

Inattentive respondents cause attenuation bias in conjoint experiments

Agabek Kabdullin¹

July, 2024

Abstract

In forced choice conjoint experiments, in expectation, the respondents should be picking between two answer options at the same rate thanks to randomization of profile attributes. Using replication data, I demonstrate that this empirical regularity does not always hold across various political science experiments: respondents pick the first option more often. I employ data from seven original conjoint experiments to show that it is inattentive respondents who are driving this result. I derive that this primacy bias leads to attenuation of estimates. Finally, I present a research design that pins down the mechanism behind primacy bias and elucidates whether it is the product of the language of instruction or survey layout.

¹Social Science Division, New York University Abu Dhabi. Email: ak5996@nyu.edu

1 Introduction

Conjoint experiments popularized by Hainmueller et al. (2014) are entering the mainstream and are slowly becoming a staple in political science. Scholars have paid close attention to the interpretation of the main metrics of interest (Abramson et al., 2022; Ganter, 2023), potential measurement problems that could occur in conjoint experiments (Clayton et al., 2023) and the issue of survey satisficing (Bansak et al., 2018). The challenge of inattentive respondents in online surveys in general has also been acknowledged in recent work (Peyton et al., 2022). Some propose ways to track attention of survey takers (Kane et al., 2023), and others caution that ignoring respondents failing manipulation checks may bias results (Aronow et al., 2019). Ultimately, the question of what to do with inattentive respondents depends on the nature of the problem at hand (Alvarez et al., 2019).

While existing work has discussed the problem of inattentive respondents in various surveys and survey experiments (Atsusaka and Kim, 2023; Blair et al., 2019), inattention poses a specific challenge in conjoint experiments. In particular, I demonstrate that in forced-choice (and rank-based conjoint experiments), inattentive respondents produce attenuation bias because they do not pick options at random – rather, they flock to the first answer option that they read. Using data from seven original forced-choice conjoint experiments (and seven rank-based conjoint experiments) in five different languages across six countries, I find evidence that inattentive respondents pick the first option that they read. I show that this results in attenuation bias. To see why, consider that if all respondents chose the same response (e.g. option A) across all conjoint rounds, the average marginal component effect (AMCE) across all attributes would go to zero. Thus, the inclusion of inattentive respondents in a forced-choice conjoint experiment attenuates the estimates.² These findings have consequences for the interpretation of the magnitude

²Since inattentive respondents tend to go for the neutral non-attitude options (Alvarez et al., 2019), this finding also applies to rank-based conjoint experiments.

of the AMCEs for both forced-choice and rating-based conjoint experiments: scholars are likely underestimating the strength of their findings.³

The rest of this research note proceeds as follows. First, using replication data from more than twenty political science forced-choice conjoint experiments, I demonstrate that there is indeed imbalance in the distribution of answer options that respondents pick. Second, by employing data from seven original forced-choice conjoints, I show that inattentive respondents are driving the imbalance. Third, I present an experimental design that is aimed to capture the mechanism behind behavior of inattentive respondents. In particular, I check whether it is the language of instruction or the survey design that effects the primacy bias.

2 Prevalence of Skewed Distributions

In a forced-choice conjoint experiment with two outcome levels (say, “Candidate A” and “Candidate B”), in expectation, thanks to randomization, respondents should pick among the two options at the same rate. However, I find that this is not always the case. Among those taking a survey in English, German, Spanish, Dutch, Chinese and other languages respondents are disproportionately likely to pick the left-most choice or top answer option (when answer options are presented vertically), while those taking the survey in Arabic and Urdu are disproportionately likely to pick the right-most choice, which corresponds with the first option they encounter in the direction in which they read: left to right, and right to left, respectively.

³There are a couple of avenues that experimenters could take to tackle the issue of inattentive respondents in conjoint experiments. In the absence of instructional manipulation checks (IMCs) (Alvarez et al., 2019) and pre-treatment mock vignette checks (MVCs) (Kane et al., 2023) and if multiple rounds have been conducted, researchers could post factum run a simple sensitivity analysis and check how their results would change with varying proportion of inattentive respondents. One could also check whether the distribution of answer options is balanced across tasks. Finally, at the survey designing stage, I suggest including IMCs and/or MVCs, and test whether survey-takers failing IMC/MVC creates an imbalance in the proportion of options picked. If imbalance is detected, researchers could exclude inattentive respondents from the analysis.

To capture evidence of primacy bias in forced choice conjoint experiments, I have examined replication data from 21 studies. I focused on recent papers from some of the top political science journals (*American Political Science Review*, *American Journal of Political Science*, *Journal of Politics*, *Political Science Research and Methods*, *Journal of Experimental Political Science*). There are many more conjoint studies published in these journals. I focus on papers in which I could identify which answer option respondents read first (i.e., I present results for papers in which the profiles were identified as “left”, “right”, “1”, “2” etc.) Whenever possible, I also teased out what the survey layout was like and if the answer options were presented horizontally (left-right) or vertically (top-down). If an experiment was run on two samples, I examine the distribution of answers for two samples separately.

For many of the studies, it is clear that the distribution of answers *in the first round* is skewed. Across the surveys, respondents disproportionately pick the first option (e.g., option on the left or top answer option). I run a simple chi-squared test to establish whether a skew is statistically significant. In 19 experiments out of 24, this result is statistically significant and persists across languages, countries, survey companies and number of attributes. My suspicion is that the first round of a conjoint experiment confuses respondents which is why they gravitate towards the first option on the screen (or the first option that they read). For instance, in Holland et al. (2021) study, the same experiment was run on Colombian citizens in Colombia and on Venezuelan migrants residing in Colombia. In the first round, 527 Colombian citizens out of a thousand, picked option A. In stark contrast, in the first round, 919 Venezuelan migrants out of 1605, picked option A. “[T]he vast majority [of the Venezuelan participants] worked in the informal economy in Colombia” (Holland et al., 2021), which may suggest that the Venezuelan sample was bearing more cognitive load during the survey and, thus, picked the first option to proceed (see Table 1).

In a few of the studies, the skew persists across multiple rounds: respondents kept picking the first option disproportionately more often. In some cases, many respondents kept “straightlining”, i.e., picking the same option (A or B) across multiple rounds. For instance, in Kao et al. (2024), the proportion of respondents that selected the first answer option across all six rounds is 3.9% (the expected value is $(\frac{1}{2})^6 = 1.5625\%$; 2.7% of the respondents picked the second answer option exclusively). Similarly, in Dai and Kustov (2023), 10.5% of respondents picked the first option in all four rounds (the expected value is $(\frac{1}{2})^4 = 6.25\%$). Interestingly, respondents flock to the first option they view even when presented with pictorial treatment instead of text even in the presence of enumerators. Clary and Siddiqui (2021) conducted a conjoint experiment in the fall of 2015 in Pakistan in which they employed pictures instead of text to present respondents with profiles. Still, there is evidence of primacy bias: 23% of the respondents picked the first option in all three rounds ($(\frac{1}{2})^3 = 12.5\%$ is the expected value).

What explains variation in primacy bias? There are too few experiments presented here to establish conclusively (I work with more replication data and I will add those results in the future iteration of the paper), but one factor that could potentially predict primacy bias is the type of sample. Largely, whenever experiment targets highly selective groups, the primacy effect dies down. For instance, one arm of Bush et al. (2023) experiment targets international development workers, a highly educated and small sample, that is less likely to be inattentive (or to be confused by the layout of a conjoint experiment). Fabbe et al. (2023) recruited Greek municipal officials via email and the response was not very high (and neither was sample size) which implies that only highly motivated and interested officials took the survey. In both cases, primacy bias was not detected (see Figure 1). Respondents not recruited via survey companies are likely not driven by a promise of monetary compensation and, therefore, they rush less and they take surveys in a more meaningful way.

paper	Journal	Publication Year	Survey time	Sample size	Country	Target population	Survey company	Language	Layout presented	number of rounds	attribute number	max level number	straightlining percentage	Round 1 Skew	Round 2 Skew	Round 3 Skew	Round 4 Skew	Round 5 Skew	Round 6 Skew	Round 7 Skew	Round 8 Skew	Round 9 Skew	Round 10 Skew	Round 11 Skew	Round 12 Skew	Round 13 Skew	Round 14 Skew	Round 15 Skew
Clary and Siddiqui 2021	FPA	2021	2015	1990	Pakistan	Adults	PIPP/Gallup	Urdu	No	3	5	4	23%	p<0.01 (Option 1)	p<0.01 (Option 1)	p<0.01 (Option 1)												
Dai and Kustov 2023	PSRM	2023	2022	1004	US	Adults	Mturk	English	Yes (top down)	4	4	3	10.50%	p<0.01 (Option 1)	no	p<0.01 (Option 1)	p<0.01 (Option 1)											
Hutchinson et al. 2022	JEPS	2022	2020	1938	US	Adults	Lucid	English	Yes (top down)	4	5	3	10.10%	p<0.01 (top)	no	p<0.01 (top)	p<0.01 (top)											
Mazumder and Yan 2023	APSR	2023	2020	1002	US	Adults	YouGov Blue	English	No	4	18	6	8.30%	p<0.05 (Option 1)	p<0.05 (Option 1)	no	no											
Stoetzer et al. 2023	PSRM	2023	2020	6415	Brazil, Germany, Italy, Poland, US	Adults	Korpus	Portuguese, German, Italian, Polish and English	Yes (top down)	4	7	5	6.30%	p<0.01 (Option 1)	no	no	no											
Silva et al. 2022	JEPS	2022	2019	1505	US	Adults	Lucid	English	Yes (top down)	5	6	8	5.80%	p<0.01 (top)	no	no	no	no										
Senninger and Bischof 2023	PSRM	2023	2019	993	Germany	Adults	Clickworker	German	Yes (L-R)	5	3	3	3.50%	p<0.01 (Option 1)	no	no	no	no										
Bremer and Burgisser 2023	PSRM	2023	2018	3600	Germany, Italy, UK	Adults	Qualtrics	German, Italian and English	Yes (top down)	5	6	3	5.40%	p<0.01 (Option 1)	p<0.01 (Option 1)	no	p<0.01 (Option 1)	no										
Peyton et al. 2022	JEPS	2022	2020	1100	US	Adults	Lucid	English	No	5	9		7.20%	p<0.05 (Option 1)	no	p<0.05 (Option 1)	no	no										
Holland et al. 2024	JoP	2024	2019	1000	Colombia	Colombian citizens	CNC	Spanish	No	5	7	4	5.60%	p<0.1 (Option 1)	no	no	no	p<0.1 (Option 1)										
Holland et al. 2024	JoP	2024	2019	1600	Colombia	Venezuelan migrants	CNC	Spanish	No	5	7	4	6.60%	p<0.01 (Option 1)	no	p<0.1 (Option 1)	p<0.05 (Option 1)	no										
Bracic et al. 2023	JoP	2023	2019	4042	US	Adults	Lucid Theorem	English	No	5	7	10	4.50%	p<0.01 (Option 1)	no	p<0.01 (Option 1)	p<0.01 (Option 1)	no										
Kao et al. 2024	PSRM	2024	2017	1100	China	Adults	Survey Sampling International	Chinese	Yes (L-R)	6	4	5	3.80%	p<0.01 (Option 1)	no	p<0.01 (Option 1)	p<0.01 (Option 1)	p<0.01 (Option 1)	p<0.01 (Option 1)									
van der Does and Kantorowicz 2023	PSRM	2023	2019	3246	Netherlands	Adults	Dynata	Dutch	Yes (L-R)	6	5	8	3.50%	p<0.01 (Option 1)	p<0.01 (Option 2)	p<0.05 (Option 1)	p<0.01 (Option 1)	no	no									
Arias and Blair 2022	JoP	2022	2019	1086	the US	Adults	Dynata	English	No	9	7	5	1.70%	p<0.05 (Option 1)	no	no	no	no	no	no	no	no						
Arias and Blair 2022	JoP	2022	2019	1074	Germany	Adults	Dynata	German	No	9	7	5	1.70%	p<0.01 (Option 1)	no	no	no	no	no	no	no	no						
Bowen et al. 2023	JoP	2023	2019	2138	US	Adults	Lucid	English	No	12	6	3	~0%	p<0.01 (Option 2)	p<0.01 (Option 2)	p<0.01 (Option 2)	p<0.01 (Option 2)	p<0.01 (Option 2)	p<0.01 (Option 2)	no	no	no	no	no	no			
Driscoll and Nelson 2023	JoP	2023	2019	1000	US	Adults	Mturk	English	No	15			~0%	p<0.1 (Option 1)	no	p<0.1 (Option 2)	no	no	no	no	no	no	no	no	p<0.01 (Option 1)	no	no	no
Bush et al. 2023	APSR	2023	2019	108	NA	international development workers	email	English	Yes (top down)	6	6	3		no	no	no	no	no	no									
Bush et al. 2023	APSR	2023	2021	631	US	Adults	Prolific	English	Yes (top down)	6	6	3		no	no	no	no	no	no									
Fabbe et al. 2023	JoP	2023	2020	586	Greece	Greek municipal officials	AAPOR	Greek	No	3	5			no	no	no												
Boittin et al. 2023	APSR	2023	2017	1065	Nepal	Nepali police officers	in person	Nepalese	No	2	6	7		no	no													

Figure 1: Across studies that target adult population, there is a statistically significant skew towards the first option that respondents read in the first round of tasks.

Inattentive respondents are rarely identified in the replication data, however, so to check whether inattentiveness is at the core of primacy bias (and the resulting attenuation), I compare the original data from Hainmueller and Hopkins (2015) to Peyton et al. (2022) replication of the same study during Covid. Peyton et al. (2022) demonstrate that Covid lockdown has led to increasing lack of attention in respondents and moderation in estimates. While Peyton et al. (2022) rightly ascribe attenuation to inattentiveness, I suggest a potential mechanism behind this result. When comparing the distribution of selected options (Immigrant One vs Immigrant Two) across rounds in the original study and the replication study, it is clear that the distributions are more skewed during Covid which one could think of as a booster of inattention (see Tables 1 and 2).

Table 1: Distribution of option picks in the original immigration study by Hainmueller and Hopkins (2015). Pre-Covid respondents were more attentive. The distribution across answer options is roughly even.

Round	Immigrant 1	Immigrant 2	Chi-squared p value
1	718	678	0.280
2	677	719	0.260
3	708	688	0.590
4	701	695	0.870
5	738	658	0.032

Table 2: Distribution of option picks in the replication immigration study by Peyton et al. (2022). During Covid respondents were less attentive. Respondents picked the first option disproportionately more often in the first three rounds.

Round	Immigrant 1	Immigrant 2	Chi-squared p value
1	597	520	0.02
2	561	538	0.50
3	579	508	0.03
4	517	564	0.20
5	540	537	0.90

Another reason behind skewed distributions could be inexperience or confusion. Bansak

et al. (2018) demonstrate that conjoint experiments do not have a concerning issue of survey satisficing since respondents do not seem to demonstrate fatigue across thirty rounds of tasks. Authors support this result by showing that AMCE remains stable across rounds: after initial slight drop, the estimates do not attenuate. While it is hard to argue with that result, it is important to note that the respondents in the survey had previous experience with conjoint experiments. All 565 respondents from the MTurk sample have seen a “similar survey” four times before. Out of 1617 respondents from the SSI sample, 716 have seen a similar survey four times, 466 have seen it five times and 435 have seen it eight times or more. I compare SSI respondents who only saw a similar survey four times before to those who saw it five times or more. Less experienced respondents gravitate toward the left option in the first seven rounds.⁴

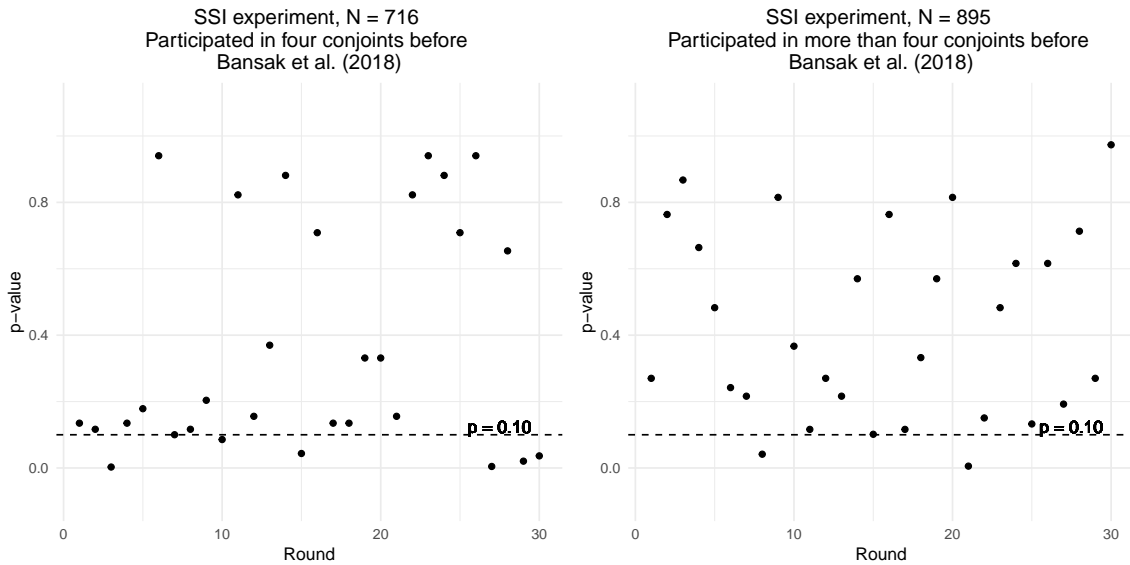


Figure 2: Presented are p-values from chi-square tests across thirty rounds of Bansak et al. (2018) study. Participants that have identified that they’ve seen a conjoint experiment only four times before flocked to the left-most answer option in the first seven rounds (left panel). For more experienced respondents the distribution between answer options is more balanced (right panel).

⁴The same is not observed for MTurk respondents potentially because the sample size is too small.

3 Inattentiveness and Skewed Distributions

It is not always tracked whether a respondent was attentive in the replication data. To explore whether disbalance in answer option choice is indeed driven by inattentiveness, I present the results on seven original forced choice conjoint experiments conducted in February of 2023 and February of 2024 in Germany, India, Spain and US (I incorporate results on seven original ranking based conjoint experiments in the future iteration of this paper). The surveys targeted the general adult population in each country and each survey covered around 2000 respondents. Each experiment was a part of a bigger survey, i.e., experiments were “stacked” and respondents took a few experiments in sequence. This led to fewer tasks for each conjoint experiment (two to three rounds of tasks), but it also gives an opportunity to utilize the position of an experiment in the survey to track respondent fatigue. I keep track of the answer option layout, i.e., whether answer options were presented horizontally (from left to right) or vertically (from top to bottom). Lucid, Cint and Cloud Research survey companies were employed.

I measure respondent inattentiveness primarily via the amount of time spent on a round. I deem respondents in the first quartile of time spent on a round as “inattentive” (this is arbitrary, I will include robustness checks in the future draft). Whenever possible, I utilize pre- and post-treatment MVCs, mock vignette checks (Kane et al., 2023), short questions checking whether respondents retain information (unrelated to treatment) shared with them.⁵ Across the surveys conducted for this research, passing rate for MVCs fluctuates around 70%, and for instructional manipulation checks (IMCs) the value is above 90%. The result observed for the general population observed in the replication data persists: there is a skew towards the first answer option in the first round across experiments (see Figure 3).

⁵While respondents’ attentiveness was also tracked via instructional manipulation check (IMC) questions (Alvarez et al., 2019), I do not use them here, because almost every respondent passes those. IMCs usually take the form of giving respondents specific instructions, e.g., “if you’re reading this, please pick ‘strongly agree’ below.”

experiment	Sample	attn check type	Survey time	Sample size	Country	Survey company	Language	Layout	Rounds	Notes	straightlining percentage	Round 1 Skew	Round 2 Skew	Round 3 Skew	Round 1 time (sec)	Round 2 time (sec)	Round 3 time (sec)
G-S	all		2023	~2000	Spain	Cloud Research	Spanish	left-right	3	second survey out of four	16.60%	$p < 0.01$ (Option 1)	no	no	35	27	15
G-S	inattent	time (Q1)	2023	~2000	Spain	Cloud Research	Spanish	left-right	3	second survey out of four		$p < 0.01$	no	no			
G-G	all		2023	~2000	Germany	Cloud Research	German	left-right	3	second survey out of four	18%	$p < 0.01$ (Option 1)	no	$p < 0.01$ (Option 1)	34	19	16
G-G	inattent	time (Q1)	2023	~2000	Germany	Cloud Research	German	left-right	3	second survey out of four		$p < 0.01$ (Option 1)	$p < 0.1$ (Option 2)	$p < 0.01$ (Option 1)			
K-U2	all		2023	~2000	US	Lucid (Cint)	English	top-down	2	first out of seven	32.60%	$p < 0.01$ (Option 1 <first outcome question)	$p < 0.01$ (Option 1 <first outcome question)		111	106	
K-U2	inattent	time (Q1)	2023	~2000	US	Lucid (Cint)	English	top-down	2	first out of seven		$p < 0.01$	$p < 0.01$				
K-U2	inattent	post-MVC (80% clearance)	2023	~2000	US	Lucid (Cint)	English	top-down	2	first out of seven		$p < 0.01$	$p < 0.01$				
K-U2	inattent	post-MVC (47% clearance)	2023	~2000	US	Lucid (Cint)	English	top-down	2	first out of seven		$p < 0.01$	$p < 0.01$				
K-U2	inattent	post-MVC (47% clearance)	2023	~2000	US	Lucid (Cint)	English	top-down	2	first out of seven		$p < 0.01$	$p < 0.01$				
T-GA	all		2024	~2000	Germany	Cloud Research	German	left-right	2	second out of three	30%	$p < 0.01$ (Option 1 <first outcome question)	no		63	32	
T-GA	inattent	time (Q1)	2024	~2000	Germany	Cloud Research	German	left-right	2	second out of three		$p < 0.01$	$p < 0.01$				
T-GA	inattent	pre-MVC (72% clearance)	2024	~2000	Germany	Cloud Research	German	left-right	2	second out of three		$p < 0.01$	$p < 0.01$				
T-GA	inattent	pre-MVC (76% clearance)	2024	~2000	Germany	Cloud Research	German	left-right	2	second out of three		$p < 0.01$	$p < 0.1$				
S-GA	all		2024	~2000	Germany	Cloud Research	German	top-down	2	third out of three	28.60%	$p < 0.01$ (Option 1)	$p < 0.01$ (Option 1)		45	36	
S-GA	inattent	time (Q1)	2024	~2000	Germany	Cloud Research	German	top-down	2	third out of three		$p < 0.01$	$p < 0.1$				
S-GA	inattent	pre-MVC (72% clearance)	2024	~2000	Germany	Cloud Research	German	top-down	2	third out of three		$p < 0.01$	$p < 0.05$				
S-GA	inattent	pre-MVC (76% clearance)	2024	~2000	Germany	Cloud Research	German	top-down	2	third out of three		$p < 0.01$	$p < 0.05$				
A-GB	all		2024	~2000	Germany	Cloud Research	German	top-down	3	first out of four	14.70%	$p < 0.01$ (Option 1)	no	no	52	27	24
A-GB	inattent	time (Q1)	2024	~2000	Germany	Cloud Research	German	top-down	3	first out of four		$p < 0.01$	no	$p < 0.01$			
A-GB	inattent	pre-MVC (68% clearance)	2024	~2000	Germany	Cloud Research	German	top-down	3	first out of four		$p < 0.01$	no	$p < 0.01$			
B-UC	all		2024	~2000	US	Cint	English	top-down	3	fifth out of five	14.40%	$p < 0.01$ (Option 1 <first outcome question)	no	no	42	32	27
B-UC	inattent	time (Q1)	2024	~2000	US	Cint	English	top-down	3	fifth out of five		$p < 0.01$	$p < 0.1$	$p < 0.05$			
A-I-TD	all		2024	~2000	India	Cloud Research	English	top-down (but also earlier)	3	fifth out of five		$p < 0.01$ (Option 1 <first outcome question)	$p < 0.01$ (Option 1 <first outcome question)	$p < 0.01$ (Option 1 <first outcome question)	71	39	31
A-I-TD	inattent	time (Q1)	2024	~2000	India	Cloud Research	English	top-down (but also earlier)	3	fifth out of five		$p < 0.01$	$p < 0.01$	$p < 0.01$			
A-I-LR	all		2024	~2000	India	Cloud Research	English	left-right (but also later)	3	fifth out of five		$p < 0.01$	$p < 0.01$	$p < 0.01$			
A-I-LR	inattent	time (Q1)	2024	~2000	India	Cloud Research	English	left-right (but also later)	3	fifth out of five		$p < 0.01$	$p < 0.01$	$p < 0.01$			

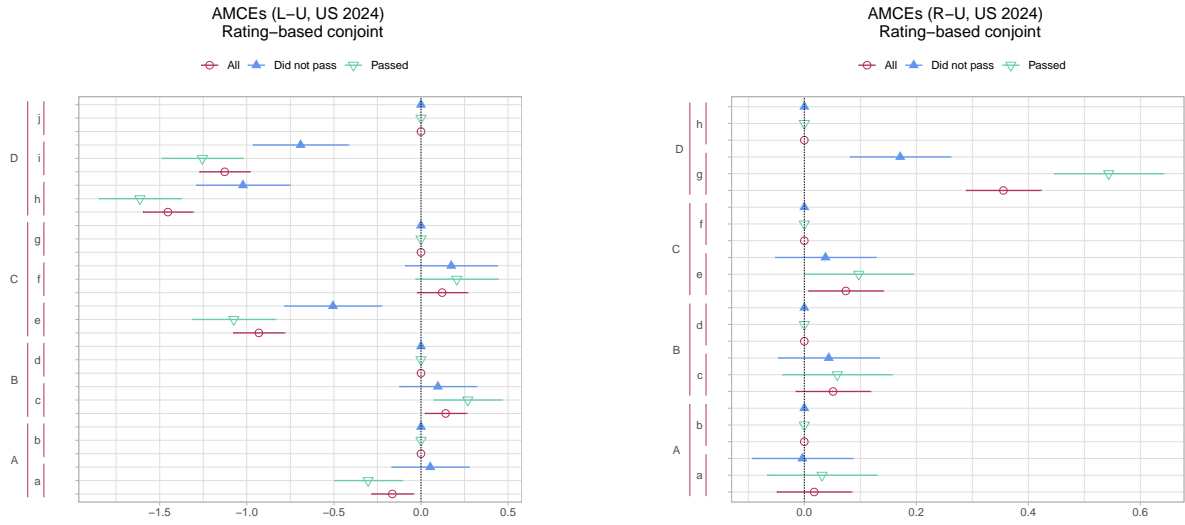
Figure 3: Across experiments and rounds, inattentive respondents flock towards the first option they read. This result is exacerbated by the time spent on each round of tasks.

Once the data are disaggregated into attentive and inattentive respondents, it is clear that the skew in answer distribution is driven by respondents who do not pay as much attention. In many cases, inattentive respondents disproportionately pick the first answer option across rounds. This result is moderated by two key factors that are both related to fatigue. First, position in the survey seems to play an important role in inducing inattention in respondents and the choice to pick the first answer option. Further down the experiment is, more time was spent on a survey, and less attention a respondent is paying. Second, the round duration also seems to affect respondents. If a task takes too much time to read and get through, the respondents are more likely to skip the question or pick the first answer option and move on. It is of interest that respondents adapt to the format of a conjoint experiment relatively fast. The time spent on a task drops across all experiments after the first round (e.g., in experiment T-GA, the average time per round halves and goes from 63 seconds in the first round to 32 seconds in the second round; see Figure 3).

Respondents are also less selective about answer options when it comes to less consequential questions. For instance, in the experiment S-GA respondents disproportionately picked the first option in both rounds potentially because the experiment asked respondents to pick between two cups of coffee (a choice that is less consequential than, say, picking between two hypothetical immigrants to be let into country). In experiment A-I-LR conducted in India, respondents were asked which of the two hypothetical migrants they would prefer to be resettled in Mumbai. When answering that question, respondents were more likely to go for the first answer option since respondents were predominantly not from Mumbai and did not really have an opinion about migration there.

Inattentive respondents also cause attenuation in rank-based conjoint experiments. Since inattentive respondents tend to go for the neutral non-attitude options (Alvarez et al., 2019), I find that inclusion of inattentive respondents in rating-based conjoint ex-

periments attenuates the results as well. In two separate rating-based conjoints conducted in the US in February of 2024, inattentive respondents were flagged via post-treatment memory checks (in experiment L-U, 60.4% passed the check and in R-U, 48.9% passed). The AMCE estimates for inattentive respondents are closer to zero than the full sample estimates, and AMCEs for attentive respondents are larger in magnitude than full sample estimates (see Figures 4a and 4b).



(a) Results for rating-based conjoint experiment, L-U. Conducted in the US in 2024. Sample split based on pre-treatment memory check which 60.4% of the respondents passed.

(b) Results for rating-based conjoint experiment, R-U. Conducted in the US in 2024. Sample split based on post-treatment memory check which 48.9% of the respondents passed.

Figure 4: Inattentive respondents lead to attenuation in rating-based conjoints as well.

4 How serious is the attenuation problem?

Having inattentive respondents that only pick the first option across all rounds in the sample leads to attenuation of AMCE. Intuitively, if all respondents picked the first option across all rounds, AMCEs would go to zero since none of the levels of any of the attributes would be effecting a preference. Also, when estimating AMCE, (this specific type of) inattentive respondents would be contributing to the denominator, but not the numerator.⁶

One way to tackle this issue would be list-wise deletion of inattentive respondents. This updated estimate would be reverse proportionate to the ratio of attentive respondents in the sample. If the proportion of inattentive respondents in the sample was p , then $AMCE_\alpha^u$ for some level α would be $AMCE_\alpha^u = \frac{1}{(1-p)}AMCE_\alpha^b$, where $AMCE_\alpha^b$ is the estimate for the full sample and $AMCE_\alpha^u$ is the updated estimate. As for standard errors, of course, $s.e.(AMCE_\alpha^u) = \frac{1}{(1-p)}s.e.(AMCE_\alpha^b)$. In other words, while the “true” AMCE is larger, so is the confidence interval around it (which follows from losing observations). If 4% of the sample are inattentive respondents that keep picking the same option⁷, then AMCE is attenuated roughly by 4% (Peyton et al. (2022) come to a similar conclusion). Importantly, if this is the type of careless respondent that is prevalent in the sample,

⁶Suppose there was an attribute with two levels: α and β . Then marginal mean for β would be $MM_\beta = \frac{0.5N_I\#\beta + \sum_A \mathbb{I}(choice=\beta)}{(N_I+N_A)\#\beta}$ where N_I is the number of inattentive respondents picking the first option and N_A is the number of attentive respondents manifesting true preferences. $\#\beta$ is the number of profiles with β level present in them (inattentive respondents pick those profiles half of the time) and $\sum_A \mathbb{I}(choice=\beta)$ is the number of profiles with level β present in them that attentive respondents picked. $MM_\alpha = \frac{0.5N_I\#\alpha + \sum_A \mathbb{I}(choice=\alpha)}{(N_I+N_A)\#\alpha}$ is the marginal mean for α . If β was the reference level, then AMCE for α is nothing but $AMCE_\alpha^b = \frac{\sum_A \mathbb{I}(choice=\alpha) - \sum_A \mathbb{I}(choice=\beta)}{(N_I+N_A)\#\alpha}$. Without inattentive respondents, the AMCE for α would be $AMCE_\alpha^u = \frac{\sum_A \mathbb{I}(choice=\alpha) - \sum_A \mathbb{I}(choice=\beta)}{(N_A)\#\alpha}$ since inattentive respondents do not add anything to the numerator. If p was proportion of inattentive respondents to the total number of respondents $p = \frac{N_I}{(N_I+N_A)}$, then $AMCE_\alpha^u = \frac{1}{(1-p)}AMCE_\alpha^b$.

⁷In three rounds the probability of observing three consecutive choices of the first option is $(\frac{1}{2})^3 = 12.5\%$; if the observed proportion is 16.5%, chances are, 4% of the respondents picked the first option due to inattentiveness

then estimates are likely attenuated, but statistical significance is probably not affected (confidence intervals scale by the same amount).

5 What Drives Primacy Bias: Research Design

While authors provide replication data for their conjoint experiments, they don’t always present the layout of the survey. Thus, although one can spot that the distributions between options A and B are not even, one can’t be sure what drives the skew. For instance, there is a clear skew in answer options distribution in Ridge (2024): Moroccan respondents who took the survey in Arabic clearly picked the right most option more often than the option on the left (see Table 3). However, it is not clear whether this is the result of the language of instruction or the survey layout. In this particular instance, the “Next question” button is situated right below “Option 1” (the right most option), which could result in inattentive respondents choosing it to save time (see Figure 5).

Table 3: Distribution of option picks in Ridge (2024). Survey conducted in Arabic. Option A (Option 1) was to the right of Option B (Option 2) in the survey (see Figure 3).

Round	Chose A	Chose B	Chi-squared p value
1	1,098	920	0.0001
2	1,078	942	0.002
3	1,096	920	0.0001
4	1,134	884	0.00000
5	1,058	960	0.029

To tease out the mechanism behind the exact observed discrepancies, I propose a simple factorial experiment. I test whether inattentive respondents are driven either by the language of instruction or by the layout of the survey. Inattentive respondents are likely either picking the first option that they read or the option that is closest to the “Next Question/Skip Question” button.

SET 1 :: Which of these two political systems of government would you prefer the most?

	OPTION 1	OPTION 2
Elections	There are elections with multiple recognized political parties	There are elections with multiple recognized political parties
Citizen participation	Some barriers to citizen political participation	Many barriers to citizen political participation
Official religion	There is no official state religion	There is no official state religion
Role for religious leaders	No formal role for religious leaders in government	No formal role for religious leaders in government
Provision of public services	Good government provision of basic items (ie housing and food) to individuals	Little government provision of basic items (ie housing and food) to individuals
Unemployment rate	Low unemployment (7%)	High unemployment (14%)
	OPTION 1	OPTION 2



مجموعة 1: أي من بين نظامي الحكومة السياسيين التاليين تفضل أكثر؟

خيار 2	خيار 1	
اجراء الانتخابات لحزب سياسي واحد معترف به	لا يوجد انتخابات	الانتخابات:
وجود بعض العوائق التي تخص مشاركة المواطنين في الحياة السياسية	وجود العديد من العوائق التي تخص مشاركة المواطنين في الحياة السياسية	مشاركة المواطنين في الحياة السياسية :
القرار دين رسمي للدولة	لا يوجد دين رسمي للدولة	الدين الرسمي للدولة:
أن يكون لدى رجال الدين دورا رسميا في الحكومة	لا يلعب رجال الدين دورا رسميا في الحكومة	دور رجال الدين :
توفير الحكومة للخدمات الأساسية (مثل: المأكل والسكن) بشكل جيد للأفراد	توفير الحكومة للخدمات الأساسية (مثل: المأكل والسكن) بشكل جيد للأفراد	توفير الخدمات العامة:
معدل البطالة منخفض (7%)	معدل بطالة منخفض (7%)	معدل البطالة:
خيار 2	خيار 1	



Figure 1. Conjoint choice task screen. NB: The tables are randomly generated. The English-language text is not a translation of the Arabic text.

Figure 5: Survey layout in Ridge (2024). Survey conducted in Morocco in Arabic. Respondents picked option A (Option 1) disproportionately more often. This could be driven by the language of instruction (Option 1 is read first in Arabic) or the position of the “Next” button (it’s closer to Option 1).

To empirically test this, I leverage the sizable bilingual population in Pakistan who speak both Urdu and English.⁸ Given that in Urdu people read from right to left, and in English they read from left to right, I manipulate the language of instruction thereby checking whether it is indeed the first read item that inattentive respondents gravitate towards. Half of the respondents that identify themselves as bilingual will be assigned to the English language of instruction and the other half will be assigned to the Urdu language of instruction. The position of the “Next/Skip Question” button will also be randomly assigned: for half of the respondents, the button will be to the right of the right most option (no matter the language)⁹ and for the other half, the button will be situated between the two choice options. The findings of the study would have implications for survey layout in conjoint experiments as well as the magnitudes of the effects derived from conjoint studies. This experiment received institutional review board (IRB) approval from New York University Abu Dhabi #HRPP-2024-79 and is to be pre-registered at OSF.

I target the population of Pakistan and assign half of the bilingual respondents to the survey in English and half of them to the survey in Urdu. The answer options will be laid out horizontally and not from top to bottom.¹⁰ I further split each language group into two. The first half will see the “Next Question/Skip Question” button between two options A and B (at the same distance away from A and B). The other half will see the “Next” button to the right of option B (see Table 4). Respondents that are not bilingual will take the survey in their preferred language and still would be randomly assigned to either (i) centered “Next” button group or (ii) group that has “Next” button to the right of option B.

⁸In Wave 6 of World Values Survey (WVS), although the most popular language spoken at home in Wave 6 of Pakistan was Punjabi, majority of the sample took the interview in Urdu. In Wave 7 of WVS, for 1796 out of the 1995 Pakistani respondents the interview language (Urdu) did not match the language they spoke at home.

⁹I may also put the “Next/Skip Question” away from the direction from which respondents begin reading, e.g., the button would be on the left for the Urdu group and on the right for the English group

¹⁰To ensure that the respondents taking the survey on mobile devices do not present a challenge to this part of the design I move away from the “mobile friendly” version of the layout.

	Urdu	English
“Next” button on the left	<i>Not needed (implicitly covered by “Next” button being on the left)</i>	<i>Not needed (implicitly covered by “Next” button being on the left)</i>
“Next” button in the middle	Under language hypothesis, more people will pick “Right”	Under language hypothesis, more people will pick “Left”
“Next” button on the right	Under layout hypothesis, more people will pick “Right”	Under layout hypothesis, more people will pick “Right”

Table 4: Some of the expectations of choice distributions under various hypotheses.

- Hypothesis 1 (not-at-random picks): Inattentive respondents pick option choices not at random
 - Hypothesis 2 a (language driven choice): Inattentive respondents pick the first option that they read
 - Hypothesis 2 b (layout driven choice): Inattentive respondents pick the option that is closest to the “Next” button on the webpage

If the inattentive respondents are driven purely by the language of instruction (and not survey layout), then the position of the “Next” button should not affect the distribution between choices A and B. If, on the other hand, the inattentive respondents are driven purely by the survey layout, they would consistently pick the option closer to the “Next” button no matter the language. Finally, if inattentive respondents are driven both by the language in which they take the survey and the survey layout, then respondents taking the survey in Urdu will select option on the right more often when the “Next” button is put between options A and B. Likewise, when the “Next” button is put between options A and B, respondents taking the survey in English will select option on the left more often (see Table 4).

Finding support for Hypothesis 1 (not-at-random picks) would have implications on how to properly adjust AMCEs. If there is no support for not-at-random picks, then

AMCEs need to be adjusted for measurement error via Clayton et al. (2023) method. However, if respondents indeed pick options not randomly, then AMCEs need to be adjusted for attenuation bias.¹¹ Importantly, identifying the mechanism behind choices of inattentive respondents would have implications for how to properly lay out conjoint experiments.

I spot inattentive respondents in the survey using a gamut of methods. Before exposing them to treatment I will present respondents with IMCs, instruction manipulation checks (Alvarez et al., 2019) as well as MVCs, mock vignette checks (Kane et al., 2023). I will also give the respondents the same task in the first and final (six) round to check if the respondents are consistent in their choices (Clayton et al., 2023). Finally, to test whether it is indeed the reading direction that shapes inattentive respondents' choices, I present the respondents with the same task in the first and an intermediate (fourth) round, but with the profiles flipped (i.e., what was scenario A in round one would be scenario B in round four; under one of the hypotheses an inattentive respondent would pick the same option – the option that they read first – in both cases). In addition, I measure how long respondents spend on each round. Finally, I check how well the respondents remember the last task (I ask what the outcome question was and whether a certain level was present in the task).

The conjoint experiment that the respondents will be exposed to would be the replication of Arias and Blair (2022) experiment on climate migration (Arias and Blair, 2022) (see Table 5). Respondents will be exposed to six tasks. Inattentiveness will be induced via positioning of the experiment in the survey: Arias and Blair (2022) replication experiment will be randomly presented to the respondents before or after an additional experiment on intimate partner violence (IPV).

Select parts of the survey instrument are presented in the next section.

¹¹Currently working on the demonstration of attenuation bias and bias-adjusted estimator.

Attribute	Levels
Reason for migration	Economic opportunity, drought, flooding, wildfires, political/religious/ethnic persecution
Language fluency	None, broken, fluent
Gender	Female, male
Occupation	Unemployed, cleaner, teacher, doctor
Religion	Atheist/agnostic, Christian, Muslim
Origin	Another region in your country, Afghanistan, Ethiopia, Myanmar, Ukraine
Vulnerability	None, posttraumatic stress disorder, food insecurity, no surviving family members, physically handicapped

Table 5: Attributes for replicating Arias and Blair (2022)

6 Survey Instrument

NB: The following is a trap question with a position of the instruction manipulation check randomized in a grid (Alvarez et al., 2019). The order of items would be randomized. Question borrowed from World Values Survey Wave 6.

C_IMC Please tell us for each of the following actions whether you think it can always be justified, never be justified, or something in between.

	Never justifiable									Always justifiable
Claiming government benefits to which you are not entitled	1	2	3	4	5	6	7	8	9	10
Avoiding a fare on public transport	1	2	3	4	5	6	7	8	9	10
Stealing property	1	2	3	4	5	6	7	8	9	10
Cheating on taxes if you have a chance	1	2	3	4	5	6	7	8	9	10
Someone accepting a bribe in the course of their duties	1	2	3	4	5	6	7	8	9	10
Homosexuality	1	2	3	4	5	6	7	8	9	10
Prostitution	1	2	3	4	5	6	7	8	9	10
Abortion	1	2	3	4	5	6	7	8	9	10
Divorce	1	2	3	4	5	6	7	8	9	10
Sex before marriage	1	2	3	4	5	6	7	8	9	10
Suicide	1	2	3	4	5	6	7	8	9	10
Euthanasia	1	2	3	4	5	6	7	8	9	10
For a man to beat his wife	1	2	3	4	5	6	7	8	9	10
Parents beating children	1	2	3	4	5	6	7	8	9	10
Violence against other people	1	2	3	4	5	6	7	8	9	10
For quality purposes, please select “always justifiable”	1	2	3	4	5	6	7	8	9	10

- **NB:** The following are mock vignette checks (Kane et al., 2023). The order of answer options would be randomized. This question is also used to spot whether respondents have natural proclivity for left-most or right-most answer options.

C_MVC Please carefully read the following stories and click either button to proceed.

Scenario A	Scenario B
Noor's brother sometimes pays for the meals at a restaurant. Noor's mother thinks that it's okay.	Noor's sister often pays for the coffee whenever they go out. Noor's father is not sure whether that's appropriate.

Scenario A

Scenario B

Next Question

- *C_MVC_Q1* In the stories you just read, the name of the main character was...
 - Kausar
 - Noor
 - Amal
 - Islam
- *C_MVC_Q2* In the stories you just read, who was NOT one of the characters?
 - Noor's brother
 - Noor's sister
 - Noor's cousin
 - Noor's father

The following is the template for the scenarios presented to the respondents in the intimate partner violence (IPV) part of the experiment (levels of attributes presented in brackets):

A friend approaches you and tells you about their cousin. Your friend tells you that the cousin's [husband/ex-husband/partner] has recently [assaulted her/cussed her out]. [This happens routinely/This happens occasionally/This is the first time it happened]. Your friend [tells you that family and friends support their cousin's decision to reach out to somebody/tells you that family and friends doubt that their cousin should reach out to somebody].

Example task below:

- Language of instruction: reading from left to right
- "Next" button position: in between Options A and B
- Outcome question: police

Scenario A	Scenario B
A friend approaches you and tells you about their cousin. Your friend tells you that cousin's husband has recently assaulted her. This happens occasionally. Your friend tells you that family and friends doubt that their cousin should reach out to somebody.	A friend approaches you and tells you about their cousin. Your friend tells you that cousin's ex-husband has recently cussed her out. This happens routinely. Your friend tells you that family and friends support their cousin's decision to reach out to somebody.

Question: Which case do you think would be better handled by the police?

Scenario A

Scenario B

Next Question

Example task below:

- Language of instruction: reading from left to right
- “Next” button position: to the right of option B
- Outcome question: imam

Scenario A	Scenario B
A friend approaches you and tells you about their cousin. Your friend tells you that the cousin’s partner has recently assaulted her. This is the first time it happened. Your friend tells you that family and friends doubt that their cousin should reach out to somebody.	A friend approaches you and tells you about their cousin. Your friend tells you that the cousin’s ex-husband has recently assaulted her. This happens routinely. Your friend tells you that family and friends support their cousin’s decision to reach out to somebody.

Question: Which case do you think would be better handled by an imam?

Scenario A

Scenario B

Next Question

Question [attention check]: The question on the previous page was about...

- imam
- police

Example task below (borrowed from Arias and Blair (2022)):

- Language of instruction: reading from left to right
- “Next” button position: center

	Migrant 1	Migrant 2
Gender	Male	Female
Occupation	Unemployed	Teacher
Reason for migration	Flooding	Economic opportunity
Religion	Agnostic	Muslim
Language Fluency	None	Fluent
Origin	Another region in Pakistan	Myanmar
Vulnerability	Physically handicapped	No surviving family members

Question: Which of these two migrants do you prefer?

Migrant 1

Migrant 2

Next Question

Question [attention check]: One of the migrants on the previous page was from Ukraine.

- True
- False

References

- Abramson, S. F., Koçak, K., and Magazinnik, A. (2022). What do we learn about voter preferences from conjoint experiments? American Journal of Political Science, 66(4):1008–1020. 2
- Alvarez, R. M., Atkeson, L. R., Levin, I., and Li, Y. (2019). Paying attention to inattentive survey respondents. Political Analysis, 27(2):145–162. 2, 3, 9, 11, 18, 20
- Arias, S. B. and Blair, C. W. (2022). Changing tides: public attitudes on climate migration. The Journal of Politics, 84(1):560–567. 18, 19, 24
- Aronow, P. M., Baron, J., and Pinson, L. (2019). A note on dropping experimental subjects who fail a manipulation check. Political Analysis, 27(4):572–589. 2
- Atsusaka, Y. and Kim, S.-y. S. (2023). Addressing measurement errors in ranking questions for social sciences. Technical report, Working Paper. 2
- Bansak, K., Hainmueller, J., Hopkins, D. J., and Yamamoto, T. (2018). The number of choice tasks and survey satisficing in conjoint experiments. Political Analysis, 26(1):112–119. 2, 7, 8
- Blair, G., Chou, W., and Imai, K. (2019). List experiments with measurement error. Political Analysis, 27(4):455–480. 2
- Bush, S. S., Donno, D., and Zetterberg, P. (2023). International rewards for gender equality reforms in autocracies. American Political Science Review, pages 1–15. 5
- Clary, C. and Siddiqui, N. (2021). Voters and foreign policy: evidence from a conjoint experiment in pakistan. Foreign Policy Analysis, 17(2):orab001. 5
- Clayton, K., Horiuchi, Y., Kaufman, A. R., King, G., Komisarchik, M., Ebanks, D., Katz, J. N., King, G., Evans, G., King, G., et al. (2023). Correcting measurement error bias

- in conjoint survey experiments. American Journal of Political Science, 12(B2):1–11. 2, 18
- Dai, Y. and Kustov, A. (2023). The (in) effectiveness of populist rhetoric: a conjoint experiment of campaign messaging. Political Science Research and Methods, pages 1–8. 5
- Fabbe, K., Kyrkopoulou, E., Matakos, K., and Unan, A. (2023). Control and fairness: What determines elected local leaders’ support for hosting refugees in their community? The Journal of Politics, 85(2):778–783. 5
- Ganter, F. (2023). Identification of preferences in forced-choice conjoint experiments: Reassessing the quantity of interest. Political Analysis, 31(1):98–112. 2
- Hainmueller, J. and Hopkins, D. J. (2015). The hidden american immigration consensus: A conjoint analysis of attitudes toward immigrants. American journal of political science, 59(3):529–548. 7
- Hainmueller, J., Hopkins, D. J., and Yamamoto, T. (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. Political analysis, 22(1):1–30. 2
- Holland, A., Peters, M., and Zhou, Y.-Y. (2021). Left out: How political ideology affects support for migrants in colombia. 4
- Kane, J. V., Velez, Y. R., and Barabas, J. (2023). Analyze the attentive and bypass bias: Mock vignette checks in survey experiments. Political Science Research and Methods, 11(2):293–310. 2, 3, 9, 18, 21
- Kao, J. C., Lü, X., and Queralt, D. (2024). Do gains in political representation sweeten tax reform in china? it depends on who you ask. Political Science Research and Methods, 12(1):146–165. 5

- Peyton, K., Huber, G. A., and Coppock, A. (2022). The generalizability of online experiments conducted during the covid-19 pandemic. Journal of Experimental Political Science, 9(3):379–394. 2, 7, 13
- Ridge, H. M. (2024). Democratic commitment in the middle east: a conjoint analysis. Political Science Research and Methods, 12(2):285–300. 14, 15