

# Continuous and Binary Sets of Responses are Not the Same: Evidence from the Field

**Noelia Rivera-Garrido**

Universidad Loyola Andalucía

**M. P. Ramos-Sosa** (✉ [mpramos@uloyola.es](mailto:mpramos@uloyola.es))

Universidad Loyola Andalucía

**Michela Accerenzi**

Fundación ETEA, Universidad Loyola Andalucía

**Pablo Brañas-Garza**

Universidad Loyola Andalucía

---

## Research Article

**Keywords:** binary, questionnaire, feasible, negative wording

**Posted Date:** June 29th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-646078/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Continuous and binary sets of responses are not the same: Evidence from the field

Noelia Rivera-Garrido<sup>1</sup>, M.P. Ramos-Sosa<sup>1,\*</sup>, Michela Accerenzi<sup>2</sup>, Pablo Brañas-Garza<sup>1</sup>

**1** Economics, *LoyolaBehLab*, Universidad Loyola Andalucía, Spain

**2** Fundación ETEA, *LoyolaBehLab*, Universidad Loyola Andalucía, Spain

\*Corresponding author: mpramos@uloyola.es

## Abstract

This paper conducts a pre-registered study aimed to compare binary and continuous set of responses in questionnaires. Binary responses consist of two possible opposed responses (Yes/No). Continuous are numerical, where respondents can indicate an answer in a 0 to 10 horizontal blind line. We study whether binary and continuous feasible sets of responses yield to the same outcome (distribution) and the same cost (duration in minutes). We collect data from 360 households in Honduras who were randomly assigned to Yes/No questions or given a slider (visual scaling 0-10) to mark their responses, therefore, we provide *causal evidence*. We find that respondents are 13% more likely to say “Yes” and spend 2.1 minutes less in the binary setting. Besides, we find that most of the differences between binary and continuous settings arise from questions which include negative wording.

June 26, 2021

## Introduction

When dealing with opinion surveys, question design is a key aspect to obtain the desired level of detailed information and to avoid difficulties in expressing opinions. Both, capture the true response of the subject and avoid any bias from the design are of utmost importance. Therefore, the type of question used in a survey is crucial and will depend on the aim of the researcher.

The literature on survey design, and more specifically, the choice of the response format, offers a wide range of typology of questions to evaluate the degree of agreement with an issue [1; 2; 3; 4; 5]. Agree-disagree scales are used to obtain the level of agreement with an issue [6]. If the scale options directly refer to the issue under evaluation, item-specific scales are then recommended [7; 8]. In social sciences, dichotomous scales (binary responses), in which two responses with opposed directions are offered (typically Yes or No) are widely spread when the goal is to evaluate the direction of the answer (agreeing or disagreeing with an issue). This type of questions are very clear in meaning. There is considerable consensus on the meanings of the options [9] and they are easy for the respondents to understand. They also imply low costs and low time for both the researcher and the respondent [10]. In addition, in case of factual questions it has been showed that people correctly answer binary questions more often when the correct answer is "Yes" [3]. This format also requires less interpretative efforts compared to longer ones [9]. However, the lack of neutral responses may alter the results. It may lead to acquiescence bias, which is the tendency for a respondent to agree with a statement without considering the content of the item or

to please the researcher [11; 12; 13; 3]. Moreover, it could be problematic for respondents with neutral attitudes because of the lack of accurate mapping between moderate responses and the dichotomous options offered. As a result, it would require a major cognitive effort [14; 9; 15].

Continuous rating scales allow the researcher to obtain the exact level of agreement with certain issue. When continuous rating scales are numerical, respondents can indicate an option in a horizontal line, marking precise numbers which express the exact level of agreement for an issue with respect to a maximum number, usually 10 or 100. The use of this metric scale allows subjects to easily transform their opinions, whether extreme or mild, into replies, reducing distortion [16]. Nevertheless, point meanings may become less clear as the respondent cannot differentiate between one number and the other [9].

When choosing among binary and continuous rating scales, it is important to evaluate the advantages and disadvantages of each type of question. In terms of statistic evaluation, continuous scales allow a wider range of procedures [17]. Regarding survey length and complexity, continuous scales require more time to explain the set of possible answers and to think about its conversion into numbers, implying higher reasoning costs [16; 10]. However, compared to binary responses, continuous set of feasible answers provide better quality of information [9]. In binary scales, where respondents have to convert mild responses into Yes-No answers, the respondent is forced to make computations to transform their opinion in the middle to an extreme answer, resulting in an inaccurate choice, which consequently may manifest a higher level of agreement that would not be reached if middle responses were available (i.e., acquiescence or satisficing [13; 3]). Moreover, binary scales do not allow for nuance in respondents' answers or to study respondents' expectations at an individual level. Therefore, if respondents have problems in computing their mild thoughts into extreme responses, biases arise.

This study investigates the extent to which the use of binary and continuous scales yields different results. Specifically, we study if the probability of agreement with a statement is the same when using binary and continuous

set of responses. To do so, we conduct a field experiment in Honduras with 360 randomly selected participants. Following a random assignment, half of the participants were assigned to the binary treatment, in which subjects has to respond Yes or No. The other half were assigned to the continuous treatment, in which they had to answer the same questions but with a different set of responses. Specifically, subjects in the continuous treatment indicate the level of agreement with an statement using a slider with a 0 to 10 scale. To compare replies among treatments, we discretize continuous answers. We assign the binary option "No" if the option indicated is below 5. Those who claim to agree by more than 5 are assigned to the binary option "Yes". In case the option indicated is 5, the answer is randomly assign to "Yes" and "No". Additionally, we study whether question wording such as the use of negative words (e.g., no, anything) or prescriptive expressions (e.g., act as, take care of, limit activities) has also an effect on the probability of agreement. We also analyze the differences in terms of survey length.

We find that respondents are more likely to say "Yes" in the binary setting compared to the continuous setting. In particular, the probability of answering "Yes" is 2.3 times higher when asking in binary instead of continuous. Regarding the survey length, binary questionnaires took 2.1 minutes less, which implies on average a reduction of 41% with respect to continuous questionnaires. Our results hold when excluding males from the sample and those subjects with lower cognitive abilities.

We extend the analysis to the wording of the questions. The literature has shown that question wording also plays an important role in acquiring accurate responses. Negative wording has been defined as a question in which disagreement would be a positive answer [12]. Negative wording is usually introduced in surveys to avoid fatigue throughout the questionnaire, acquiescence [18; 19; 12] and to avoid respondents choosing the same answer [20]. Nevertheless, [12; 21; 22], among others, raise the attention to the additional processing effort that negative questions imply, overall in binary sets, where the respondent have to decide to agree or disagree with a question that includes negation, what may generate confusion (see [23]). In the same line, our data show that when using negative wording in ques-

tions, the probability of agreement jumps to 41% in binary sets compared to continuous.

## Results I: Aggregate data

We start showing results of the behavior of participants along the entire experiment. Recall that every participant replied to 30 questions in a random order (52% starting from block *A*, and 48% starting from block *E*). Figure 1 shows the average number of answers “Yes”, along with the average time (in minutes) respondents spent on the entire questionnaire in both, continuous and binary answers.

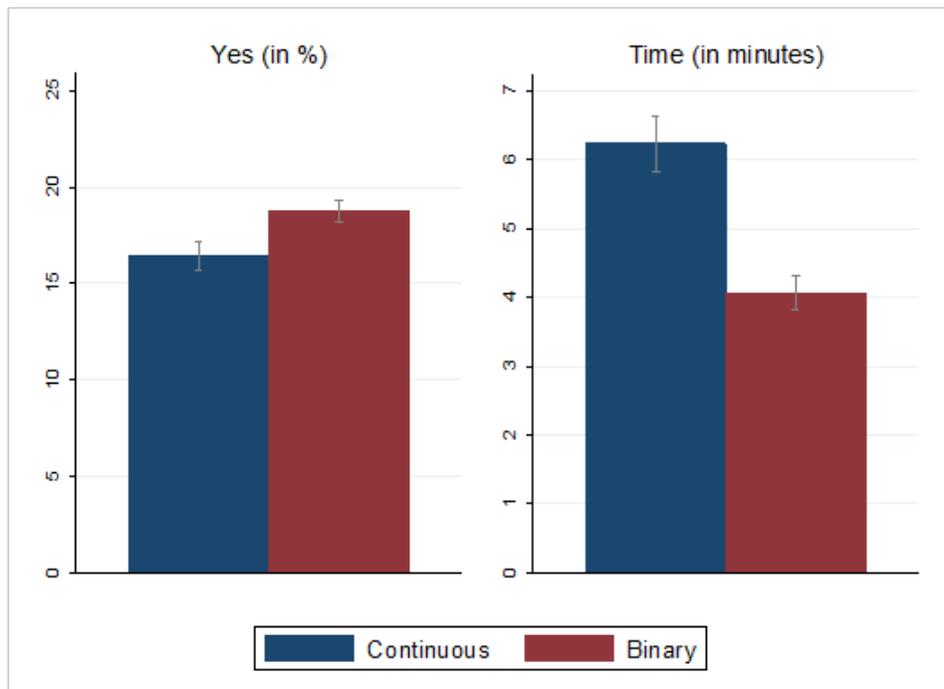


Figure 1: **Binary vs. Continuous Settings.** On the left: Average number of times the subjects agree throughout the questionnaire. On the right: Response time in minutes. Sample: All.

We observe that the probability of agreement with the survey question is higher when asking in binary, and that the average time in the binary setting is lower. Table 1 shows the regression analysis. Model 1 does not include controls. Model 2 includes controls for age, gender, and ethnic. Model 3 includes controls for age, gender, ethnic, school, sufficient income, having a daughter, education, and the order of the questions in the survey. In addition, after controlling for these characteristics, the difference between binary and continuous is statistically significant at the 1 % level (Table 1).

It is important to recall that subjects were assigned to treatments randomly. Hence our results imply *causality*. Therefore, we conclude that:

**Result 1:** Binary settings produce higher shares of Yes.

**Result 2:** Continuous settings are more time consuming.

Before moving to question by question data we want to explore whether Result 1 and Result 2 are robust to certain variations. First, we remove the few males in the sample. Second, we remove those subjects facing problems with probabilities.

First, we consider only women in the sample (87%) and repeat the same exercise in order to check whether these results are driven by the very few men in the sample (13%).

Figure 2a repeats previous Figure 1 using the restricted sample of women only. Table 1 (middle) repeats the regression analysis for the sub-sample of women. Results are similar than when considering the whole sample. Then, our results are not driven by men. Therefore, *Results 1 and 2 are confirmed for women*.

Now, we focus on respondents' cognitive abilities. We want to see whether Result 1 and Result 2 hold once we consider that subjects may face problems when are asked about probabilities and expectations

Figure 2b reflects the same than Figure 1 for those individuals who show medium or high ability according to the extension of [24] developed by [25]. Then, we conclude that *Results 1 and 2 are not driven by the respondents' cognitive abilities*.

In conclusion, the use of binary (instead of continuous) set of responses make subjects respond more “*Yes*” (or less “*No*”) and complete the survey much faster. These results are robust to excluding men and low ability respondents.

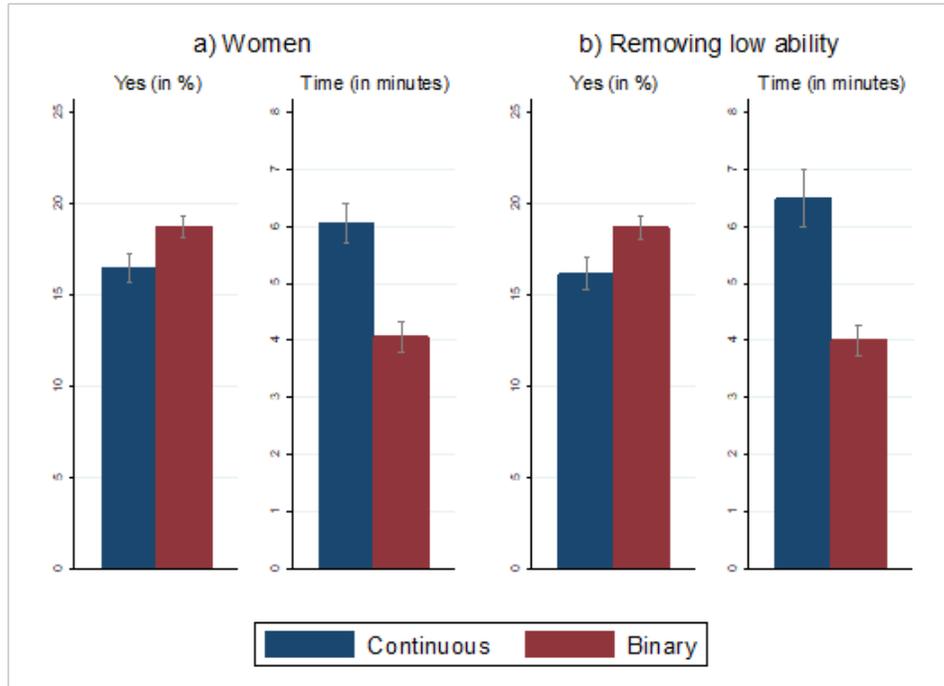


Figure 2: **Binary vs. Continuous Settings.** On the left: Average number of times the subjects answer agree throughout the questionnaire. On the right: Response time in minutes. a) women, b) low ability.

Table 1: Binary vs. Continuous Set of Responses. The effect of the set of responses (aggregate level).

<b>Outcome:</b>	<b>Model (1)</b>	<b>Model (2)</b>	<b>Model (3)</b>	<b>Model (1')</b>	<b>Model (2')</b>	<b>Model (3')</b>
<b>Sample: All</b>	<b>Aggregate Number of "Yes"</b>			<b>Time (min)</b>		
Binary	2.317*** (0.459)	2.333*** (0.458)	2.297*** (0.284)	-2.159*** (0.238)	-2.127*** (0.230)	-2.120*** (0.169)
Dep. Var (Mean)	17.66	17.66	17.66	5.11	5.11	5.11
Observations	353	353	353	338	338	338
R-squared	0.069	0.076	0.131	0.201	0.217	0.291
<b>Sample: Women</b>						
Binary	2.253*** (0.486)	2.255*** (0.488)	2.166*** (0.383)	-1.991*** (0.225)	-2.007*** (0.226)	-2.030*** (0.276)
Dep. Var (Mean)	17.72	17.72	17.72	4.97	4.97	4.97
Observations	306	306	306	295	295	295
R-squared	0.068	0.075	0.150	0.218	0.224	0.319
<b>Sample: Medium/ High Ability</b>						
Binary	2.534*** (0.544)	2.569*** (0.551)	2.488*** (0.327)	-2.486*** (0.291)	-2.407*** (0.278)	-2.352*** (0.180)
Dep. Var (Mean)	17.56	17.56	17.56	5.18	5.18	5.18
Observations	249	249	249	242	242	242
R-squared	0.082	0.088	0.169	0.243	0.267	0.354

Note: The dependent variable is a dummy equal one if agree with the survey question. The table reports the estimated coefficients for a dummy variable equal one if binary set of responses. Model (1) without controls. Model (2) includes controls for age, gender, and ethnic. Model (3) includes controls for age, gender, ethnic, school, sufficient income, have a daughter, education, and the order of the questions in the survey. Clustered standard error by school in Model 3.

Robust standard error in parenthesis.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

## Results II: Question by question

Figure 3 shows the average number of “*Yes*” (and “*No*”) question by question. It shows that Result 1 (higher number of “*Yes*” when asking in binary) does not hold for every single question. Indeed we see that many items are almost identical in binary or continuous.

In particular, we find that 10 out of 30 questions are different and most importantly the direction of the difference is always the same (see Model 3 in Table S7,  $p < 0.05$ ). That is, from the ten questions that report different numbers of “*Yes*”, all of them report higher probability of agreeing with the survey question when asking in binary instead of in continuous. This difference is statistically significant at a 0.05 level.

**Result 3:** In 63% of the cases there are not differences between binary and continuous settings. In the other cases, the differences are always in the same direction (positive) as in Result 1.

For sake of completeness we repeat the analysis for women only and for the sub-sample of subjects who passed the cognitive test. Results remain in the same direction (Table S8).

Our results also hold when we remove those individuals responding the neutral option 5 (see Table S9) or when we consider as a cut-off the value of 7 (see Table S10). We selected this value since it is the minimum qualification to pass exams at school in Honduras (over a total of 10) [26] and it could be considered as a turning point between agreement and disagreement.

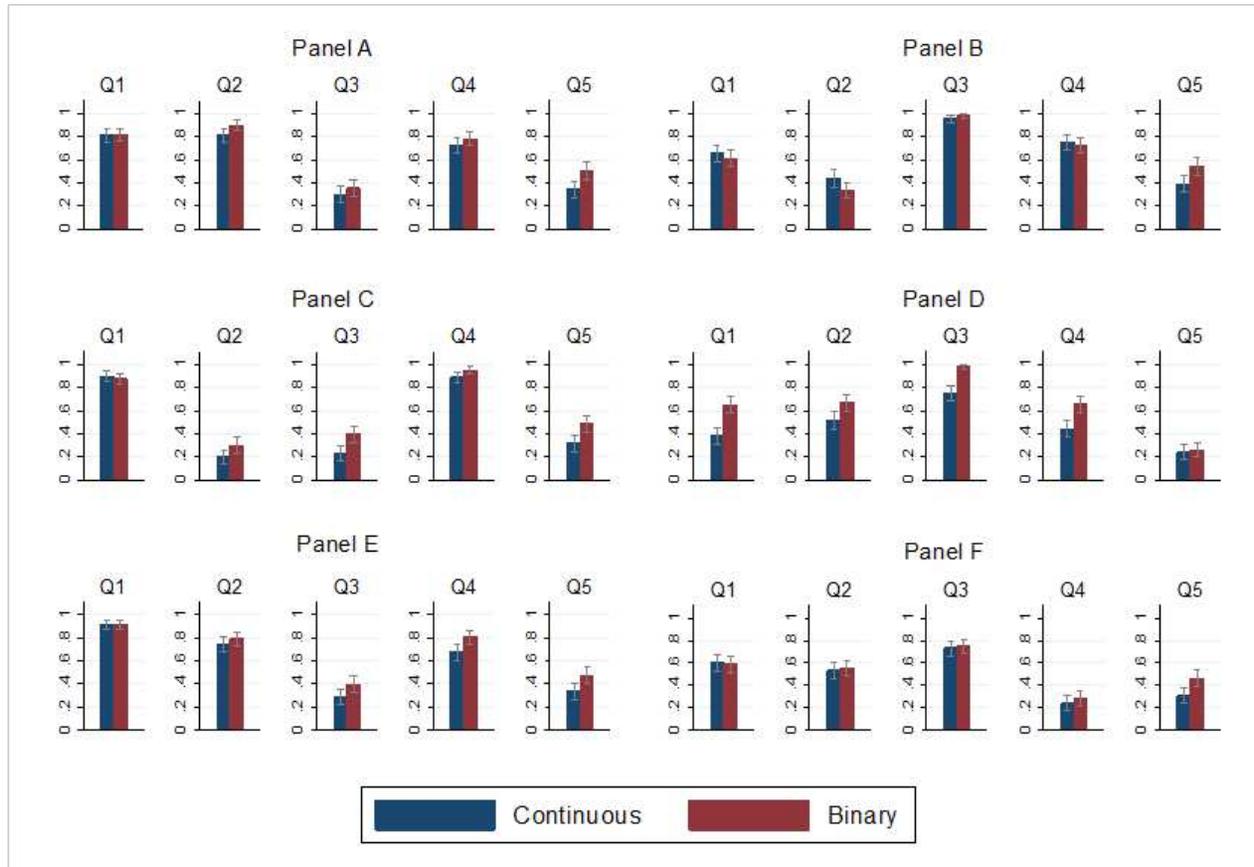


Figure 3: Binary vs. Continuous Settings. Item by item.

## The impact of wording

We next explore the role of question wording in explaining Results 3. To do this, we distinguish three types of question: negative, prescriptive, and other. *Negative* refers to questions which include words like “no” or “not”. *Prescriptive* refers to those questions that include words like “have to” or “must be”. *Other* questions refer to those questions that cannot be classified as negative or prescriptive. Negative questions are the questions SQ5 in Panel A, SQ5 in Panel B, SQ5 in Panel C, SQ1, SQ4, SQ5 in Panel D, SQ5 in Panel E, and SQ5 in Panel F. Prescriptive questions are the questions SQ2-SQ4 in Panel B.

First, we analyze the differences at the aggregate level when asking in binary and continuous for each group of questions (negative, prescriptive, and other). That is, we compare the total number of “Yes” that the subjects respond in binary and continuous. We start with the whole sample. Table S4 reports the average difference in the number of “Yes” when asking in binary instead of continuous for each group of questions. Regarding the group of negative questions, those who answered binary questionnaires said “Yes” 1.4 times (or 41%) more than those who answered the continuous questionnaires. The average number of “Yes” is also higher in binary when considering the group of other questions (1.02 times or 8.43%). The difference is not statistically significant when considering the group of prescriptive questions.

**Result 4:** The probability of agreeing with the survey question is higher when using binary set of responses and negative wording (41%).

We repeat our analysis for women and removing those subjects with low ability according to the results obtained from [24] (see Table S4 middle and bottom) and Result 4 is confirmed.

## Discussion

We investigate how answers change when questions are asked using binary Yes-No answers or by offering a continuous rating scale.

Using data from 360 households in Honduras where subjects were randomly assigned to treatments this paper provides *causal evidence* regarding differences in the probability of agreement and the time length of the questionnaire. We also provide non-causal evidence of the impact of negative wording.

We compare binary and continuous responses by discretizing the continuous variable, collapsing values into 0 (“No”) and 1 (“Yes”) answers. That is, those which indicates an option below 5 are assigned to 0, while those indicating more than 5 were assigned to 1. Thus, those answering the neutral option 5, are randomly assigned either to 0 or 1.

Our first result shows that, on average, the probability of agreement is 13% higher in binary sets compared to continuous sets. The fact that the binary setting yields to higher agreement rates is not new since this typology of questions is affected by acquiescence, the tendency to agree with the question without considering the content of the item (see [13; 3]).

Second, we control for the length of the interviews and, as expected, Result 2 shows that continuous settings require more time than binary settings. Specifically, opting for Yes-No responses reduced the average length of the survey in 2.1 minutes, which implies a reduction of 41.5%. This result is also expected given the wider range of options offered to the respondents, which may need more time to accurately map their opinion in a 0 to 10 interval. Furthermore, the use of the slider to collect answers is also time consuming. This is in line with [16; 10], that declared longer completion times by using sliders. Sliders are more engaging than other visual scales [27; 28], and increase the time and effort needed to provide an answer [15]. Sliders also allow respondents to communicate exact values and its use simultaneously convey the respondent how precise the expected answer should be [29]. In addition to the longer completion times, there are other challenges to tackle, such as how to handle the starting position of the marker [30; 17], or the inclusion or not of anchors and labels [31; 17]. The manipulation of sliders by enumerators remove all these difficulties, explaining respondents how to use it properly, and avoiding issues with incomplete data or with the slider bar movement and interpretation [30].

To check whether Results 1 and 2 are robust, we analyze the probability of agreement focusing on two particular sub-samples: women and considering respondents' ability. When we consider only women in the sample, which accounts for 87% of the sample, we find similar results, being the probability of agreement 2.1 times higher in binary compared to continuous sets. The length of the questionnaire is also 2 minutes shorter using binary questions for this sub-sample. Hence, the probability of agreement and the time length differences between binary and continuous set of responses are not driven by men.

Additionally, we consider respondents' ability using an extension of [24] (see also [25]). We find that 71% of the sample responded correctly to these questions. Using this sub-sample, we test our data and find similar results: high ability respondents require on average 2.3 minutes less when using a binary rather than a continuous questionnaire, which implies a reduction of 45%, 2pp less compared to the whole sample. We also find that the probability of agreement is higher in binary sets compared to continuous sets. Specifically, respondents answer "Yes" 2.4 times more (or 14.1%) in binary sets. [13; 3].

Third, we also estimate the differences of responses for every single question. Result 3 shows that 37% of total questions present significant differences between binary and continuous sets, showing in all cases higher agreement in the binary setting. These results hold when we only consider women and when we take the high-ability respondents' sub-sample. To go a step further, we analyze whether the role of question wording has an effect in our results. In particular, we distinguish *negative* wording questions, which include negative adverbs in the sentence. Specifically, 8 out of 30 questions include negative words, from which 7 of them show significant differences when we compare the binary and the continuous treatment. This specific group of questions report 41% more agreement in binary than in continuous sets. The group of questions classified as "*other*" also report 8.43% more agreement in binary than in continuous sets. However, the group of questions with *prescriptive* words show no differences.

## Concluding remarks

This paper tests whether asking questions using binary or continuous sets provide different responses. To do so, we design a quasi-experimental survey where subjects were randomly assigned to the binary or continuous treatment, hence we provide here causal evidence. The experiment was conducted in Honduras where 360 participants were randomly selected. At the aggregate level, we find that subjects are more likely to say “Yes” in the binary setting (Result 1). We also find that subjects in the binary treatment spend 41.5% less time to respond the whole questionnaire (Result 2). Thus, continuous set of responses comes at a higher cost. We test whether both results still hold when the sample is only composed by women (87% of the sample) and results are similar. We also test whether results remain constant when we control for cognitive abilities. The proportion of subjects classified as low ability is 29%. Results removing this subgroup of participants are similar to the ones obtained using the whole sample. Then we estimate the difference between binary and continuous treatment question by question. We find two relevant results. First, the differences between binary and continuous sets do not occur in every single question but in 37% of them. However, the difference between binary and continuous always goes in the same direction. That is, the probability of agreeing with the survey question is higher when using binary set of responses instead of continuous. Finally, we analyze the impact of question wording. To that end, we differentiate questions with negative, prescriptive and other wording. We find that the negative wording makes subjects increase the probability of answer “Yes” in a 41%. We find no differences between binary and continuous when using prescriptive wording; in sum, our paper provides causal evidence of the difference between binary and continuous set of responses. We find that the type of answer is important and could yield different results. In addition, question wording could play an important role in explaining its results.

## **Methods**

### **Protocol**

The experiment took place in Santa Rosa de Copán (Honduras) between May 1 and 14, 2019. It was conducted by a consultancy firm (PILARH) in four districts (Osorio, El Carmen, Prado Alto and Santa Teresa) in which 360 households were randomly selected based on the eligibility criterion of having a child between 6 and 9 years old in one of the 12 selected schools, being their relatives over 18 years old, the ones interviewed. The selected schools are: Divina Providencia, Jerónimo J. Reina, Jesús Banegas Membreño, Jorge Portillo, José Cecilio del Valle, José María Medina, Los Ángeles, Manuel Bonilla, Membreño, San Antonio, San José and Santo Domingo Savio. The entire experiment consisted of four experimental tasks and a questionnaire, which was always placed at the end. This study is entirely based on the questionnaire.

The experiment was conducted by enumerators who were specifically trained for the field work. Enumerators were in charge of reading and explaining the instructions, thus avoiding self-management. They received a list of households they had to visit. The random allocation of households to treatments was done prior to the visit of the communities and the enumerators did not have any influence on such selection. Enumerators conducted all face-to-face interviews in the households of the participants. Questionnaire was paper-based. It consisted on 30 questions about menstruation and sexual and reproductive rights. It was divided into 6 blocks (A-F) of 5 items each. The original questionnaire in Spanish and the translated questionnaire in English are in the Supporting Information sections S1 and S2, respectively. Only one experimental subject was interviewed per household.

### **Ethics and pre-registration**

This study was approved by the Ethics Committee of Universidad Loyola Andalucía and was conducted according to the principles expressed in the Declaration of Helsinki. All participants signed an informed consent prior to

participate. The field study was pre-registered in April 15, 2019 and made public in April 25, 2019 in: <https://aspredicted.org/ps766.pdf>.

## Design

To assess the effect of binary versus continuous set of responses in a questionnaire, subjects were randomly allocated to  $2 \times 2$  treatments (between-subjects):  $\{B, C\} \times \{\text{Order1}, \text{Order2}\}$ .

- $\{B, C\}$  refers to binary and continuous set of responses, respectively. The binary set of responses corresponds to the case in which the set of possible answers are limited to “Yes” or “No” (i.e., dichotomous question). The continuous set of responses allow subjects to indicate an option within a given interval. Specifically, enumerators capture respondents’ answers using sliders, which include a 0-to-10 point range and are partially labeled, only indicating the extreme points (0 and 10). Respondents can drag the marker and drop it to a desired location in the horizontal line. Figure 4 shows the slider. Respondents see the front side of the slider (panel A). At the same time, the enumerator observes the back side (panel B).
- $\{\text{Order1}, \text{Order2}\}$  refers to the order in which the 30 questions (organized in 6 blocks of 5 items each) were presented. 52% of the participants started by blocks A, B, C (and then D, E, F) while the remaining 48% started by blocks D, E, F and then A, B, C. Within each block, the order of the items does not change.

We also are interested in comparing the time expended when using binary instead of continuous set of responses. To that end, we recorded the time (in minutes and seconds) devoted to complete the whole questionnaire by each participant (note that, given the paper-based nature of the questionnaire, we could only account for the time of the whole questionnaire, not question by question). Finally, to account for respondents’ ability, we make use of an extension of [24] developed by [25]. This test consist of several questions



**Panel A. Respondents view of the slider (front side).**



**Panel B. Enumerator view of the slider (back side).**

Figure 4: Slider

that allow us to test whether respondents are able to deal with probabilities (See Supporting Information S3 for further discussion on this point).

## Sample

The total sample consists of 353 subjects. However, 7 of them did not answer all questions. Table S5 presents descriptive statistics of the key variables in our analysis. All the subjects surveyed were residents in Honduras at the time of the survey and had a child between 6 and 9 years old in one of the 12 schools considered (see Table S6 for further information). The average age of the participants is 34 years. They were mostly women (87%) with secondary or less than secondary education, and with a daughter living at home (65%). Most of the individuals have not a particular ethnicity (70%), and 29% are low ability (see S3 for more details about measuring ability). In addition, 77% of the individuals in the sample indicate that they have enough money to feed their children.

All the subjects answered the same questionnaire with only one difference, the type of answer (binary vs continuous). 52% of the participants

answered binary questions (Yes/No questions), and the remaining 48% answered continuous questions.

The average number of times that the subjects answered “Yes” is higher in binary questionnaires (see standardization below). In addition, the response time in binary answers is lower (about 2 minutes per questionnaire less).

Finally, we provide the balance of the randomization across treatments in our experiment (Table S5, column 6). Our sample is balanced in all the individual characteristics (age, ethnicity, sufficient income, ability, questions order, and education) except gender ( $p = 0.008$ ) and marginal differences in secondary education.

## Standardization

In order to analyze whether individuals respond differently to the same question depending on whether it is asked in binary (*Yes/No*, where 1 means *Yes* and 0 means *No*) or continuous (scale from 0 to 10, from disagree to agree), we need the responses to the two types of questionnaires (binary and continuous) to be comparable. Following the existing literature [32; 33], we discretize the continuous responses, and then we compare them with the binary responses.

Let  $r \in [0, 10]$  be the set of feasible answers for the continuous questionnaire. We generate the following binary variable,  $d$ .

$$d = \begin{cases} 0, & \text{if } r < 5 \\ 1, & \text{if } r > 5 \\ 0, & \text{with } prob = 0.5 \text{ if } r = 5 \\ 1, & \text{with } prob = 0.5 \text{ if } r = 5 \end{cases}$$

If the individual claims to agree with a statement by indicating more than 5 ( $r > 5$ ), the binary equivalent is “Yes” and it is assigned to 1. In case of being less than 5 ( $r < 5$ ), the equivalent is “No”, and it is assigned to 0. Finally, if the individual indicates the neutral response 5 ( $r = 5$ ), we

randomly assign the answers to 0 or 1.

We opt for splitting 5 answers as it represents the mean point in the continuous set. Using the slider instead of a “Yes/No” question, we can measure the intensity of an opinion. Similarly to our case, the well-known Likert scale questions [34] also measure this intensity. The original 5-point Likert scale questions use the mean value to indicate “neither agree nor disagree”, values above the mean indicate agreement with the statement, and values below the mean indicate disagreement with the statement [35]. Following a similar reasoning and using our continuous questionnaire, we assume that: (1) when the answer is greater than 5, it means agreement with the survey question; (2) when the answer is less than 5, it means disagreement with the survey question; (3) when the answer is equal to 5, it means neither agree nor disagree. Then, as a result, when the individual answers 5, we randomly assign him/her to “Yes” or “No”.

## **Acknowledgments**

We would like to thank the participants at the 2021 IAREP SABE, the UIB webinar and the 2020 ESA Global Online Around-the-Clock Meetings for useful comments. This research was supported by the Spanish Ministry of Economy and Competitiveness (PGC2018-093506-B-I00) and the Excelencia-Junta (PY-18-FR-0007).

## **Authors contribution**

M.A. conceived, designed and performed the experiments. P.B. conceived, designed and performed the experiments and wrote the paper. N.R. and M.R. conceived and designed the experiments, analyzed the data and wrote the paper.

## **Competing interests**

The authors declare no competing interests.

## Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on request.

## References

- [1] Krosnick JA, Fabrigar LR. Designing rating scales for effective measurement in surveys. *Survey Measurement and Process Quality*. New York, John Wiley 1997: 141-164.
- [2] Converse JM, Presser S. Survey Questions: Handcrafting the Standardized Questionnaire. *Journal of Marketing Research*. 1987. doi:10.2307/3151644
- [3] Krosnick JA. Survey research. *Annu Rev Psychol*. 1999;50: 537–567. doi:10.1097/PRS.0b013e3181ea44f9
- [4] Schaeffer NC, Presser S. The Science of Asking Questions. *Annu Rev Sociol*. 2003;29: 65–88. doi:10.1146/annurev.soc.29.110702.110112
- [5] Krosnick JA, Presser S. Question and Questionnaire Design.. J D Wright P V Marsden (Eds), *Handb Surv Res (Second Ed West Yorkshire, Engl Emerald Group*. 2010.
- [6] Revilla MA, Saris WE, Krosnick JA. Choosing the Number of Categories in Agree-Disagree Scales. *Social Methods Res*. 2014;43: 73–97. doi:10.1177/0049124113509605
- [7] DeCastellarnau A. A classification of response scale characteristics that affect data quality: a literature review. *Qual Quant*. 2018;52: 1523–1559. doi:10.1007/s11135-017-0533-4
- [8] Saris WE, Revilla M, Krosnick JA, Shaeffer EM. Comparing questions with agree/disagree response options to questions with item-specific response options. *Surv Res Methods*. 2010;4: 45–59. doi:10.18148/srm/2010.v4i1.2682

- [9] Krosnick JA, Judd CM, Wittenbrink B. Measurement of Attitudes. *Handb attitudes attitude Chang.* 2005; 21–76.
- [10] Hilbert S, Küchenhoff H, Sarubin N, Nakagawa TT, Bühner M. The influence of the response format in a personality questionnaire: An analysis of a dichotomous, a likert-type, and a visual analogue scale. *TPM - Testing, Psychom Methodol Appl Psychol.* 2016;23: 3–24. doi:10.4473/TPM23.1.1.
- [11] Billiet JB, Mcclendon MJ. Modeling Acquiescence in Measurement Models for Two Balanced Sets of Items. *Structural Equation Modeling.* 2009;7:4, 608-628. doi:10.1207/S15328007SEM0704.
- [12] Colosi R. Negatively Worded Questions Cause Respondent Confusion. *Proc Surv Res Methods Sect.* 2005;20746: 2896–2903.
- [13] Krosnick JA. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Appl Cogn Psychol.* 1991;5: 213–236.
- [14] Birkett NJ. Selecting the number of response categories for a Likert-type scales. *Proc Am Stat Assoc.* 1986; 488–492.
- [15] Kunz T. Rating Scales in Web Surveys: A Test of New Drag-and-Drop Rating Procedures. mimeo 2015.
- [16] Givon MM, Shapira Z. Response to Rating Scales: A Theoretical Model and Its Application to the Number of Categories Problem. *J Mark Res.* 1984;21: 410. doi:10.2307/3151467
- [17] Chyung SYY, Swanson I, Roberts K, Hankinson A. Evidence-Based Survey Design: The Use of Continuous Rating Scales in Surveys. *Perform Improv.* 2018;57: 38–48. doi:10.1002/pfi.21763
- [18] Allen M (Ed.) *The SAGE encyclopedia of communication research methods.* Sage Publications.; 2017.
- [19] Jang Y, Kwag KH, Chiriboga DA. Not saying I am happy does not mean I am not: cultural influences on responses to positive affect items

in the CES-D. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 65(6); 2010. pp. 684-690.

- [20] Weisberg, H. F. *The total survey error approach: A guide to the new science of survey research*. University of Chicago Press.
- [21] Holleman B, Kamoen N, Krouwel A, Van De Pol J, De Vreese C. Positive vs. Negative: The impact of question polarity in voting advice applications. *PLoS One*. 2016;11: 1–17. doi:10.1371/journal.pone.0164184
- [22] Kamoen N, Holleman B, Mak P, Sanders T, Van Den Bergh H. Why Are Negative Questions Difficult to Answer? on the Processing of Linguistic Contrasts in Surveys. *Public Opin Q*. 2017;81: 613–635. doi:10.1093/poq/nfx010
- [23] Dillman, D. A., Smyth, J. D., Christian, L. M. *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. John Wiley & Sons.
- [24] Delavande, Adeline, and Hans-Peter Kohler. Subjective expectations in the context of HIV/AIDS in Malawi. *Demographic Research* 20. 2009: 817.
- [25] Estepa, L., Jorrat, D., Orozco, V., Rascón-Ramirez, E. Beyond beans: Eliciting probabilities using slider in the field. Mimeo 2020.
- [26] Secretaría de Educación. República de Honduras. Acuerdo No. 1796-SE-2017: Lineamientos sobre la Evaluación de los Aprendizajes para los Años Lectivos 2018-2019. República de Honduras - Tegucigalpa, M. D. C., 6 de enero del 2018 No. 34,535.
- [27] Funke F, Reips UD, Thomas RK. Sliders for the smart: Type of rating scale on the web interacts with Educational Level. *Soc Sci Comput Rev*. 2011;29: 221–231. doi:10.1177/0894439310376896
- [28] Roster CA, Lucianetti L, Albaum G. Exploring slider vs. Categorical response formats in web-based surveys. *J Res Pract*. 2015;11: 1–19.

- [29] Kuhlmann T, Dantlgraber M, Reips UD. Investigating measurement equivalence of visual analogue scales and Likert-type scales in Internet-based personality questionnaires. *Behav Res Methods*. 2017;49: 2173–2181. doi:10.3758/s13428-016-0850-x
- [30] Funke F. A Web Experiment Showing Negative Effects of Slider Scales Compared to Visual Analogue Scales and Radio Button Scales. *Soc Sci Comput Rev*. 2016;34: 244–254. doi:10.1177/0894439315575477
- [31] Bourdel N, Alves J, Pickering G, Ramilo I, Roman H, Canis M. Systematic review of endometriosis pain assessment: How to choose a scale? *Hum Reprod Update*. 2015;21: 136–152. doi:10.1093/humupd/dmu046
- [32] Komorita SS. Attitude Content, Intensity, and the Neutral Point on a Likert Scale. *J Soc Psychol*. 1963;61: 327–334. doi:10.1080/00224545.1963.9919489
- [33] Matell MS, Jacoby J. Is there an optimal number of alternatives for likert scale items? study 1: Reliability and validity. *Educ Psychol Meas*. 1971;31: 657–674. doi:10.1177/001316447103100307
- [34] Taherdoost H. What Is the Best Response Scale for Survey and Questionnaire Design; Review of Different Lengths of Rating Scale / Attitude Scale / Likert Scale. *Int J Acad Res Manag*. 2019;8: 2296–1747.
- [35] Likert R. A Technique for the Measurement of Attitudes. *Arch Psychol* 22(140), 1–55. 1932. doi:10.4135/9781412961288.n454

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Suplmaterial.pdf](#)