




Formatting and Presentation as Confounds in Online Studies

Jessica L. Engelbrecht^a  and John E. Edlund^a

^aRochester Institute of Technology

Abstract ■ Participant biases are a well-documented part of psychological research, and accounting for common biases (e.g., participant acquiescence) is now commonplace. However, the role that online survey construction plays in handling or aggravating these biases has not been as deeply investigated. The current research asked participants to complete a simple trait attribution task in the form of an online survey in which the scale type, order of stimuli presentation, and presence of page breaks varied. As hypothesized, these varying presentation features impacted participant responses. Precautions for avoiding these errors are discussed, as well as the possible implications for replicability in social psychology.

Keywords ■ participant biases, online studies, survey construction, order effects.

 engej084@newschool.edu

 *JLE*: 0000-0003-2397-2892; *JEE*: 0000-0003-3868-1844

 [10.20982/tqmp.12.3.p243](https://doi.org/10.20982/tqmp.12.3.p243)

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

Reviewers

■ One anonymous reviewer

Introduction

Researchers have identified a number of potential confounds and concerns to consider when designing studies for psychological research (e.g., McDermott & Miller, 2007). Ensuring the reliability and validity of measures and stimuli are the foremost goals of researchers, but these factors can be adversely affected by unforeseen biases during the execution of the study (Nichols & Edlund, 2015). Biases that stem from the expectations of the researchers (Strickland & Suben, 2012) or the design of the materials can be anticipated and controlled in many cases, but often participants bring with them their own thinking biases that can impact results on a large scale (e.g., Twedt, Crawford, & Proffitt, 2015). Eliminating these biases in order to assure the validity of the data collected can be challenging, as many heuristics, cognitive biases, and logical fallacies are nearly ubiquitous, independent of intelligence (Stanovich & West, 2008).

Surveys are particularly easy to execute and yet difficult to design in a fashion that controls all possible confounds. Surveys are a highly attractive option for psychological research, as they require minimal interaction between researchers and participants and provide data in formats convenient for analysis. Several guidelines for

their use are typically followed; wording should be understandable to participants, questions should not bias participants towards a specific answer with their wording, and response items should be standardized and presented in a single index when possible (Krosnick, 1999). Although these basic guidelines have been recognized as effective for some time, they cannot prevent all possible flaws that a particular survey design may have. Next, we briefly review some of common biases found in survey research.

Acquiescence

One core flaw inherent in survey design is found in the manner of data collection. As researchers cannot ask follow-up questions to participants when responses are unclear, ensuring that specific answers are received is imperative. Keeping survey questions as simple as possible is desirable; however, this doesn't always provide the results researchers hope for. Binary questions (that seek answers such as yes/no, agree/disagree, true/false) lead to a phenomenon known as acquiescence, wherein participants are ultimately more likely to answer in the positive than the negative (Krosnick, 1999). Thus, researchers will end up collecting more "yes", "true", or "agree" answers in response to survey items, and the results are not necessarily informative. This confirmation bias is endemic and well



documented; a variety of measures have led to the conclusion that acquiescence generally occurs nearly 10% of the time, but can be reduced across trials with the same participants (e.g., Hoffman, 1960).

A number of possible reasons for acquiescence have been suggested, from a desire to be socially agreeable to a mere preference for saying confirmatory words (Krosnick, 1999). Interestingly, questions phrased such that participants should agree or disagree as to whether something should be allowed or forbidden produce an opposite effect, with participants far more likely to answer “no” regardless of whether the issue was to “allow” or “forbid” something (Hippler & Schwartz, 1986). However, in this case, the behavior seems to be motivated by a desire to remain neutral on serious issues, as the questions were often political in nature and “no” seemed to be the least assertive response option. While inversions of the acquiescence effect are less common, they appear to be moderated by similar principles. Research supports the notion that acquiescence stems from a desire to satisfy even in situations where participants do not have an opinion on a topic or don’t have enough cognitive resources available to devote deep thought to a topic, a phenomenon dubbed “satisficing” (Krosnick, Narayan, & Smith, 1996).

Order Effects

When binary answer formats will not elicit detailed enough responses, or there is a risk of the acquiescence effect hampering results, researchers may choose to use a more complex form of response measurement. Rating scales are a popular alternative for those who wish for more subtle gradation of responses, especially Likert-type scales. These scales also provide data in a convenient format for analysis, but they come with their own problems as well. Scales which do not provide a label for each selectable point along the continuum can often confuse participants, who end up guessing as to where their answer belongs on the array (e.g., Krosnick & Berent, 1993; Lodge, 1981). In addition, scales which use numeric labels in between semantically labeled endpoints lead participants to disambiguate the mid points of the scale with more nuance (Schwarz, Knauper, Hippler, Noelle-Neumann, & Clark, 1991). For example, a participant offered a scale from 1 (never) to 5 (always) may be uncertain what frequency of behavior is represented by a 3 on the scale if such point is not labeled as well. An optimized ratings scale of 5 (for unipolar) or 7 points (for bipolar) with complete labels is suggested as the best approach (Krosnick, 1999). The order in which the scale endpoints are oriented can lead to another participant bias. When reading a survey, participants are likely to select an option close to the beginning of the scale (a primacy effect), while those who hear op-

tions read aloud are more likely to select an option close to the end of the scale (a recency effect) (e.g., Krosnick & Alwin, 1987; McClendon, 1991). Having participants receive scales in alternating orders between subjects can account for this, but is not commonly practiced. Memory bias towards the easiest remembered item, social bias towards confirmation, and a lack of cognitive resources to pursue deeper thought on the issue all play a part (Krosnick, 1999).

Anchoring

Anchoring can be loosely defined as the tendency to come to a decision that is unduly influenced by a previously considered piece of information known as an anchor. This can occur regardless of whether the anchor value is germane to the question at hand. For example, participants who write down their own ID number for an experiment several times begin to give numerical estimates similar to their ID number in response to general knowledge questions (Wilson, Houston, Etling, & Brekke, 1996). Although anchoring is one of the more robust cognitive biases (e.g., Mussweiler, 2002), it is not typically discussed in the context of biases which can disrupt data collection. Importantly, anchoring stands out in a variety of mental tasks, from basic retrieval of general knowledge to negotiations and legal judgments (Mussweiler & Strack, 1999).

Anchoring is presumed to be a side-effect of the common use of comparison-adjustment in cognition; people commonly determine values by comparing them against known quantities (e.g., Simmons, LeBoeuf, & Nelson, 2010). Indeed, people are generally biased towards comparison as opposed to contrast (Gentner & Markman, 1997). Asking participants to use an anchor as a contrast point, or priming them to do so using a contrast task beforehand, is the most expedient way to defeat the anchoring effect (e.g., Mussweiler, 2002; Wilson et al., 1996). Attempts to motivate participants to overcome the effect consciously by informing them of it and offering cash rewards for accuracy in an experiment does not achieve results, as participants cannot accurately gauge the impact of the bias or how to correct for it (e.g., Joslyn, Savelli, & Nadav-Greenberg, 2011).

The Current Study

The current study aimed to examine how these biases interact in online survey research. The aforementioned biases have long been studied and are known obstacles to survey construction; however, when scales that have been validated on pen and paper are put online for convenience of use, new confounding variables can emerge. We predicted that sub-optimal use of survey response scales would influence the responses given by participants. Specifically, we hypothesized that page breaks, order of



presentation, and scale type would radically affect the validity of results by changing how participants judged the stimuli.

Methods

Participants

A total of 257 participants (129 female, 127 male) were recruited to take part in an online study. Of these, 200 participants were recruited through psychology classes and compensated with class credit, while the remaining 57 participants (29 men, 27 women, and one respondent who declined to share their gender) were recruited through Amazon's mTurk (a crowdsourcing research marketplace: mturk.com) and given monetary compensation for participation.

Procedure

Participants were provided a link to one of eight different surveys, randomized by a PHP script. The content of the surveys was identical in all versions; only the presentation order of the stimuli and question response scales differed. All surveys asked participants to watch two video clips of individuals ostensibly engaging in a conversation with an off-screen partner and answer questions about what they saw. Each video featured the target facing the camera and displaying a clear facial emotion. One video featured a crying target, the other a laughing target. Both video clips were ten seconds in length, and featured no sound. Participants answered three questions after viewing each of these clips, on 7-point Likert-type scales. The questions asked participants to diagnose the emotion they had seen the individual in the video display, and make inferences about the personality of that individual and the conversation they were having. After answering these questions, participants filled out the Need for Cognition scale (Cacioppo & Petty, 1982).

The eight versions of the survey differed in presentation in three different ways, corresponding to the independent variables. Participants may have encountered the two stimuli videos in either possible order; some viewed the happy video first, others the sad video. Page breaks were also present in half of the surveys administered, dividing the first video stimulus and its accompanying questions from the second; when a page break was present, participants had no way of knowing a second video would follow the first, and they were unable to return to the first stimulus to compare the two. Finally, the Likert scales used to answer the three questions about each video stimulus took two different forms. Some participants used a 7-point bipolar Likert scale ranging from “very sad” to “very happy” to answer the questions, each of which was presented only

once. The other half of participants would instead rate the video they saw along two sets of 7-point unipolar Likert scales, one ranging from “not at all sad” to “very sad”, the other from “not at all happy” to “very happy”.

Regardless of which of these versions participants encountered, the survey always concluded with demographic questions and the Need for Cognition Scale questions. In total, participants spent between 10-15 minutes completing the study.

Results

Prior to analysis, all participant responses were transformed to allow for a combined analysis (although means are presented in an untransformed manner in Table 1 and Figure 1). Unipolar responses were combined such that responses to the sad scale were subtracted from the happy scale. This score was then z-transformed to create a single standard z-score for the unipolar scales. Independently, responses to the bipolar scales were z-transformed. These independent transformations resulted in a single z-score for each question for each participant where higher scores were indicative of a participant seeing more happiness in the target video.

The primary analysis was an evaluation of the three question types (emotion evaluation, personality evaluation, and topic evaluation), the type of response scale (uni- or bi-polar), the presence of a page break (present, absent), and the ordering of the videos (happy or sad first). This analysis was significant, $F(2, 238) = 7.37, p < .001, \eta_{\text{partial}} = .058$. Given the nature of the response scales, we decided to decompose this interaction based on the individual questions.

Looking at the emotion evaluation item (specifically asking how happy the individual was behaving), we found a significant interaction between the type of response scale, the presence of a page break, and the ordering of the videos across the two videos, $F(1, 245) = 4.23, p = .043, \eta_{\text{partial}} = 0.017$. Specifically, we found that when the happy video was presented first, participants showed a significant bias towards rating both of the videos as being happy, except when unipolar scales were used with a page break. Please see table 2 for this interaction's z-scores

Looking at the topic evaluation item (specifically asking the type of topic being discussed), we found a significant interaction between the type of response scale, the presence of a page break, and the ordering of the videos across the two videos, $F(1, 242) = 5.58, p = .019, \eta_{\text{partial}} = 0.023$. This interaction took the form that when the happy video was presented first, both the bipolar scale (with no page break) and the unipolar scale (with a page break) resulted in a negativity bias; whereas the bipolar scale (with a page break) and the unipolar scale (with no page break) resulted



Table 1 ■ Pre-transformation Descriptive Statistics: Participant Judgments

Scale	Happy First		Sad First	
	Break	No Break	Break	No Break
<i>Happy: Emotion Evaluation Question</i>				
Bipolar	6.15	5.97	6.04	5.97
Unipolar Happy	3.85	4.20	4.06	4.07
Unipolar Sad	1.09	1.17	1.00	1.07
<i>Happy: Personality Evaluation Question</i>				
Bipolar	5.32	5.36	5.81	5.35
Unipolar Happy	3.59	3.94	3.72	3.97
Unipolar Sad	1.31	1.54	1.22	1.10
<i>Happy: Topic Evaluation Question</i>				
Bipolar	6.29	5.72	6.04	5.91
Unipolar Happy	3.88	4.06	4.14	4.17
Unipolar Sad	1.12	1.09	1.00	1.00
<i>Sad: Emotion Evaluation Question</i>				
Bipolar	1.86	1.79	1.46	1.73
Unipolar Happy	1.16	1.11	1.06	1.24
Unipolar Sad	4.18	4.14	4.28	4.03
<i>Sad: Personality Evaluation Question</i>				
Bipolar	2.61	3.06	3.21	3.15
Unipolar Happy	1.61	1.60	1.61	1.24
Unipolar Sad	3.24	3.47	3.46	3.83
<i>Sad: Topic Evaluation Question</i>				
Bipolar	1.60	1.70	1.57	1.91
Unipolar Happy	1.15	1.12	1.06	1.38
Unipolar Sad	4.06	4.03	4.11	3.93

in a positivity bias. The exact opposite pattern of results emerged when the negative video was presented first. See table 3 for this interaction’s z-scores.

Looking at the personality item, we found a significant order effect, where participants showed a positivity bias when the sad video was presented first ($z_{mean} = .10, z_{sd} = .05$) whereas a negativity bias emerged when the happy video was presented first ($z_{mean} = -.08, z_{sd} = .05$), $F(1, 246) = 4.73, p = .03, \eta_{partial} = 0.019$. All other non-reported effects did not reach statistical significance (p 's > .05).

Discussion

As expected, suboptimal construction of the online surveys significantly impacted participant responses. Response scale type, video presentation order, and the presence of a page break interacted across conditions to lead participants to succumb to different biases. Typically, the presentation order of the videos led participants to anchor their perception of the displayed emotion in the subsequent video on the first clip viewed, unintentionally using it as a baseline for comparison as in earlier research (Simmons et al., 2010). When used together, unipolar scales and

page breaks mitigated the effect of this anchoring; however, this also led to a new bias in their perception of the topic of subsequent videos that opposed the expected anchoring bias. While it is clear that formatting changes can be used to correct for confounding biases like anchoring, it is also possible to overcorrect and cause new unforeseen problems, just as participants themselves do when instructed to correct for common heuristics and errors.

Ultimately, the former results address the conditions under which participants were influenced in the way they perceived the target emotions; separate effects emerged in the results addressing the personality and topic evaluation questions. These attributions involve more complex social cognitive processes in which some measure of conscious thought goes towards making a decision about the person in the video. The setup in our experiment was direct enough that participants were likely aware that making strong dispositional judgments from a highly emotional behavior sample was foolish. When a happy video was shown first, participants demonstrated a negativity bias in rating the personality of the target, and a positivity bias when a sad video was shown first. This suggests a conscious effort to consider the video as being not representa-

**Table 2** ■ Z-scores for Participant Judgments: Emotion Evaluation Question

Scale	Happy First		Sad First	
	Break	No Break	Break	No Break
Bipolar	.18	.04	-.14	.01
Unipolar	-.09	.13	.05	-.09

Table 3 ■ Z-scores for Participant Judgments: Topic Evaluation Question

Scale	Happy First		Sad First	
	Break	No Break	Break	No Break
Bipolar	.13	-.10	-.03	.12
Unipolar	-.08	.05	.12	-.13

tive of the target's ordinary mannerisms. Suboptimal survey construction further complicated these effects on the questions addressing the topic of the conversation. Bipolar scales or unipolar scales with a page break replicated the results found on the dispositional question, suggesting that the opposite arrangements had some impact on the participants' efforts to control for the probable abnormality of the target's behavior in the video.

Participants in this experiment performed a simple task brimming with opportunities for biases to slip in. Some participants may have put forth extra effort to ensure that they treated each stimulus the same way, but even then the presentation of the survey may have hindered their efforts or caused them to overcorrect. Future research should investigate how participants' own independent attempts to correct for bias are impacted by suboptimal survey construction; up to this point the focus has been mostly on investigating what biases exist and participant's abilities to overcome them when explicitly instructed to do so. Other known social cognition biases not investigated by the current study may also have had an impact on results. For example, tasks assessing social cognition in regards to dispositional and situational attributions are impacted by the race and gender of the target, with women viewed as being more motivated by emotional reactivity and the negative behavior of racial out-group members being attributed more to disposition (e.g., Barrett & Bliss-Moreau, 2009; Pettigrew, 1979). Our study showed no clear gender or race effects, but this cannot be taken as an indication that none exist that would interact with our findings. In particular, anchoring effects may interact with attribution biases about race and gender, leading a whole series of answers to be anchored to an example which elicited particularly extreme attributions errors from participants.

The specific manipulation of biases in these results demonstrates that these common features of online surveys can have a strong impact on the results. Even when

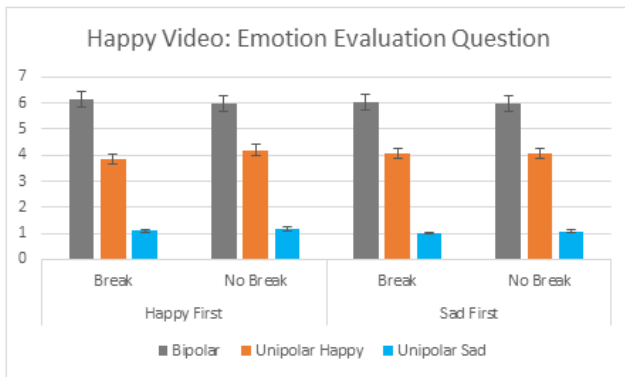
researchers have designed high-quality stimuli, questions, and measures, the format of their presentation online can adversely affect the validity of the findings, making them difficult to replicate or apply practically; consideration of these issues is especially critical in light of recent concern about how psychology as a whole deals with failure to replicate results (Klein et al., 2015; Collaboration, 2015). Replication can potentially become very tricky when two researchers think they're doing everything the same way, but have actually made critical distinctions in the way they formatted their surveys for online use that lead to different results. As new technology continues to provide more expedient methods to connect with participants over the internet, researchers would do well to use caution in converting tried and true surveys to an online format, as it's entirely possible to lose things in translation or introduce new bugs into the system.

In light of the results of this study, a number of recommendations are worthy of mention. First and foremost, we believe that researchers should pay attention to the full design of their study rather than the "key" operational decisions (e.g., Nichols & Edlund, 2015). Second, we believe that researchers should be aware that the first stimuli participants see serve as the point of comparison or contrast for all subsequent stimuli, so extra effort should be undertaken to ensure participants do not all undertake the stimuli and questions in the same order. If counterbalancing is feasible, a test of order effects should be performed first. If counterbalancing is not feasible, the impact of such biases should be considered and accounted for carefully (ideally using multiple studies that feature internal replications). Finally, we believe that the evidence suggests that unipolar scales and page breaks are good ways to correct for such biases, although our results suggest there is a possibility of overcorrection when using both. Researchers who wish to take advantage of the convenience of online surveys would do well to consider using at least one of these methods in accordance with the content of their survey.

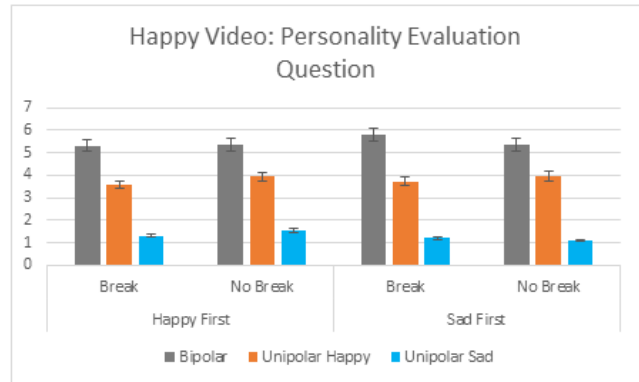


Figure 1 ■ Mean Plots of Raw Participant Ratings in Table 1

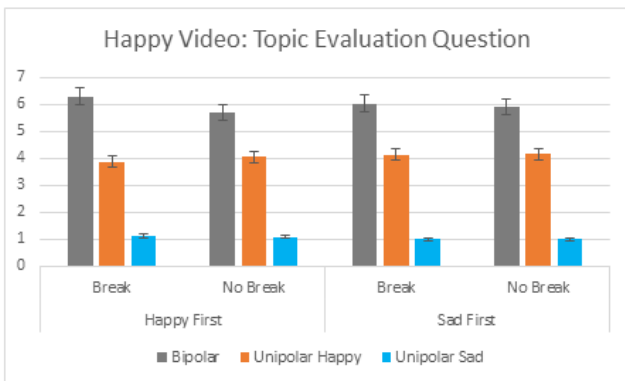
(a)



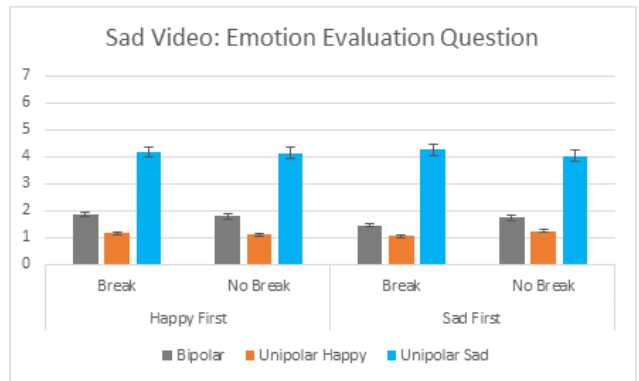
(b)



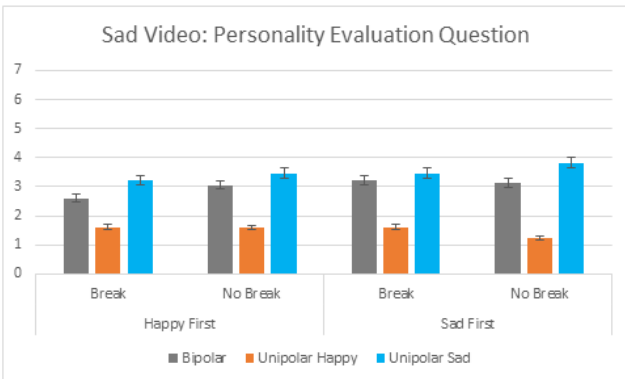
(c)



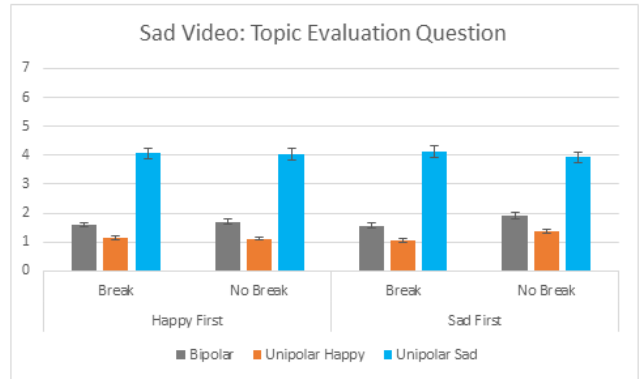
(d)



(e)



(f)





References

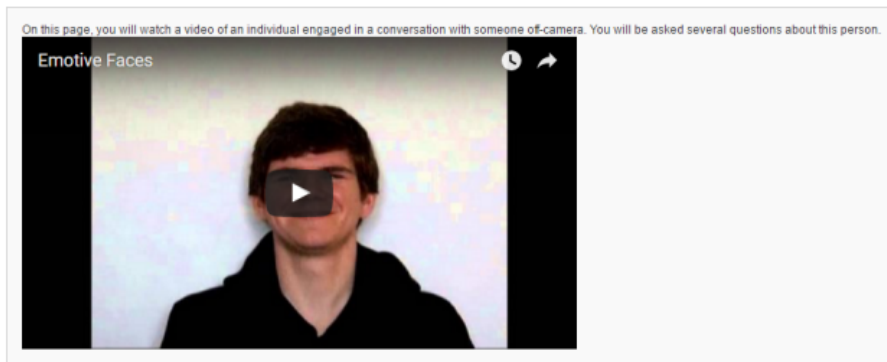
- Barrett, L. F. & Bliss-Moreau, E. (2009). She's emotional. he's having a bad day: attributional explanations for emotion stereotypes. *Emotion*, 9, 649–649. doi:[10.1037/a0016821](https://doi.org/10.1037/a0016821)
- Cacioppo, J. T. & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–116. doi:[10.1037/0022-3514.42.1.116](https://doi.org/10.1037/0022-3514.42.1.116)
- Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.
- Gentner, D. & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1), 45–56. doi:[10.1037/0003-066X.52.1.45](https://doi.org/10.1037/0003-066X.52.1.45)
- Hippler, H. J. & Schwartz, N. (1986). Not forbidding isn't allowing: the cognitive biases of the forbid-allow asymmetry. *Public Opinion Quarterly*, 50, 87–96. doi:[10.1086/268961](https://doi.org/10.1086/268961)
- Hoffman, P. J. (1960). Social acquiescence and “education”. *Educational and Psychological Measurement*, 20, 769–776. doi:[10.1177/001316446002000413](https://doi.org/10.1177/001316446002000413)
- Joslyn, S., Savelli, S., & Nadav-Greenberg, L. (2011). Reducing probabilistic weather forecasts to the worst-case scenario: anchoring effects. *Journal of Experimental Psychology: Applied*, 17(4), 342–353. doi:[10.1037/a0025901](https://doi.org/10.1037/a0025901)
- Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B. J., Bahnik, S., Bernstein, M. J., & Woodzicka, J. A. (2015). Investigating variation in replicability. *Social Psychology*, 9. doi:[10.1027/1864-9335/a000178](https://doi.org/10.1027/1864-9335/a000178)
- Krosnick, J. A. (1999). Maximizing questionnaire quality. In Robinson, J. P., P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of political attitudes* (pp. 37–57). New York, NY, US: Academic Press, 1999.
- Krosnick, J. A. & Alwin, D. F. (1987). An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51, 201–219. doi:[10.1086/269029](https://doi.org/10.1086/269029)
- Krosnick, J. A. & Berent, M. K. (1993). Comparisons of party identification and policy preferences: the impact of survey question format. *American Journal of Political Science*, 37, 941–964. doi:[10.2307/2111580](https://doi.org/10.2307/2111580)
- Krosnick, J. A., Narayan, S., & Smith, W. R. (1996). Satisficing in surveys: initial evidence. *New Directions for Evaluation*, 1996(70), 29–44. doi:[10.1002/ev.1033](https://doi.org/10.1002/ev.1033)
- Lodge, M. (1981). *Magnitude scaling: quantitative measurement of opinions*. Beverly Hills, CA: Sage.
- McClendon, M. J. (1991). Acquiescence and recency response-order effects in interview surveys. *Sociological Methods and Research*, 20, 60–103. doi:[10.1177/0049124191020001003](https://doi.org/10.1177/0049124191020001003)
- McDermott, K. B. & Miller, G. E. (2007). Designing studies to avoid confounds. In R. J. Sternberg, H. L. I. Roediger, & D. F. Halpern (Eds.), *Critical thinking in psychology* (pp. 131–142). New York, NY, US: Cambridge University Press, 2007. doi:[10.1017/cbo9780511804632.009](https://doi.org/10.1017/cbo9780511804632.009)
- Mussweiler, T. (2002). The malleability of anchoring effects. *Experimental Psychology*, 49(1), 67–72. doi:[10.1017/cbo9780511804632.009](https://doi.org/10.1017/cbo9780511804632.009)
- Mussweiler, T. & Strack, F. (1999). Comparing is believing: a selective accessibility model of judgmental anchoring. In Stroebe, W., & M. Hewstone (Eds.), *European review of social psychology* (vol. 10) (pp. 1–44). Chichester, England: Wiley.
- Nichols, A. L. & Edlund, J. E. [J. E.]. (2015). Practicing what we preach (and sometimes study): methodological issues in experimental laboratory research. *Review of General Psychology*, 19(2), 191–202. doi:[10.1037/gpr0000027](https://doi.org/10.1037/gpr0000027)
- Pettigrew, T. F. (1979). The ultimate attribution error: extending allport's cognitive analysis of prejudice. *Personality and social psychology bulletin*, 5(4), 461–476. doi:[10.1177/014616727900500407](https://doi.org/10.1177/014616727900500407)
- Schwarz, N., Knauper, B., Hippler, H. J., Noelle-Neumann, E., & Clark, L. (1991). Rating scales: numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55(4), 570–582. doi:[10.1086/269282](https://doi.org/10.1086/269282)
- Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: do people adjust from provided anchors? *Journal of Personality and Social Psychology*, 99(6), 917–932. doi:[10.1037/a0021540](https://doi.org/10.1037/a0021540)
- Stanovich, K. E. & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Personality Processes and Individual Differences*, 94(4), 672–695. doi:[10.1037/0022-3514.94.4.672](https://doi.org/10.1037/0022-3514.94.4.672)
- Strickland, B. & Suben, A. (2012). Experimenter philosophy: the problem of experimenter bias in experimental philosophy. *Review of Philosophy and Psychology*, 3(3), 457–467. doi:[10.1007/s13164-012-0100-9](https://doi.org/10.1007/s13164-012-0100-9)
- Twedt, E., Crawford, E., & Proffitt, D. R. (2015). Judgments of others' heights are biased toward the height of the perceiver. *Psychonomic Society Bulletin Review*, 22, 566–571. doi:[10.3758/s13423-014-0689-z](https://doi.org/10.3758/s13423-014-0689-z)
- Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 125(4), 387–402. doi:[10.1037/0096-3445.125.4.387](https://doi.org/10.1037/0096-3445.125.4.387)



Appendix: Screenshot of Survey in no-break bipolar scale condition

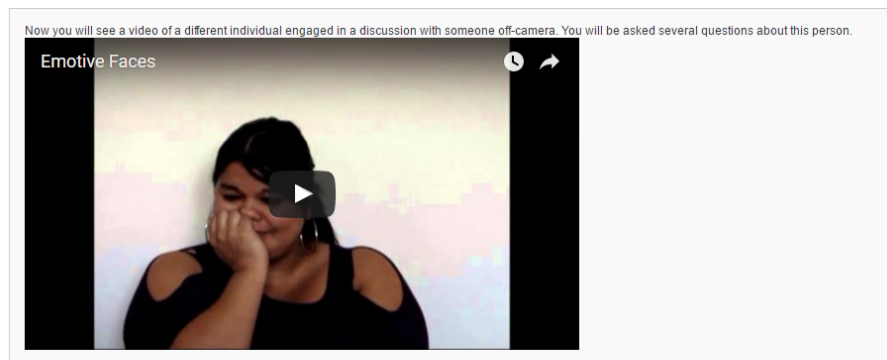
Figures 2 and 3 shows the top and bottom part of the display presented to the participants.

Figure 2 ■ Screenshot of Survey in no-break bipolar scale condition (top part)



2. Please answer the following questions as quickly and accurately as possible.

	Very sad						Very happy
On a scale from one (very sad) to seven (very happy), how was the individual behaving?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
On a scale from one (very sad personality) to seven (very happy personality) what kind of personality does this person have?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
On a scale from one (a very sad topic) to seven (a very happy topic) what kind of topic is this individual discussing?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



3. Please answer the following questions as quickly and accurately as possible.

	Very sad						Very happy
On a scale from one (very sad) to seven (very happy), how was the individual behaving?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
On a scale from one (very sad personality) to seven (very happy personality) what kind of personality does this person have?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
On a scale from one (a very sad topic) to seven (a very happy topic) what kind of topic is this individual discussing?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Figure 3 ■ Screenshot of Survey in no-break bipolar scale condition (bottom part)

4.

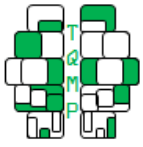
	Strongly Disagree				Neither agree nor disagree				
1. I would prefer complex to simple problems.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. I like to have the responsibility of handling a situation that requires a lot of thinking.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Thinking is not my idea of fun.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. I try to anticipate and avoid situations where there is likely chance I will have to think in depth about something	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. I find satisfaction in deliberating hard and for long hours.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. I think only as hard as I have to	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. I prefer to think about small, daily projects to long-term ones.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. I like tasks that require little thought once I've learned them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. The idea of relying on thought to make my way to the top appeals to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. I really enjoy a task that involves coming up with new solutions to problems.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. Learning new ways to think doesn't excite me very much.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. I prefer my life to be filled with puzzles that I must solve.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. The notion of thinking abstractly is appealing to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. I would prefer a task that is intellectual, difficult and important to one that is somewhat important but does not require much thought.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. I feel relief rather than satisfaction after completing a task that required a lot of mental effort.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. It's enough for me that something gets the job done; I don't care how or why it works.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. I usually end up deliberating about issues even when they do not affect me personally.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. What is your gender?

- Female
- Male
- Prefer not to answer

7. Which of the following ethnic backgrounds best describes you?

- White (including Middle Eastern)
- Black or African American (including Africa and the Caribbean)
- Asian (including Indian subcontinent and the Philippines)
- Native American or Native Alaskan
- Hawaiian or Pacific Islander
- Hispanic or Latino



Citation

Engelbrecht, J. L. & Edlund, J. E. [John E.]. (2016). Formatting and presentation as confounds in online studies. *The Quantitative Methods for Psychology*, 12(3), 243–252. doi:[10.20982/tqmp.12.3.p243](https://doi.org/10.20982/tqmp.12.3.p243)

Copyright © 2016, *Engelbrecht and Edlund*. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 31/03/2016 ~ Accepted: 10/06/2016