

January 31, 2020

NOTE TO THE REVIEWER OF: OMB CLEARANCE 1220-0141  
“Cognitive and Psychological Research”

FROM: Jean Fox and Robin Kaplan  
Office of Survey Methods Research

SUBJECT: Submission of Materials for Evaluating Qualifiers  
in Rating Scales

Please accept the enclosed materials for approval under the OMB clearance package 1220-0141 “Cognitive and Psychological Research.” In accordance with our agreement with OMB, we are submitting a brief description of the study.

The total estimated respondent burden for this study is 238 hours.

If there are any questions regarding this project, please contact Robin Kaplan at 202-691-7383.

## Background

The Bureau of Labor Statistics (BLS) Behavioral Science Research Center (BSRC) conducted [previous research](#) to help understand the use of survey scales that are used to study a variety of constructs of importance to BLS including, satisfaction and feedback surveys (e.g., data user surveys, stakeholder surveys), respondent burden, and usability. These scales typically have five or seven response options, and are generated by pairing the construct being measured with a series of qualifiers. The scales can be **unipolar**, which prompts respondents to think about the presence or absence of a quality or attribute, and where the response options use qualifiers ranging from something like *Not at all* to *Extremely*. Or scales can be **bipolar**, where a respondent is asked to balance two opposite attributes, and where the response options range from something like *Very* of one construct to *Very* of the opposite construct (e.g., *Very easy* to *Very difficult*).

Ideally, the scales would provide equal intervals across response options that provide distinctions which most respondents would find meaningful, to obtain the most valid and reliable data possible. Researchers attempt to do this by selecting qualifiers that appear to create equi-distant response options, and then assume they are interval scale for analysis.

Our previous research explored the impact of using different qualifiers in 5-point unipolar response scales (as they were the most common in our data) in two ways: (1) by assessing the values that people assign to qualifiers commonly used in social science research in an online study where participants assigned values (0 to 100) to each qualifier (e.g., the qualifier *Very* might present on average 80 on 0-100 point scale), and via paired comparisons, asking participants to rate which of two closely related qualifiers “represents a greater quantity”; and (2) by using Item Response Theory (IRT) to evaluate how scales using the qualifiers perform in real datasets.

IRT can be used to compare the distribution of scale responses over the latent variable continuum. The latent variable continuum is a standardized scale, where extremely low levels of the trait fall below negative two and extremely high levels fall above positive two (like a z score). An ideal scale would have response options that were equally utilized and evenly distributed across the entire continuum, with very little overlap. A response option is considered to be under-utilized or suppressed, when it is overshadowed by surrounding response options. In an ideal scale, each response option should peak as the most likely response option at some point along the latent continuum, however, when a response option’s peak is suppressed by the peak of a neighboring response option, the response options provide less distinguishable estimates of the latent trait. A presentation detailing the findings of this research can be found in the link on Attachment 1. As seen in those slides, a visual analysis of IRT item characteristic curves during our exploratory research phase suggested the following about response scales:

- *Very* as an endpoint may not capture the full range of responses, but
- Adding *Extremely* may suppress the use of *Very* (that is, including the qualifier *Extremely* in the same scale with the qualifier *Very* may cause suppression of *Very*, providing less information).
- The value assigned to a qualifier by a respondent for the qualifiers *A little* vs. *Somewhat* may depend on the other responses in the scale

Findings suggested that the following factors may have impacted the interpretation of individual scale items as patterns in the data differed, including:

- The construct being measured
- The other response items used in the scale
- The context of the survey item
- The respondent population (e.g., stakeholders or data users of different BLS programs or products)

We also found that the average values assigned to the qualifiers collected in the online study were similar to the distribution of responses we'd expect to see based on the IRT findings. However, the IRT analysis revealed that some scales that should have been well-distributed were not (e.g., the qualifier *Moderately* was rated as close to the mid-point (50 out of 100) in the online study, but the qualifier *Somewhat* was closer to the mid-point in scales measuring burden and concern in the IRT analysis). The previous study was also restricted in that we had a very limited set of datasets available for analysis, we were unable to assess a wide range of response options, and we had a very small sample size. To build on the previous study, we propose a follow-up study to improve on these limitations using more robust datasets and data collection with a larger sample size.

## Study Goals

The goal of the current study is to build on and expand our previous work described above. In the proposed study, we will continue both the IRT and the online research, using the patterns observed in the previous data and the results of the latest IRT analysis so far to inform a new online study. The online study will evaluate additional qualifiers, scales, constructs, and surveys. After collecting the online data, we will again use IRT methods to compare the findings with the prior analyses.

In the first phase of the proposed study, we are analyzing new existing external datasets, including the American Association for Public Opinion Research (AAPOR) membership surveys from 2017 to 2019 ( $N = 2686$ ) and the 2019 Census Barriers, Attitudes, and Motivators Survey (CBAMS;  $N = 50,000$  households). Both surveys use the qualifiers of interest (5-point unipolar scales) across a variety of constructs of relevance (i.e., satisfaction, concern, burden, and importance, which were included on both surveys and are very common across BLS and other surveys). As such, we focus on these commonly used constructs. The surveys ask multiple

questions for each construct, which should allow us to cross validate some of our findings, especially when scales appear to collapse or fail to measure the full spectrum of the latent trait. Additionally, these datasets have larger sample sizes from more diverse populations than the previous datasets used in the first iteration of this research.

The IRT analysis of these datasets has revealed several response patterns so far:

- With the AAPOR datasets (which used the response scale *Not at all, A little, Somewhat, Very, Extremely*), our initial IRT analysis showed that respondents did not select the *A little* option very often, they more often used *Somewhat*. This suggests that inclusion of the qualifier *A little* did not add much information, implying that the qualifier *Somewhat* may better capture respondents' ratings, suppressing use of *A little*.
- We found a similar pattern in the CBAMS data (which used a similar scale, but used the qualifier *Not too* instead of *A little*), where respondents are selecting *Somewhat* more frequently than *Not too*, again showing suppression of the second qualifier in the response scale.
- Similar to our previous study's findings, the qualifier *Moderately* appeared more in the middle of the 5-point scale than *Somewhat*.
- Also, in scales measuring burden, the qualifier *Somewhat* performed closer to the mid-point of *Moderately*, whereas *A little* performed closer to *Not at all*
- With the burden questions, *Somewhat* spread the responses out more than expected, given that the values assigned in the previous online study were closer to *Moderately*.

Given these results, we identified several outstanding research questions for additional IRT analyses, including:

- Are *Not too* and *A little* suppressed when the qualifier *Somewhat* is in the same scale?
- Would inclusion of the qualifier *Moderately* reduce suppression of the qualifiers *Not too* and *A little*, or would including *Moderately* lead to suppression of the next qualifier (*Very*)?
- How do the above qualifiers operate when measuring other constructs, for example:
  - Is *Not too* suppressed in positive constructs like Satisfaction?
  - Is *Very* suppressed in negative constructs like Concern?

Because the AAPOR and CBAMS datasets did not include response scales that used each of the qualifiers of interest in a single scale, we are unable to explore conditions in which particular qualifiers are suppressed and do not perform well (i.e., do not provide information) in the existing datasets. This would require an experiment where participants are randomly assigned to answer the same survey questions using different response scale qualifiers across the constructs of interest. To answer the above remaining questions, we propose an online study using the scale qualifiers of interest to assess how they perform across different constructs and scale types in contexts that are similar to the external datasets analyzed.

In addition to the main goal of studying response scales and qualifiers, the demographic section of the proposed survey will embed questions about sexual orientation and gender identity (SOGI). In a continued effort to better understand measures of sexual orientation, BLS staff participate in the [Measuring Sexual Orientation and Gender Identity Research Group \(SOGI\)](#). In this group, we conduct research to collect feedback on terminology for sexual orientation and terms used to describe one’s orientation. Those measures will be discussed in more detail in the following sections. A similar add-on was approved in the [previous study](#).

## Study Methods

We propose an experiment that varies scale type systematically, using questions similar to those included on the AAPOR survey and CBAMS, as well as two additional scales that were not used in the AAPOR or CBAMS surveys (and thus not included in the IRT analysis), but would inform when different qualifiers perform well or get suppressed when presented together.

Participants will complete a survey where they answer questions related to the constructs of **importance, concern, burden, and satisfaction** about their experiences completing online surveys, a topic they will all be familiar with. Participants will answer the same exact set of survey questions, the only difference is that they will be randomly assigned to one of the four response scales (see Table 1); once assigned, the same response scale will carry through the entire survey. We will then analyze the data using IRT to assess how the scales compare across each of the four constructs and to one another, as well as to the external datasets (AAPOR and CBAMS), and to the findings in our previous study.

**Table 1.** Response scales that participants will be randomly assigned to; once assigned to one of the response scale conditions, that scale will be displayed through the remainder of the survey.

Condition 1	Condition 2	Condition 3	Condition 4
<b>AAPOR scale</b>	<b>CBAMS scale</b>	<b>Alternative 5-point scale</b> (performed well for burden in previous study)	<b>6-point scale</b> (Designed to explore differences in the middle qualifiers)
Not at all	Not at all	Not at all	Not at all
	Not too		
A little			A little
Somewhat	Somewhat	Somewhat	Somewhat
		Moderately	Moderately
Very	Very	Very	Very
Extremely	Extremely	Extremely	Extremely

- **Condition 1 (AAPOR scale).** The first experimental condition will include the scale used on the AAPOR survey with the qualifiers *Not at all*, *A little*, *Somewhat*, *Very*, and *Extremely*. This is a common scale used across many surveys and social science research.
- **Condition 2 (CBAMS scale).** The second is the scale used on the CBAMS, which is identical to the AAPOR survey scale, but used the less common second qualifier *Not too*.
- **Condition 3 (Alternative 5-point scale).** Following up on the results from the previous study, the third scale is an alternative that includes both *Somewhat* and *Moderately* within the same scale to assess the finding that *Somewhat* performs closer to the mid-point than *A little* in burden scales, and whether this finding holds across the other three constructs.
- **Condition 4 (Alternate 6-point scale).** A final fourth scale includes both of these qualifiers (*A little* and *Somewhat*) to determine the impact of including both of these on a singular scale on overall scale performance.

In addition to the survey questions, we will also ask participants to complete a series of paired comparisons, where they are shown two qualifiers and asked to identify the member of each pair that represents “more” or “a greater quantity.” The pairs include the qualifiers used in the survey questions (e.g., comparing *Not too* to *A little*), and the task is designed to cross-validate which qualifier, on average, represents “more” to participants. We will compare distributions across the four different types of response scales used in the experimental conditions to determine which scale performed the best, and which qualifiers performed well or were suppressed given a particular response scale. For the paired comparisons, we will determine which of two adjacent qualifiers represents “more” to participants to cross-validate the data. Afterward, participants will complete demographic questions and final questions pertaining to the burden construct.

## **SOGI Analysis**

In a continued effort to better understand measures of sexual orientation, BLS staff participate in the [Measuring Sexual Orientation and Gender Identity Research Group](#). In this group, we conduct research to collect feedback on terminology for sexual orientation and terms used to describe one’s orientation. As such, the demographic section of this protocol contains additional questions designed to test a new way to measure sexual orientation using a series of “unfolding” questions that get progressively more specific. The need for this research was identified in a recent review paper written by the FCSM SOGI Terminology subgroup as crucial to assess whether this more inclusive question can function as an alternative to the one-step sexual orientation questions that may not have all of the options available for sexual minorities to capture valid data. The data obtained from this study will be compared to prevalence rates of SOGI identity in other online studies and general collections. The exact wording of the question

can be found in the full instrument (Attachment 2). A similar analysis was also approved in our [previous study](#).

## **Survey Instrument**

Participants will complete an online survey with a series of questions about their experiences completing online surveys across the four constructs of interest (importance, concern, burden, and satisfaction). These questions mimic the wording of the questions from the AAPOR and CBAMS surveys as closely as possible, but relate to a domain that is relevant to participants (who all have experience completing online surveys), scales commonly used on BLS stakeholder surveys, and to survey methodology research more generally. Some questions were also designed to assess concerns about privacy in submitting data online, and were drawn from a previously validated questionnaire developed by Buchanan et al. (2007), as well as a new scale titled the International Survey Attitude Scale (de Leeuw et al., 2019), which was designed to assess participants' attitudes towards three components of surveys that measure respondent burden (interest, value, and general burden) and to obtain valid and reliable measures to help understand the decline in survey response rates. The questions were adapted into 5-point unipolar scales so that they are more consistent with the present research and can be included in the experiment and IRT analyses.

At the beginning of the survey, participants will first be prompted to think about the survey platforms they use to answer online surveys, as many participants in one platform also tend to complete online surveys in other platforms (Hillygus et al., 2014). This question is designed to help participants recall the breadth of online surveys that they have completed in the past and to prime them to think about their general experiences taking online surveys. This question will not be used for analysis purposes (that is, we will not make comparisons based on which survey platforms participants have used). Afterward, participants will complete a set of questions related to the four constructs (importance, concern, burden, and satisfaction). The questions will be presented in random order (organized by construct) to reduce order effects. The survey includes five to nine questions per construct to help ensure we obtain a variety of responses and to obtain scale adequate reliability statistics.

Following the rating scales, participants will answer a series of questions asking them which of two qualifiers "represents more" of something (paired comparisons) for each of the closely related qualifiers in the response scales included in this study. After answering these questions, participants will complete demographic questions and a set of questions on burden pertaining to this survey (to be included in the IRT analyses, as burden is one of the four constructs of interest). The full instrument is available in Attachment 1.

## Participants

In this research, we will use online data collection with participants recruited from Amazon.com’s Mechanical Turk (MTurk; see Berinsky et al., 2012; Paolacci & Chandler, 2014). MTurk is an online marketplace where individuals can sign up to participate in short online research tasks for nominal compensation. Although the MTurk population may not be representative of the entire U.S. population, studies using MTurk samples obtain similar results to surveys using population-based samples (e.g., Mullinix et al., 2015). Samples obtained from MTurk are more representative of the general population than other convenience samples, such as university students (e.g., Berinsky et al., 2012), or using the OSMR participant database which only contains volunteers in the DC metro area. Further, the results of this study are more concerned with internal validity than the representativeness of any one population.

We will recruit a total of 1428 participants from MTurk. This sample size was determined to sufficiently explore the range of variables of interest across the four experimental conditions (n=357 per condition), and because we expect a very small effect size as the study manipulations are subtle for online surveys of this nature (e.g., Paolacci & Chandler, 2014). This sample size also takes into account any potential break-offs, incomplete data, and participants who do not follow the task instructions, similar to other OMB-approved samples used for studies of this nature listed in the introduction.

## Burden Hours

The survey is expected take approximately 10 minutes to complete for a burden of up to 238 total burden hours.

Table 1. Estimated Burden Hours

# of Participants Screened	1428
Maximum number of Participants	1428
Minutes per participant for data collection	10
Total Collection Burden Hours	238

## Payment

Participants will receive \$1.00 for participating in the survey, which is a typical rate for similar MTurk tasks (Paolacci & Chandler, 2014). The estimated maximum total for participant fees is \$2000. This includes a commission fee that Amazon requires.



Recruiting of participants is handled by MTurk. Once participants are recruited into the study, they will receive a link to the survey, which is hosted by SurveyMonkey.com. The data collected as part of this study will be stored on SurveyMonkey servers.

Participants will be informed of the OMB number and the voluntary nature of the study with the following statement which will appear on the first page of the online instrument:

This voluntary study is being collected by the Bureau of Labor Statistics under OMB No. 1220-0141 (Expiration Date: March 31, 2021). Without this currently-approved number, we could not conduct this survey. We estimate that it will take on average 10 minutes to complete this survey. Your participation is voluntary, and you have the right to stop at any time. This survey is being administered by SurveyMonkey and resides on a server outside of the BLS Domain. The BLS cannot guarantee the protection of survey responses and advises against the inclusion of sensitive personal information in any response. By proceeding with this study, you give your consent to participate in this study.

## References

- Berinsky, A.J., Huber, G.A. and Lenz, G.S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3), 351–368. doi: 10.1093/pan/mpr057.
- Buchanan, T., Paine, C., Joinson, A. N., & Reips, U. D. (2007). Development of measures of online privacy concern and protection for use on the Internet. *Journal of the American Society for Information Science and Technology*, 58(2), 157-165.
- de Leeuw, E., Hox, J., Silber, H., Struminskaya, B., & Vis, C. (2019). Development of an international survey attitude scale: measurement equivalence, reliability, and predictive validity. *Measurement Instruments for the Social Sciences*, 1(1), 1-10.
- Hillygus, D. S., Jackson, N., & Young, M. (2014). Professional respondents in non-probability online panels. *Online panel research: A data quality perspective*, 1, 219-237.
- Mullinix, K.J., Leeper, T.J., Druckman, J.N. and Freese, J. (2015). The generalizability of survey experiments. *Journal of Experimental Political Science*, 2(2), 109–138. doi: 10.1017/XPS.2015.19.
- Paolacci, G., and J. Chandler. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23(3), 184–188.

**Attachment 1: Previous study findings**



QualifiersESRA  
presentation.pdf

## Attachment 2: Survey instrument

Welcome! Thanks for your interest in this study.

This study is part of our research into how to improve surveys. Please do your best to respond to the questions accurately. We are not collecting any personally identifiable information in the survey questions.

The study should take about 10 minutes. Please only start the study when you will be able to complete the whole study without interruption.

Please do not use your browser's back button.

This voluntary study is being collected by the Bureau of Labor Statistics under OMB No. 1220-0141 (Expiration Date: March 31, 2021). Without this currently-approved number, we could not conduct this survey. We estimate that it will take on average 10 minutes to complete this survey. Your participation is voluntary, and you have the right to stop at any time. This survey is being administered by SurveyMonkey and resides on a server outside of the BLS Domain. The BLS cannot guarantee the protection of survey responses and advises against the inclusion of sensitive personal information in any response. By proceeding with this study, you give your consent to participate in this study.

<page break>

We are interested in learning more about your experiences completing online surveys for pay or other compensation (for example, cash, gift cards, points, charitable donations, etc.)

This does **not** include classification, transcription, or other non-survey HITS on MTurk.

Aside from MTurk, please select all of the survey platforms that you have used to complete online surveys for pay or other compensation: [randomize order]

- SurveyMonkey
- Qualtrics
- Pinecone research
- Ipsos (Knowledge panel)
- YouGov
- Nielsen
- QuestionPro
- Swagbucks
- Opinion Outpost
- American Consumer Opinion
- Survey Junkie
- Toluna
- ClixSense
- GlobalTestMarket (Lightspeed)

- AmeriSpeak panel (NORC)
- Federal Government surveys or panels
- Gallup poll
- American Trends (Pew)

Please provide the names of any other survey platforms you have participated in:

\_\_\_\_\_

Approximately how many online surveys have you completed from the list above and MTurk? (Do not include classification, transcription, or other non-survey HITs from MTurk).

- 0-10
- 11-50
- 51-100
- 101-500
- 501-999
- 1000 or more

<page break>

Now we have some general questions about your experience completing online surveys for pay or other compensation.

1. How important is the compensation amount offered to you when selecting surveys to complete?
2. How important is the topic of the survey when selecting surveys to complete?
3. How important is the amount of time it will take to complete a survey when selecting surveys to complete?
4. How important is the type of compensation you will receive (for example, cash, gift card, charitable donations) when selecting which surveys to complete?
5. How important is it to receive fair pay for your completed surveys?

*\*Participants will answer these questions using the response scale they were randomly assigned to:*

Condition 1	Condition 2	Condition 3	Condition 4
Not at all important	Not at all important	Not at all important	Not at all important
	Not too important		

A little important			A little important
Somewhat important	Somewhat important	Somewhat important	Somewhat important
		Moderately important	Moderately important
Very important	Very important	Very important	Very important
Extremely important	Extremely important	Extremely important	Extremely important

<page break>

The following are some ways in which online survey platforms may use your data from the surveys you complete.

How important, if at all, is each of these uses to you personally? (EXCLUDING the survey you are currently completing).

- Public opinion on social issues
- Demographic changes
- Trends in the economy, such as employment, unemployment, and job growth
- How people spend their free time
- Health trends, such as disability, mental health issues, or physical exercise
- Consumer information, such as purchasing preferences or how people spend their money
- Market trends
- Transportation trends
- The link between education and income

*\*Participants will answer these questions using the response scale they were randomly assigned to:*

Condition 1	Condition 2	Condition 3	Condition 4
Not at all important	Not at all important	Not at all important	Not at all important
	Not too important		
A little important			A little important
Somewhat important	Somewhat important	Somewhat important	Somewhat important
		Moderately important	Moderately important
Very important	Very important	Very important	Very important
Extremely important	Extremely important	Extremely important	Extremely important

<page break>

Now we have some more general questions about your experiences completing online surveys for pay or other compensation (EXCLUDING the survey you currently completing).

1. How concerned are you, if at all, that online survey platforms will not keep your answers confidential?
2. How concerned are you, if at all, that online survey platforms will share your answers without your consent?
3. How concerned are you, if at all, that online survey platforms will use your answers against you?
4. How concerned are you, if at all, that online survey platforms will use your personal data in their research?
5. How concerned are you, if at all, with your privacy when completing online surveys?
6. How concerned are you, if at all, about being asked too much personal information when completing online surveys?
7. How concerned are you, if at all, about who might access your survey answers?

*\*Participants will answer these questions using the response scale they were randomly assigned to:*

Condition 1	Condition 2	Condition 3	Condition 4
Not at all concerned	Not at all concerned	Not at all concerned	Not at all concerned
	Not too concerned		
A little concerned			A little concerned
Somewhat concerned	Somewhat concerned	Somewhat concerned	Somewhat concerned
		Moderately concerned	Moderately concerned
Very concerned	Very concerned	Very concerned	Very concerned
Extremely concerned	Extremely concerned	Extremely concerned	Extremely concerned

<page break>

Now we have some questions about **the most recent** online survey you completed for compensation (for example, cash, gift cards, charitable donations, etc.)

1. Overall, how satisfied were you with the most recent online survey you completed?
2. How satisfied were you with the amount of compensation you received for the most recent online survey you completed?
3. How satisfied were you with the length of time it took to finish the most recent online survey you completed?
4. How satisfied were you with the topics you were asked about on the most recent online survey you completed?
5. How satisfied were you with the type of compensation you received for participating in the most recent online survey you completed? (for example, cash, gift cards, charitable donations, token of acknowledgement)?

6. How satisfied were you with the instructions provided on the most recent online survey you completed?
7. How satisfied were you with the amount of time it took to receive compensation for the most recent online survey you completed?

*\*Participants will answer these questions using the response scale they were randomly assigned to:*

Condition 1	Condition 2	Condition 3	Condition 4
Not at all satisfied	Not at all satisfied	Not at all satisfied	Not at all satisfied
	Not too satisfied		
A little satisfied			A little satisfied
Somewhat satisfied	Somewhat satisfied	Somewhat satisfied	Somewhat satisfied
		Moderately satisfied	Moderately satisfied
Very satisfied	Very satisfied	Very satisfied	Very satisfied
Extremely satisfied	Extremely satisfied	Extremely satisfied	Extremely satisfied

<page break>

Now we have some more questions about **the most recent** online survey you completed for pay or other compensation (EXCLUDING the survey you are currently completing).

1. How burdensome was the most recent online survey you completed?
2. How enjoyable was the most recent online survey you completed?
3. How effortful was the most recent online survey you completed?
4. How sensitive was the most recent online survey you completed?
5. How interesting was the most recent online survey you completed?
6. How difficult was the most recent online survey you completed?

*\*Participants will answer these questions using the response scale they were randomly assigned to across each burden item (i.e., burdensome, enjoyable, effortful, sensitive, interesting, and difficult):*

Condition 1	Condition 2	Condition 3	Condition 4
Not at all	Not at all	Not at all	Not at all
	Not too		
A little			A little
Somewhat	Somewhat	Somewhat	Somewhat
		Moderately	Moderately
Very	Very	Very	Very
Extremely	Extremely	Extremely	Extremely

<page break>

Now, we have some questions about your opinions on surveys in general. (These questions are taken from the [Survey Pulse scale](#)).

1. How enjoyable do you find responding to surveys through the mail or Internet?
2. How enjoyable do you find being interviewed for a survey?
3. How interesting are surveys in themselves?
4. How important are surveys for society?
5. How effective are surveys at collecting useful information?
6. To what extent is completing surveys a waste of time?
7. How excessive is the number of survey requests you receive?
8. How invasive do you believe opinion surveys are?
9. How exhausting is it to answer a survey with a lot of questions?

*\*Participants will answer these questions using the response scale they were randomly assigned to across each burden item (i.e., enjoyable, interesting, important, effective, wasteful, excessive, invasive, and exhausting):*

Condition 1	Condition 2	Condition 3	Condition 4
Not at all	Not at all	Not at all	Not at all
	Not too		
A little			A little
Somewhat	Somewhat	Somewhat	Somewhat
		Moderately	Moderately
Very	Very	Very	Very
Extremely	Extremely	Extremely	Extremely

<page break>

Next, you will look at pairs of words that represent different amounts:

- o Your task is to select the word that you think suggests more, or a greater quantity.
- o Some of the pairs of words may be very close in meaning.
- o Please do your best to determine which of the words suggests more, or a greater quantity.

**REMEMBER: go with your initial impression.**

Which word or phrase suggests more, or a greater quantity? [each pair will be asked individually; each item within the pair will be displayed in a random order]



Not at all A little	A little Not too	A little Somewhat	Somewhat Moderately	Moderately Extremely
Not at all Not too	A little Slightly	Not too Somewhat	Somewhat Very	Very Extremely
Not at all Slightly	Not too Slightly	Slightly Somewhat	Moderately Very	Extremely Completely

<page break>

Now, we have some questions to complete about yourself.

Do you consider yourself to be straight (that is not gay, lesbian, or bisexual?)

- Yes
- No

(If no) Which of the following best represents how you think of yourself?

- Gay
- Lesbian
- Bisexual
- Something else

(If something else) Which of the following best represents how you think of yourself? (mark all that apply)

- Queer
- Pansexual
- Asexual
- Two-Spirit
- Demisexual
- Same-gender loving
- Fluid
- No label
- Questioning/uncertain
- None of these

(if none of these) How would you describe yourself? \_\_\_\_\_

<page break>

Was your sex recorded as male or female at birth?

- Male
- Female

Do you consider yourself as male, female, or transgender?

- Male
- Female

- o Transgender
- o None of these

(If None of these) Which of the following best represents how you think of yourself? (mark all that apply)

- Male
- Female
- Transgender
- Trans male
- Trans female
- Nonbinary
- Gender fluid
- Genderqueer
- Gender nonconforming
- None of these

(if none of these) How would you describe yourself? \_\_\_\_\_

<page break>

What is your age? \_\_\_\_

What is the highest level of education you've completed?

- o Less than high school
- o High school diploma or GED
- o Some college
- o Associate degree
- o Bachelor's degree
- o Graduate school degree

Are you of Hispanic, Latino, or Spanish origin?

- o Yes
- o No

Below is a list of five race categories. Please choose one or more races that you consider yourself to be.

- White
- Black or African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Other Pacific Islander

<page break>

1. How burdensome was this survey to complete?
2. How enjoyable was this survey to complete?
3. How effortful was this survey to complete?
4. How sensitive was this survey to complete?
5. How interesting was this survey to complete?
6. How difficult was this survey to complete?

*\*Participants will answer these questions according to which of the four response scale conditions they were randomly assigned to, across the different burden items (i.e., burdensome, enjoyable, effortful, sensitive, interesting, and difficult):*

Condition 1	Condition 2	Condition 3	Condition 4
Not at all	Not at all	Not at all	Not at all
	Not too		
A little			A little
Somewhat	Somewhat	Somewhat	Somewhat
		Moderately	Moderately
Very	Very	Very	Very
Extremely	Extremely	Extremely	Extremely

<page break>

Thank you for completing this study!

To receive payment for participating in this study, please copy the validation code shown below and paste it in the form on this page for this HIT and submit the task.

We will check that code you paste in matches our own database and then approve your payment.

Your code is: \_\_\_\_\_.