questions that require an answer. Given 40 minutes of questions, as we have seen, the respondent will dutifully answer them. Respondents do not become too physically tired to continue, and they like answering questions, so will keep going. But we have demonstrated that the data suffers. This could be serious, as seen in the “short break holiday” example in which the market penetration varied substantially.

Knowing that physically and cognitively long surveys damage the data, how can we reduce the burden on the respondent?

1. **Become a data curator:** Ask when constructing a questionnaire: Did we ask this question before? Is there data available that could be used instead of re-asking a question? Even if the study is not a tracker, there may be previous ad hoc studies which asked similar questions.

2. **Question the usefulness of every data point:** When creating a questionnaire, ask: What does this question tell us? How will it be used in analysis? Each question added towards the end of the survey will diminish the quality of answers to other nearby questions. Has the question ever changed? Have we been asking this question over and over again and getting the same response? Does it remain in the questionnaire “just in case” we perhaps happen to need it?
3. **Consider asking some questions to a sample subset only:** Does every respondent always have to answer every question? Could a subset answer some question blocks?

4. **Rethink what is being measured:** In an attitude battery, what are we trying to measure? For example, in an image section for a luxury car brand study we may measure beliefs, aiming to understand whether people perceive a particular brand as safe, durable or luxurious. As we ask about more and more of these attributes and descriptions, people are more and more likely to start responding with their attitude towards the brand. If they really like the brand, they will answer positively about all the different attributes presented to them. The battery, which is trying to measure different beliefs is in reality uncovering a latent construct. Therefore in this example we could have used factor analysis to remove similar attributes. This is a good example of not only reducing the physical size of the questionnaire, but also its cognitive burden. This type of content reduction leads not only to better data for the survey as a whole, but better data within the single question itself. It is important to remember that we are not just concerned with survey length overall, but also about the length of individual questions and how that cognitive burden can play a role in the question itself.

5. **Ask for help from the sample provider:** The sample provider may have information, which can help reduce the number of questions asked, such as profile information and demographics.

6. **Break the survey into pieces:** We have seen that respondents today are willing to keep going through long questionnaires. But, with the advent of mobile, can we rely on their willingness to do so in the future? Data about respondent behavior on SSI’s proprietary panels\(^5\) shows that mobile respondents are usually using their mobile devices to take surveys at home rather than on the go. In the future, this may not be the case and short bursts of survey activity may become the new norm in tomorrow’s mobile environment.

There are two ways of approaching the idea of breaking up or “modularizing” surveys: SSI, in partnership with Gongos Research, has tested each of these and named them “Cyborgs” and “Monsters”\(^6\). In the SSI test, three short survey modules were conducted, and then the data combined into one complete data set, using each of the two approaches in turn. The “Cyborg” technique for doing this uses data imputation or hot decking, i.e. scientifically estimating how a respondent might have replied by using data in the data set for people who did answer the question. With this technique we estimate and predict what the person is going to say even though we do not have all of their data. The “Monster” technique uses respondent matching or data stitching, i.e. instead of imputing data, tying actual data from different but similar respondents in the three data sets together to produce one single answer record.

Neither technique is simple to execute, but with adequate planning, preparation and knowledge either technique can produce results which are similar to those which would have been achieved by requiring all the respondents in the study to complete all three modules.
GAMIFICATION AS TOOL TO REDUCE COGNITIVE QUESTIONNAIRE LENGTH

Gamification is another technique that can be used to reduce cognitive burden on respondents. Gamification’s role in questionnaires has sometimes been misunderstood so it is useful to define what we mean by gamification at the outset. Researchers are not in the business of making games; they are in the business of asking people questions to predict things about their behavior. Gamification can play a role in making the question-answering experience engaging and compelling. Game designer Danny Day says

“I don’t necessarily believe games have to be fun to be games. I think a game is chiefly the decision to keep playing and the context of that play.”

“Games don’t necessarily have to be fun to be games” is an interesting statement. As far as Day is concerned, a game is about the decision to keep playing in the context of the game. The decision to “keep playing” or rather the decision to keep answering questions is something that researchers are particularly interested in.

One example of gamification is using words within the questionnaire to appeal to the motivations of autonomy, value relatedness and competence. This type of wording can motivate people to continue with the survey:

- **Autonomy:** Thanks for choosing to complete the survey so far. Or: Please press the > button when you are ready to continue with the survey.
- **Relatedness:** The answers you’ve given are going to be fascinating to compare to all the other people who have taken the survey.
- **Value:** Our client is going to be able to get so much out of it.
- **Competence:** You really are the experts here.

Another gamification technique is to transform the “framing” of questions: for example to replace “What did you do?” with “Can you estimate what you did?” or “Can you guess what you did?” The same data will result, but the framing makes the experience more fun.

Placing questionnaires into scenarios is a similar gamification technique. For example, an open end question which asks “Please describe yourself” is less likely to deliver high quality data. Contrast this with “You may use exactly seven words to describe yourself.” This scenario makes the respondent stop and think, which is exactly our goal for the research participant.

Of course, whenever gamification techniques are used the questions must be correctly written and must measure something that we can use before they are “gamified.”

A final aspect of gamification is adding some “bells and whistles” to the questionnaire: giving people a mental break, or varying the pace and type of question, or the judicious use of images or video.
SUMMARY AND CONCLUSIONS

- Three studies across multiple countries and cultures and across more than a decade have demonstrated that there are fatigue effects when taking surveys and that data quality suffers as a result.
- People do not drop out of long surveys; they keep going but tire and begin to satisfice
- Fatigue effects are caused not just by physical questionnaire length but also by cognitive length
- Today’s average survey length of 23 minutes is too long
- There are a variety of strategies and techniques available to reduce physical and cognitive questionnaire length

Making progress requires a joint effort by researchers, their clients and sample, panel, questionnaire design and fielding providers. Panel companies can help. For example, SSI has introduced the SSI QuestTest tool which scores questionnaire quality and gives a length estimate and mobile-friendliness score\(^7\). Researchers and their clients in turn have a variety of opportunities to reduce the amount of data collected wherever possible, while questionnaire designers can use gamification and other design techniques to reduce cognitive burden.

As we enter the mobile era, now is the time for all research stakeholders to take a hard look at surveys and see what can be done to make them shorter. The best way to do this is by taking the survey as a respondent would, rather than estimating the length and quality of experience from the questionnaire designer’s point of view. It is in the interest of all of us in the long term that respondents are happy to take surveys and that they give us high quality information.
REFERENCES


3 http://www.theinvisiblegorilla.com/videos.html

4 Rathod’s research is described in more detail in this white paper: https://www.surveysampling.com/ssi-media/Corporate/White%20Paper%202012/SSI-Questionair-Length.image

5 The Changing Landscape of Technology White Paper, Nicole Mitchell, SSI


7 Advice on how to make questionnaires mobile friendly can be found in this article for Quirks Magazine: https://www.surveysampling.com/learning-center/mobile-research/how-to-make-your-surveys-mobile-friendly

ABOUT SSI

SSI is the premier global provider of data solutions and technology for consumer and business-to-business survey research, reaching respondents in 100+ countries via Internet, telephone, mobile/wireless and mixed-access offerings. SSI staff operates from 30 offices in 21 countries, offering sample, data collection, CATI, questionnaire design consultation, programming and hosting, online custom reporting and data processing. SSI’s 3,600 employees serve more than 2,500 clients worldwide. Visit SSI at www.surveysampling.com.