Survey Research Methods (2023)
Vol. 17, No. 2, pp. 205-217
doi:10.18148/srm/2023.v17i2.7943
European Survey Research Association

Pseudo-Opinions in Online Surveys: Evidence to Recontextualize the Imputed Meaning Hypothesis

Henrik K. Andersen[1], Jochen Mayerl[1], Felix Wolter[2], and Justus Junkermann[3]
[1]Chemnitz University of Technology
[2]University of Konstanz
[3]University of Leipzig

Pseudo-opinions refer to survey respondents giving answers to topics they are unfamiliar with. They are widespread but the reasons why respondents do not just admit they "don't know" are not well-understood. We investigate the underlying mechanisms for pseudo-opinions in online surveys: do respondents satisfice and perform a "mental coin-flip," or do they optimize and attempt to "impute a meaning" to the unknown question and answer accordingly? And can we reduce the prevalence of pseudo-opinions by expressing to respondents that it is okay not to have an opinion? To do so, we use fictitious issues. These are survey questions about nonexistent topics and things. We use response latencies as an indicator for the mode of responding, on a continuum from automatic-spontaneous to controlled-deliberate, to investigate whether pseudo-opinions are the result of satisficing or optimizing. We also conduct a survey experiment in which the presence of an explicit "don't know" category is randomly assigned. The sample (n = 1288) consists of data collected in August 2019 from an online panel provider. The target population was defined as adults between 18-69 years old with internet access residing in Germany. Quotas were put in place for age and sex. We find pseudo-opinions are predicted by faster, automatic responses. This contradicts the widely-assumed imputed meaning model of pseudo-opinions. The presence of an explicit "don't know" category reduces pseudo-opinions dramatically but does not moderate the effect of deliberate or automatic responding on pseudo-opinions.

Keywords: Pseudo-opinions, fictitious issues, survey methodology, online surveys, response latencies

1 Introduction

It is widely recognized that survey respondents sometimes offer opinions on things they are wholly or mostly unfamiliar with. These are referred to as *pseudo-opinions* or *nonattitudes* (Bishop et al., 1983; Converse, 1964, 1970). When a respondent gives a substantive answer to a topic they are unfamiliar with, their response deviates from the true value, which is "I don't know." Thus, pseudo-opinions have the potential to threaten the validity and reliability of survey research.

However, the underlying mechanisms responsible for pseudo-opinions are not well-understood. Some argue they are the result of satisficing respondents who believe giving pseudo-opinions is the easiest way to get through the survey without much effort. This is referred to as the *mental coin-flip hypothesis*. Others maintain pseudo-opinions are the result of optimizing behavior and that respondents look for cues to interpret the unfamiliar question and form an ad hoc opinion. This alternative is referred to as the *imputed meaning hypothesis*.

Even though the imputed meaning hypothesis is currently favored in the literature, the empirical evidence for either model is rather sparse. And the evidence up until now focuses mostly on traditional computer-assisted personal interviewing (CAPI) or computer-assisted telephone interviewing (CATI) survey modes with interviewer interaction, where social desirability pressures likely play a greater role. But as online surveys grow increasingly popular in the social sciences, we need to scrutinize whether these previous findings are transferrable.

This is the research desideratum this paper wishes to address. The guiding research question for this paper concerns whether *pseudo-opinions* are the result of satisficing or optimizing behaviour? Answering the question of what drives pseudo-opinions is particularly relevant, because it casts them in one of two very different lights. If pseudo-opinions are linked to satisficing, then they are a source of random or even systematic error (i.e., response effects) and

Contact information: Henrik K. Andersen, Chemnitz University of Technology, Institute of Sociology, Thüringer Weg 9, 09126 Chemnitz (E-mail: henrik.andersen@soziologie.tu-chemnitz.de)

are therefore indeed a troubling issue for survey research. If pseudo-opinions result when respondents attempt to construct meaning around unfamiliar topics, then the responses might arguably be seen as valid opinions to more general constructs.

We use data collected in an online survey, and hypothesize that without an interviewer present (either in person or on the telephone), pseudo-opinions should primarily be driven by automatic, satisficing behavior. We use fictitious issues, i.e., questions about nonexistent topics, to be able to discern when a respondent is giving a pseudo-opinion; how can the respondent have an opinion when the object in question does not exist? Further, we use response latencies as a measure of the respondent's current motivation and opportunity and thus mode of responding, on a continuum between automatic/satisficing and deliberate/optimizing (Fazio, 1990), to investigate the underlying mechanisms for pseudo-opinions. Finally, we perform a survey experiment by randomly assigning respondents to an explicit vs. implicit "don't know" category to investigate the effect on pseudo-opinions. This has been shown in the past to be effective at reducing pseudo-opinions; a result we intend to replicate in an online survey setting. Further, it is possible that an explicit "don't know" category may change the way respondents react to unfamiliar topics, so we test whether the experiment moderates the effect of response latencies on pseudo-opinions.

The article is structured as follows: first we discuss the issue of pseudo-opinions and the two dominant mechanisms discussed in the literature. We argue that the evidence for the imputed meaning hypothesis is sparse and based on survey modes where the interviewee interacts directly with the interviewer; modes that are increasingly being abandoned in favor of online surveys. We then discuss our data and the items of interest: six fictitious organizations and institutions for the respondents to evaluate. Then, we explain our approach to the questions of whether pseudo-opinions are driven by satisficing or optimizing, and whether they can be mitigated using an explicit "don't know" category. We then present the results and discuss the implications.

2 Theoretical Background

Pseudo-opinions describe the situation in which respondents give substantive answers to topics that they know little to nothing about (Bishop et al., 1980; Schuman & Presser, 1980; Sturgis & Smith, 2010; Wolter & Junkermann, 2019). They are often seen as a serious problem as they threaten the validity and reliability of survey data (Grant & Patterson, 1975; Wolter & Junkermann, 2019). In other words, when a respondent gives a substantive answer to a topic that they are unfamiliar with, their response deviates from the "true" value, which should be "I don't know." In what way the substantive responses deviate from this true value is often unknown, however. Thus, random deviations will cause the aggregate data to be less reliable while systematic deviations will cause the aggregate data to be less valid.

Pseudo-opinions are troubling as a concept yet their prevalence may be even more distressing. Indeed, Sturgis and Smith (2010) found between 10–15% of respondents in their CAPI-based study gave pseudo-opinions, while earlier CATI-based studies by Bishop et al. (1980), Bishop et al. (1983), and Schuman and Presser (1980) found upwards of 30–50%, depending on the question content and specific experimental condition. An older study by Le Payne (1951) as well as a quite recent one from Wolter and Junkermann (2019) each found ca. 70% of responses were pseudo-opinions. Why the prevalence of pseudo-opinions in these studies varies so greatly likely has to do with the subject matter, mode and survey situation. But it seems plausible that most if not all survey-based studies are faced with the issue of pseudo-opinions.

2.1 Pseudo-opinions as a mental coin-flip

Converse (1964) first put forth the "mental coin-flip" model of pseudo-opinions. This was based on a panel study in which he observed very little stability in individual's attitudes while noting that the aggregate attitudes changed very little over time (Sturgis & Smith, 2010, p. 68). He concluded that the responses he was observing were nonattitudes and that the only way in which the individual's attitudes could change over time without the aggregate attitudes changing as well is if the individual changes were essentially random.

So, the mental coin-flip model of pseudo-opinions says that, when faced with an unfamiliar topic, respondents essentially choose randomly between response options (Converse, 1964, 1970). This view comes from the idea that pseudo-opinions are a way for respondents to conform with survey norms while avoiding probing by the interviewer (Grant & Patterson, 1975; Sturgis & Smith, 2010). In effect, they could be seen as a type of *satisficing behavior* (Krosnick, 1991; Vannette & Krosnick, 2014); an essentially random answer—a *mental coin-flip*—can be seen as the path of least resistance to some respondents (Converse, 1964, 1970; Sturgis & Smith, 2010).

Satisficing in the context of surveys refers to a set of strategies employed by respondents who are *weakly motivated, lacking opportunity*, or both, to get through the survey with as little effort as necessary. That is, a popular model of response behavior encompasses four stages: (1) reading and understanding the question, (2) retrieving relevant information from memory, (3) forming a judgment and (4) translating the judgment to fit one of the available response categories (Roberts et al., 2019; Tourangeau et al., 2000). So-called "weak satisficing" refers to responses that result when a respondent goes through all four stages but in a cursory manner, leaving room for cognitive biases (such as acquiescence, primacy, etc.). The respondents may begin to search their memory for salient attitudes or relevant information in order to form an ad hoc judgment, but they end the process once a "first-best" or satisfactory result is reached. In other words, instead of optimizing their responses by attempting to consider all relevant information and perspectives, they gather some salient impulses and cut the process off. So-called "strong satisficing," on the other hand, refers to a behavior in which the respondent skips either the information retrieval or judgment formation stages (Vannette & Krosnick, 2014). The respondent reads and comprehends the question and forms an immediate response without attempting to form a connection between the question and stored information and memories. If the respondent is unable to intuit any easy *cues* for an "effortless" response (Leiner, 2019; Roberts et al., 2019), they may select a response at random (i.e., a mental coin-flip), skip the question, or provide another kind of nonsubstantive response (e.g., "don't know" category).

Now, a relevant question is why a satisficing respondent would choose to provide a pseudo-opinion instead of simply skipping the question or answering honestly with "I don't know." One reason could be that respondents want to conform to survey norms (Grant & Patterson, 1975; Lenzner, 2011). In other words, the respondent expects that the interviewer, or by extension, survey researcher will only ask pertinent questions. In return, the respondent is expected to answer them (Lenzner, 2011). Another perhaps even more relevant reason could be that satisficing respondents know that a substantive response is the path of least resistance. Because "I don't know" responses may be perceived by the respondent as inadequate, they anticipate further probing from the interviewer or questionnaire, which, as satisficers, they want to avoid (Lenzner, 2011).

2.2 Pseudo-opinions as imputed meaning

More recently, however, it was suggested that pseudo-opinions may actually result from highly motivated respondents. One source of motivation could be social desirability pressures, for example (Sturgis & Smith, 2010). Respondents may try to avoid looking uninformed when they respond truthfully that they have no opinion on a given topic. Further, respondents might feel pressure to deliberate carefully about the potential meaning of the question in order to avoid giving an *undesirable* response, whatever they deem that to be in the situation. It is argued that pseudo-opinions are an attempt by respondents to understand what the question is about and answer accordingly.

Along this line of reasoning, the imputed meaning hypothesis refers to when respondents "make an educated (though wrong) guess as to what the obscure acts represent, then answer reasonably in their own terms about the constructed object"(Schuman & Presser, 1980, p. 1223). Schuman and Presser (1980) proposed this model of pseudo-opinions in direct contrast to Converse (1964, 1970) mental coin-flip hypothesis. They came to this alternative model when they observed respondents making spontaneous asides, apparently trying to verbally suss out the meaning of a question regarding the fictitious "Agricultural Trade Act." Not only that, the distribution of the responses for the fictitious act was *not* split 50:50 down the middle; rather, about two-thirds of respondents supported the act. Schuman and Presser (1980) reasoned that, just as there is an equal probability of getting either heads or tails, the mental coin-flip model should result in a uniform distribution and that this unequal distribution spoke against the model.

2.3 Sparse empirical evidence

These two competing views of pseudo-opinions are both arguably plausible, though the imputed meaning hypothesis seems to have won out in recent years (Bishop & Jabbari, 2001; Bishop et al., 1980; Schuman & Presser, 1980; Sturgis & Smith, 2010). However, a review of the literature reveals that empirical evidence for either hypothesis is rather sparse.

First of all, most of the literature mentioning the imputed meaning hypothesis points back to a limited number of rather old studies. The one that is cited perhaps the most is Schuman and Presser (1980). There, they argue that a mental coin-flip should mean that opinions for and against unfamiliar questions should be equal, just as a flip of a fair coin results with equal probability in either heads or tails. They note in their study that pseudo-opinions were not, however, equally expressed and that nearly two-thirds of respondents supported the fictitious "Agricultural Trade Act". This, they concluded, spoke against the mental coin-flip model.

But just because the pseudo-opinions were not equally distributed across response categories does not necessarily imply that respondents attempt to construct meaning when responding to an unfamiliar topic. Rather, they may employ any number of other satisficing-related heuristics to conform to survey norms (i.e., answering questions when asked without putting in too much effort. Roberts (2016) provides an overview of response styles and notes that random responding (another term for the mental coin-flip) is just one form of satisficing behavior besides acquiescence, mid-point and extreme responses, mild responses, etc. A mixture of these satisficing behaviors would lead to pseudo-opinions that were not uniformly distributed across response categories.

In another example, Sturgis and Smith (2010) note that substantive responses to obscure items were correlated with confidence in the government, and they took this as evidence against the mental coin-flip hypothesis. But this is hardly convincing on its own, as the mere presence of a correlation between pseudo-opinions and confidence in government says nothing about the underlying mechanisms of pseudo-opinions. Surely it is possible that satisficing respondents could be those who tend to express confidence in their government.

Beyond that, Schuman and Presser (1980) also mention anecdotal evidence in which the majority of 35 individuals recorded during the interview made asides, suggesting they were attempting to construct some meaning to the unfamiliar topics. However, Sturgis and Smith (2010) suggest that social desirability may motivate respondents to try to appear better informed than they are. It may also increase pressure on the respondent to at least appear to be conforming to survey norms by putting thought into their responses and not giving obviously bogus answers. Indeed, most previous research is based on survey modes with interaction between interviewee and interviewer. In online surveys, where perceived anonymity should be higher, social desirability pressures should play less of a role (Joinson, 1999). And online surveys are becoming increasingly popular in academic contexts in the social sciences (Bandilla, 2016; Shin et al., 2011; Silber et al., 2013). Due partially to the Coronavirus pandemic, for example, the German ALLBUS study[1] even moved away from face-to-face interviewing towards postal and web-based surveys for the first time in the 2021 release.

There are other plausible mechanisms that might increase a respondent's motivation to generate an ad hoc opinion on an unfamiliar issue, such as need for cognition, but so far social desirability is the only cause mentioned in the literature in the context of the imputed meaning hypothesis. We hypothesize, therefore, that respondents in online surveys have little incentive to give pseudo-opinions based on social desirability pressures and that the more automatic (satisficing) the response, the higher the likelihood of a pseudo-opinion. If pseudo-opinions are, contrary to our expectation, linked to deliberate/optimizing responses, then other mechanisms, like need for cognition, should be investigated in the future.

3 Data and Measures

To investigate the underlying mechanisms of pseudo-opinions, we conducted an online survey using an online panel provider from 16th to 25th August 2019. The target population was defined as adults between the age of 18 and 69 years with internet access residing in Germany with quotas for age and sex in place to ensure the sample was representative of the target population on those characteristics.

The pseudo-opinions portion of the study consisted of a question battery in which the respondents were asked using a binary scale to give us their opinion (mostly positive vs. mostly negative) on 14 different organizations and institutions. Eight of these truly existed: The German Armed Forces (Bundeswehr), Doctors without Borders, The United Nations, Greenpeace, The German Federal Criminal Police (Bundeskriminalamt), The Rosa-Luxemburg-Foundation (associated with the German left-wing political party, Die Linke), The Konrad-Adenauer-Foundation (associated with German Christian Democratic party, CDU), and the Intergovernmental Panel on Climate Change. Six were fabricated by us:

- Environmental Court (EC),
- Coastal Aid Agency (CAA),
- Prague Energy Transition Initiative (PETI),
- German Nuclear Forum (GNF),
- Herbert-Schmaar-Foundation (HSF), and
- World Space Agency (WSA).

This part of the study encompassed a split of 1,288 randomly chosen respondents out of the full sample of 3,044. At the end of the survey, we de-briefed the respondents and revealed to them that several of the items were fabricated.[2]

If the respondent gave a substantive answer (i.e., the reported opinion was either "mostly positive" or "mostly negative"), then this was considered a pseudo-opinion. Non-substantive responses were those in which the respondent skipped the question by clicking on the "continue" button or, if they were in the explicit experimental group, selecting the "don't know" option.

3.1 Measuring the mode of responding

We employ response latencies (RL) as a measure of the respondent's mode of responding, on a continuum from automatic/satisficing to deliberate/optimizing. Response latencies were measured passively through time stamps in milliseconds (each item was shown on a separate screen). To deal with outliers, we first removed responses over 2,000 seconds and then also those that were two standard deviations above the mean, as suggested by Mayerl and Urban (2008, p. 59).[3] This resulted in 76 individual observations

[1] ALLBUS is the abbreviation for "Allgemeine Bevölkerungsumfrage der Sozialwissenschaften" (English: General Population Survey of the Social Sciences), a regularly-conducted large-scale survey of the German population.

[2] The debriefing text, translated from German read: "You may have noticed that we asked you about several things that you did not recognize. This is because they do not exist. These were the Environmental Court, the Coastal Aid Agency, the Prague Energy Transition Initiative, the German Nuclear Forum, the Herbert-Schmaar-Foundation and the World Space Agency. We did this because often people express opinions about existent things even if they are unfamiliar with them. We want to find out why people sometimes give answers anyways. We hope that this will let us improve future surveys."

[3] Mayerl and Urban (2008) give an overview of various ways to treat response latency outliers found in the literature, as there is little consensus on which method to use. For example, another approach is to remove response latencies larger than 2 standard deviations above the median. We tested this method, as well. It resulted in 80 rather than 76 individual observations being removed and the results of the rest of the analysis are almost identical.

being removed and a range for the outlier-treated response latencies of [0.019, 35.640] seconds. We demean the individual response latencies (per individual, per item) so that they can be interpreted as responses for particular items that were faster/slower than the individual's average speed ("within-unit RL"). These are shown in a histogram on the left of Figure 1. As can be seen, the overwhelming majority of the responses deviate only slightly from the individual's average speed (standard deviation of 2.19 seconds).

The average response latencies per individual ("between-unit RL") are used to rule out confounding of the effects of the demeaned response latencies by unobserved item-invariant confounders (in the form of a "correlated random effects model," discussed below). Their effects can be interesting in their own right, however, as it is also possible that the individuals' average responses predict pseudo-opinions, rather than item-specific deviations from the average speed. These are shown on the right of Figure 1. The average response speed over all of the fictitious items was about 2.96 seconds with a standard deviation of 1.28 seconds.

Note that the within- and between-unit response latencies offer the opportunity to test slightly different hypotheses:

- Within-unit: when faced with a fictitious issue, fast/slow responses predict pseudo-opinions.
- Between-unit: fast/slow respondents tend to give pseudo-opinions to fictitious issues.

Fast responses are taken to indicate automatic/satisficing behavior, and slow responses indicate deliberate/optimizing behavior.

In the following section, we outline the research design with regard to the question of whether satisficing or optimizing responses are linked with pseudo-opinions, and whether an explicit "don't know" category can help mitigate them.

4 Research Design

In the following section, we outline the research design with regard to the question of whether satisficing or optimizing responses are linked with pseudo-opinions, and whether an explicit "don't know" category can help mitigate them.

4.1 Satisficing vs. optimizing

The dependent variable is a binary outcome (0: nonsubstantive, 1: substantive/pseudo-opinion), so we turn to a probit regression model. The level-one (within-unit) response latencies are continuous and vary across both individuals and items, while the level-two (between-unit) average response latencies per respondent vary only across individuals but not items. Likewise, the experimental treatment varies between individuals but does not change within individuals. With only six items, we can include dummy variables to control for the item specific effects, i.e., characteristics of particular items that inuence the way in which all respondents approach the question. We demean the level-one response latencies by subtracting the level-two response latency (the individual's average response latency) from each. This gives the effects of the level-one response latencies their interpretation as "within-effects," i.e., the effect of a response latency that deviates from the person's usual speed on the probability of a substantive response (Wooldridge, 2002). The effects of the level-two response latencies can be interpreted as "between-effects," i.e., the effect of a person that answers more or less quickly compared to other respondents on the probability of a substantive response (Bell & Jones, 2015; Schunck & Perales, 2017).

The modeling strategy can be described as a multilevel probit regression with correlated random effects (Mundlak, 1978; Wooldridge, 2002). We assume there is individual-specific unobserved heterogeneity that affects each individual's propensity for pseudo-opinions. Some of the factors making up the unobserved heterogeneity could be correlated with the independent variables of central interest; the response latencies. So, we include the person mean response latencies—the average speed per individual over the six items—to account for the correlation between response latencies and the unobserved heterogeneity. In a stripped-down model (ignoring control variables, interactions and item dummies for a moment), we would say

$$y^*_{ij} = \beta_1 d_i + \beta_2 \mathrm{rl}_{ij} + \alpha_i + \epsilon_{ij} \quad (1)$$

where y^*_{ij} is the continuous unobserved outcome from which we only observe the dichotomous y_{ij} and rl_{ij} are the response latencies per individual, per item. d_i is a dummy variable indicating the experimental group (implicit or explicit "don't know") and α_i is the unobserved individual-specific heterogeneity. We assume $\mathrm{Cov}\left(\alpha_i, \mathrm{rl}_{ij}\right) \neq 0$, the response latencies are correlated with the unobserved stable factors.

We decompose the correlation between the response latencies and the unobserved heterogeneity into the effect of the average response latency and a residual, $\alpha_i = \gamma\overline{\mathrm{rl}_i} + \nu_i$. The residuals are uncorrelated with $\overline{\mathrm{rl}_i}$ and thus $Cor\left(\nu_i, \overline{\mathrm{rl}_i}\right) = Cor\left(\nu_i, \mathrm{rl}_{ij}\right) = 0$ (since the correlation between rl_{ij} and α_i plays out on the between-individual level).

We are interested in the probability of someone giving a substantive response to a fictitious issue given the response latency; a measure of one's current motivation and opportunity or degree of mental elaboration (Mayerl, 2013), and the experimental treatment (implicit vs. explicit "don't know"). For the response latencies, we also include a squared term to allow for effects that are not monotonic.

Controls were included to address potential confounding of the between-person average response latency scores. While the demeaned item- and individual-varying response latencies are unconfounded by stable individual-specific

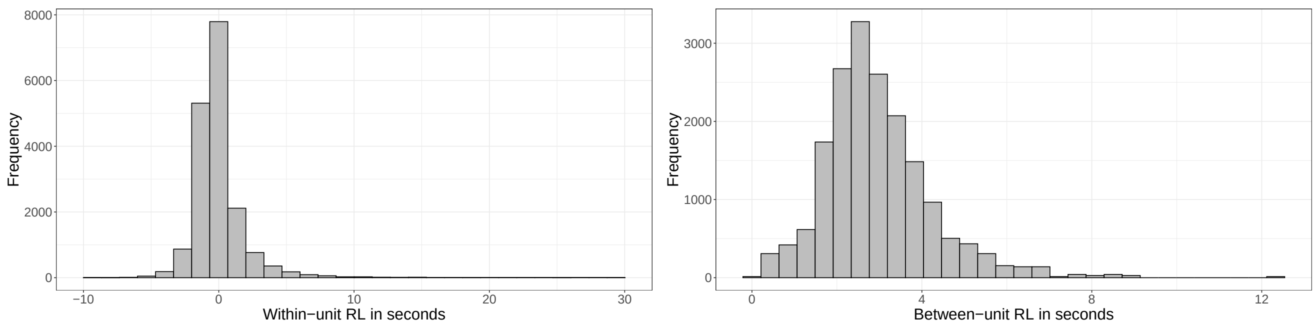

Figure 1

Response latency histograms

characteristics through the correlated random effects specification, the between-unit RLs are not exempt from confounding. The set of controls includes variables that could plausibly affect both the person means, i.e., the average speed with which the respondent answers, and the occurrence of pseudo-opinions. These controls are also plausibly "upstream" of both the person mean response latencies and the responses; we want to avoid controlling for mediators of the effect of average latencies on the responses. We chose year of birth, sex, German citizenship, education, political interest, academic degree (bachelor or higher), employment status, number of previous surveys completed within the last year, and political orientation. Note that the imputed meaning hypothesis specifically points to social desirability as a motivating factor. To test whether pseudo-opinions are associated with satisficing or optimizing, we thus do not want to control for social desirability, otherwise we would be stacking the odds artificially in favor of the mental coin-flip hypothesis.

4.2 Mitigating Pseudo-Opinions: An Explicit vs. implicit "don't know" experiment

It has been argued that nonsubstantive answers can be perceived as inadequate by respondents. An explicit "don't know" category could signal to the respondent that it is alright not to have an opinion, thereby lessening the pressure to answer substantively to obscure topics (Grant & Patterson, 1975). And there is some empirical evidence to support this idea. For example, Bishop et al. (1983) saw pseudo-opinions drop in an experiment from around 30% to only 10% when an explicit "don't know" category was offered. We therefore randomly selected half of the respondents to be shown an explicit "don't know" category throughout, while the other half was given no such option. To express their lack of an opinion, they could simply skip the question by clicking on "continue," though we did not highlight this fact to the respondents.

Further, we test whether the "don't know" option moderates the effect of response latencies on pseudo-opinions: signalling to satisficing respondents that it is okay not to have an opinion may increase the likelihood of them taking this exit option. We do so by introducing interaction effects between the experimental "don't know" variable and the response latencies (both within- and between-unit).

4.3 Modelling strategy

We specify three multilevel correlated effects probit models to assess the viability of specific assumptions. First, in Model 1, we estimate the effect of within-unit, as well as between-unit response latencies on pseudo-opinions. We allow for squared effects of both in this and subsequent models to account for effects that may not be monotonic. This and subsequent models also include the experimental "don't know" variable, as well as item dummies to account for unobserved factors unique to the specific items.

In Model 2, we introduce the set of control variables noted above. We do so in order to better interpret the effect of the between-person response latencies on pseudo-opinions, as the correlated random effects model does nothing to stop confounding of these effects.

Finally, in Model 3, we also introduce interaction terms between the response latencies and the experimental "don't know" variable. This allows for the possibility that the effect of response latencies on pseudo-opinions is moderated by an explicit "don't know" category. After all, it is plausible that the effects of optimizing/satisficing are moderated by telling the respondent it is alright not to have an opinion.

5 Results

We begin by presenting some descriptive results before interpreting the results of the multivariate regression models.

5.1 Descriptive Results

Before examining the results of the multivariate models, let us look descriptively at the distributions and overall prevalence of substantive responses to the fictitious organization questions; see Table 1.

From this, we see the extent of pseudo-opinions in our survey. Especially in the implicit "don't know" group, a large proportion of respondents gave pseudo-opinions (substantive responses, rows "negative" and "positive"). For the "Environmental Court," only about 15% of respondents in the implicit "don't know" group indicated they had no opinion by clicking the "continue" button. The other around 85% apparently had an opinion towards this fictitious organization.

We see from this table already that the experimental condition had a large impact on responses. When respondents were explicitly given the opportunity to admit they had no opinion, many indeed opted to use it. For the "Environmental Court," nonsubstantive responses increased from about 15% to about 36% without an explicit "don't know" option (rows "continue" and "don't know" taken together).

We also see a great deal of variation not just between experimental groups, but also between the items. This seems to suggests respondents do, in fact, interpret "cues" and attempt to answer accordingly. The first two items, the "Environmental Court" and "Coastal Aid Agency," ostensibly have to do with protecting the environment and perhaps dealing with refugee situations, respectively. And these two items have among the highest pseudo-opinion rates. Again, over 85% of respondents in the implicit "don't know" group had an opinion towards the "Environmental Court," and nearly 80% in that group had an opinion towards the "Coastal Aid Agency." The next two items, the "Prague Energy Transition Initiative" and the "German Nuclear Forum" would seem to have something to do with energy policy, although the former seems to suggest a pro-environmental group, whereas the latter seems to be concerned with advocating for nuclear energy in Germany. About 60% of respondents in the implicit "don't know" group gave opinions towards these organizations. The "Herbert-Schmaar-Foundation," on the other hand, provides essentially no cues as to the goals and interests of the organization. And this is reflected in the substantive responses: only about 8% of respondents in the explicit "don't know" group gave pseudo-opinions to this organization; in the implicit group it was only around 40%. And while the "World Space Agency" would obviously have something to do with "space," it is also not clear what the goals of the agency are. This is reflected in the proportions of pseudo-opinions, which are generally lower than the first four items, as well.

The differences between items speaks to the imputed meaning hypothesis, at least initially. However, we will come back to this point when we look at the results of the response latency analysis in the multivariate models.

5.2 Results of multivariate models

We now move on to the results of the multivariate regression models. Table 2 shows the average marginal effects (AMEs) for the three correlated random effects probit models. These allow for a better interpretation of the model results, as models 2 and 3 feature interaction effects and the models with binary outcomes are themselves nonlinear, meaning estimated probabilities hold only for specific combinations of values on the independent variables. The "raw" results can be found in Appendix A. We report the results with regard to the satisficing/optimizing research question (response latency), the explicit "don't know" experiment, and the item dummies. The coefficients of the control variables are not shown.

The effects of the response latencies, both within- and between-individual, are negative and significant, at least at the 5% level.[4] This holds for all model specifications, whether with or without controls, and regardless of the interaction terms specified. The slower the response (RL increases), the lower the likelihood of a pseudo-opinion. This actually speaks against the imputed meaning hypothesis, as it is the *faster, presumably automatic responses* that are associated with pseudo-opinions.

Note that in all models, we allow for nonlinear effects of response latencies (both within- and between-unit) by including squared terms. The average marginal effects give us a single coefficient for the within- and between-unit response latency effects. Thus, the AMEs give us a summary of the effect of the response latencies, averaged over all the observed values. And in fact, if we look at the "raw" results, shown in Appendix A, we see a significant positive effect for the squared between-unit response latencies in each model (though the effect is only significant at the 5% level in Models 2 and 3). This indicates that the probability of a pseudo-opinion actually increases again for slow average response latencies. While this indeed contradicts the impression given by the AMEs, these effects only hold for specific combinations of the independent variables, and we thus prefer to interpret the AMEs, which provide an effect averaged over all the observed combinations of predictor values.

Finally, we see the experiment of offering an explicit "don't know" category had the predicted effect: offering such an option reduces pseudo-opinions dramatically, with significant negative effects in each model specification. The AMEs for each model are displayed graphically in Figure 2.

Since our sample is a convenience sample drawn from a pool of online access panel participants, we must be careful in generalizing our findings. As a form of robustness check

[4]We report and interpret p-values here even though they are problematic with such convenience samples. We take use them as a way to discern an effect that is likely not simply due to chance, though we weary of making generalizing statements as the population and sampling procedure make these difficult. To our knowledge, the imputed meaning hypothesis has not yet been investigated in the context of online samples. As such, our results are preliminary and should be further tested in future studies.

Table 1

Responses to fictitious issues in percent

	Environmental Court		Coastal Aid Agency		Prague Energy Transition Initiative		German Nuclear Forum		Herbert-Schmaar Foundation		World Space Agency	
	expl.[a]	impl.[b]	expl.[a]	impl.[b]	expl.[a]	impl.[b]	expl.[a]	impl.[b]	expl.[a]	impl.[b]	expl.[a]	impl.[b]
negative	15	17	9	13	8	22	12	33	4	17	7	20
positive	48	69	32	65	14	43	12	33	3	23	14	39
continue	2	15	1	22	2	35	2	34	1	61	2	41
don't know	34	-	58	-	76	-	74	-	91	-	77	-
n	645	643	645	643	645	643	645	643	645	643	645	643

[a] explicit don't know [b] implicit don't know

Table 2

Average marginal effects, correlated random effects multilevel probit models

	Model 1		Model 2		Model 3	
	AME	SE	AME	SE	AME	SE
Experiment						
Explicit don't know	−0.35*	0.02	−0.39*	0.02	−0.39*	0.02
Response latency (in sec.)						
Within-unit RL	−0.01*	0.00	−0.01*	0.00	−0.01*	0.00
Between-unit RL	−0.03*	0.01	−0.04*	0.01	−0.04*	0.01
Ref.: Environmental Court						
Coastal Aid Agency	−0.16*	0.01	−0.16*	0.02	−0.16*	0.02
Prague Energy Transition Initiative	−0.31*	0.01	−0.33*	0.02	−0.33*	0.02
German Nuclear Forum	−0.30*	0.01	−0.33*	0.02	−0.33*	0.02
Herbert-Schmaar-Foundation	−0.52*	0.02	−0.57*	0.02	−0.56*	0.02
World Space Agency	−0.35*	0.01	−0.34*	0.02	−0.34*	0.02
Observations	7676		3800		3800	
Individuals	1287		637		637	
Model specification						
Controls	No		Yes		Yes	
Random slope RL	No		Yes		Yes	
Interactions, RL, exp. DK	No		No		Yes	

Controls not shown. Controls were: year of birth, sex, German citizen, school education dummies, political interest, academic (university degree), employment status, number of surveys completed in the past, and political orientation.
* $p < 0.05$.

for the generalizability of the results, we look at the effects of response latencies on pseudo-opinions for various subgroups of our sample. We ran Model 1 (with no controls, no interaction terms between response latencies and experimental treatment) separately based on sex, year of birth (categories), citizenship, school education, academic degree, employment status, and number of surveys completed in the past. This is shown in Table 3.

The results of this subgroup analysis show that the AMEs for both the within- and between-unit response latencies tend to be negative throughout. In the few examples where the AME is positive (those born between 1988 and 1974, non-German citizens, part-time employed, those with between 0–10 and 21–50 surveys completed in the past), the effects

Table 3

Subgroup analysis (Model 1 specification)

	Within-unit RL		Between-unit RL		
	AME	SE	AME	SE	*n*
Sex					
Female	0.006	0.004	−0.013	0.012	3894
Male	0.007	0.004	−0.048*	0.010	3798
Year of birth					
≥ 1987	−0.004	0.007	−0.063*	0.017	1926
1986–1974	0.005	0.005	−0.042*	0.017	1986
1973–1962	−0.003	0.005	−0.049*	0.017	1980
< 1962	−0.020*	0.006	−0.015	0.016	1800
Citizenship					
Not German	0.005	0.020	−0.011	0.046	198
German	−0.006*	0.003	−0.035*	0.008	7464
Education					
Abitur[b]	−0.008	0.005	−0.047*	0.014	3246
Fachhochschulreife[c]	−0.013	0.011	−0.051*	0.023	618
Realschule[d]	−0.007	0.005	−0.030*	0.013	2694
Hauptschule[e]	0.003	0.008	−0.035	0.019	966
Academic					
No	−0.006	0.003	−0.023*	0.009	5640
Yes	−0.008	0.006	−0.075*	0.015	1914
Employment					
Full-time	−0.002	0.004	−0.045*	0.012	3786
Part-time	0.005	0.007	−0.064*	0.020	1020
Mini-job	−0.026	0.019	−0.056	0.051	246
Unemployed/student	−0.008	0.014	−0.014	0.031	294
Student/similar	−0.010	0.013	−0.037	0.038	618
Retired	−0.020*	0.007	−0.002	0.017	1278
Housewife/-husband	−0.013	0.019	−0.064	0.033	378
Nr. surveys completed					
0-10	0.002	0.006	−0.072*	0.016	1194
11-20	−0.001	0.010	−0.049*	0.024	906
21-50	−0.015*	0.007	0.018	0.020	1386
51+	−0.016	0.009	−0.029	0.028	852

[a] Person-item observations in the long-format dataset.
[b] University entrance qualification [c] Technical/college entrance qualification.
[d] Higher vocational degree [e] Lower vocational degree
* $p < 0.05$

are very small and likely due to chance variation (note that none of these would typically be considered significant at the usual levels). From this, we suggest our finding that pseudo-opinions are associated with satisficing behavior may be transferable to other samples, though this must be investigated empirically.

6 Discussion and Conclusion

This study looked to investigate whether pseudo-opinions in online surveys are the result of satisficing or optimizing behavior. We operationalized response behavior in terms of passive response latencies and used fictitious issues to know when a respondent was giving a pseudo-opinion. We found the probability of observing a pseudo-opinion tends to de-

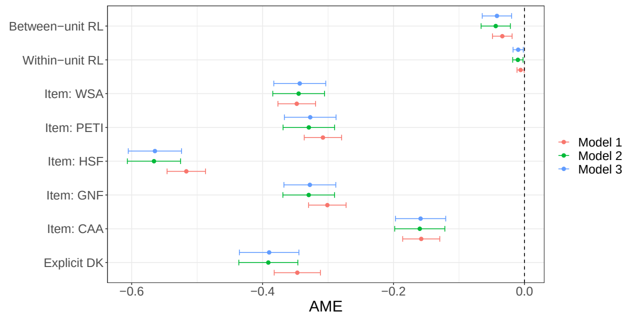

Figure 2

Average marginal effects, probability of pseudo-opinion with confidence intervals

crease the longer the respondent takes to answer. This applies at both the within-unit and between-unit level. At the same time, we observe sometimes large differences in terms of the proportions of pseudo-opinions depending on the question content. This suggests respondents perceive the item content and attempt to answer accordingly.

We take this as evidence that recontextualizes the imputed meaning hypothesis that has established itself as the dominant model of pseudo-opinions in the literature. The notion of a respondent laboring over their answer, carefully trying to suss out the meaning of an unfamiliar object, as described by Schuman and Presser (1980), does not hold up in our online survey. While respondents perceive the contents of the item, pseudo-opinions are the result of automatic satisficing. And this fits the definition of "weak satisficing" described by Vannette and Krosnick (2014), among others: respondents read, retrieve information, form a judgement and respond in a very cursory manner. If the item provides no clear cues for the respondent to latch on to, they tend to give up at a higher rate; take the "Herbert-Schmaar-Foundation," for example. This interpretation of an automatic imputed meaning is in line with previous research by Andersen and Mayerl (2019), for example, who showed that socially desirable responses (another response effect in which responses deviate from the "true" value) can occur automatically if the norms surrounding an attitude or behaviour are clear enough.

And we find this result rather intuitive. Why should survey respondents waste their time generating an ad hoc opinion towards something they are unfamiliar with? In survey modes with direct interaction between interviewer and interviewee, it is indeed possible that social desirability pressures may motivate respondents to want to appear informed, as the literature has suggested. But there is no interviewer present in online surveys, so who would the respondents be trying to impress? Cooperative respondents can conform with subjective expectations about the survey situation by simply giving any substantive answer; they do not need to deliberate in order to conform to survey protocol.

The implications of the findings are troubling for online surveys. The imputed meaning hypothesis is rather optimistic in that it paints a picture of respondents deliberating carefully about the supposed meaning of a question before giving their (presumably honest but misguided) opinion. If this were the case, then pseudo-opinions could arguably be seen as valid responses to more general objects. For example, a respondent may not know what "nanotechnology" is, but their pseudo-opinion may represent their honest feelings towards "innovative" or "new" technologies. In that case, we could expect a clear correlation between the observed ("new technologies") and latent ("nanotechnologies") attitudes/opinions. But our study suggests it is the satisficing respondents that tend to give pseudo-opinions. Respondents, when faced with an unfamiliar topic, tend to respond spontaneously and, especially for the extremely fast responses, it is hard to imagine they would be correlated with any underlying valid opinions. Depending on the satisficing heuristic the respondents tended to favor (random response, acquiescence, mainlining, etc.), results based on the data may be both unreliable and invalid.

This may simply have to do with the fact that deliberate, slow responses almost did not occur (after removing extreme outliers, the average raw (not demeaned) response was just 2.95 seconds, and 75% of responses were faster than 3.5 seconds, 90% were faster than 5.25 seconds). This means that the standard errors for the slow respondents may have just been too large, and that both the satisficing and optimizing hypotheses may be true, after all (see again results shown in Table 4, where the probability for pseudo-opinions first decreases and then starts to increase again for slow responses). But the fact remains that the *overwhelming majority of responses occurred very quickly with little obvious deliberation*. Even if some respondents do try to optimize and impute a meaning, they are few and far between and the problem of pseudo-opinions is one of satisficing respondents, first and foremost.

The results of this study are particularly relevant because of the growing popularity of online surveys. Online surveys are generally cheaper and faster to carry-out than most other modes (Bandilla, 2016). And some large-scale ongoing surveys, like the German ALLBUS, have taken the Coronavirus pandemic as an opportunity to move away from CAPI/CATI surveying towards online surveying.

Our study used a panel-provider to supply the respondents for our online survey. These respondents receive repeated invitations to participate in online surveys for a small monetary incentive. It is possible that such respondents may differ in important ways from, say, offline-recruited respondents. Then the results of the study would be applicable to online access panel recruited surveys, but perhaps not to offline-recruited online surveys with random sampling. But even if that is the case, the overwhelming majority of online surveys tend not to recruit offline, relying instead on such online access panel providers. Indeed, according to *Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.* (ADM), 78% of quantitative interviews in 2020 recruited participants in this way in Germany.[5] Further, we control for the number of surveys completed by the respondent in the multivariate models above. If the main difference between an offline-recruited participant and an online access panel participant is experience with surveys, then the results should be generalizable.

Pseudo-opinions are indeed a widespread problem in survey research. And we argue that much more work is needed to better understand the mechanisms behind them. The imputed meaning hypothesis paints a picture of survey respondents that may be overly optimistic. Our study shows that in online surveys—which continue to grow in popularity—pseudo-opinions are linked more with satisficing than optimizing. Giving the respondents an explicit exit option in the form of a "don't know" category can help mitigate pseudo-opinions in this mode, as well. But upwards of 65% of respondents in the explicit "don't know" treatment still give pseudo-opinions to the "Environmental Court," for example. So, what can be done about this sizable proportion of respondents? We encourage future research to replicate these results in other online surveys. Even when it comes to offline surveys, the evidence for the imputed meaning hypothesis is sparse and old at this point. New work should take a critical look at the conventional wisdom surrounding the mechanisms responsible for pseudo-opinions in various survey modes. And while we argue pseudo-opinions are *mostly* a problem of satisficing in this study, the results of the multivariate analyses show the effect of RLs may not be monotonic. If this is so, then we should investigate what differentiates the satisficing and optimizing pseudo-opinions. Finally, we assume the differing proportions of pseudo-opinions for the fictitious items has to do with the "cues" they provide to respondents (or lack thereof). We do not, however, have empirical evidence on the way in which the fictitious items were perceived by the respondents. This is something that should be investigated in the future. Respondents could be asked, perhaps after the debriefing, about what they thought the fictitious items were referring to. This may help us better understand the underlying thought processes of respondents when faced with unfamiliar topics.

We hope this study helps to invigorate the research on pseudo-opinions in surveys, and that it contributes to a better understanding of the problem and ways to mitigate it.

Funding

This study is part of the project Sensitive Questions and Social Desirability: Theory and Methods, funded by the German Research Foundation (grant WO 2242/1-1 to F.W. and MA 3267/7-1 to J.M.).

References

Andersen, H. K., & Mayerl, J. (2019). Responding to socially desirable and undesirable topics: Different types of response behaviour? *Methods, Data, Analyses (mda)*, *13*(1), 7–35. https://doi.org/10.12758/MDA.2018.06

Bandilla, W. (2016). Web surveys: GESIS survey guidelines [GESIS—Leibnitz Institute for the Social Sciences.].

Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, *3*(1), 133–153.

Bishop, G. F., & Jabbari, B. J. (2001). *The internet as a public opinion laboratory: Experiments with survey questions* [Paper presented on the fifty-sixth annual AAPOR conference].

Bishop, G. F., Oldendick, R. W., & Tuchfarber, A. J. (1983). Effects of filter questions in public opinion surveys. *Public Opinion Quarterly*, *47*(4), 528–546. https://doi.org/10.1086/268810

Bishop, G. F., Oldendick, R. W., Tuchfarber, A. J., & Bennett, S. E. (1980). Pseudo-opinions on public affairs. *Public Opinion Quarterly*, *44*(2), 198–209.

Converse, P. (1964). The nature of belief systems in mass publics. In D. Apter (Ed.), *Ideology and discontent* (pp. 206–261). Free Press.

Converse, P. (1970). Attitudes and nonattitudes: Continuation of a dialogue. In *The quantitative analysis of social problems* (pp. 168–189). Addison-Wesley.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology*, *23*, 75–109. https://doi.org/10.1016/S0065-2601(08)60318-4

Grant, L. V., & Patterson, J. W. (1975). Non-attitudes: The measurement problem and its consequences. *Political Methodology*, *2*(4), 455–481.

Joinson, A. (1999). Social desirability, anonymity, and internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 433–438. https://doi.org/10.3758/bf03200723

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236.

Le Payne, S. B. (1951). *The art of asking questions*. Princeton University Press.

[5] See https://www.adm-ev.de/die-branche/mafo-zahlen/, table "Quantitative Interviews nach Befragungsart".

Leiner, D. J. (2019). Too fast, too straight, too weird: Non-reactive indicators for meaningless data in internet surveys. *Survey Research Methods*, *13*(3), 229–248. https://doi.org/10.18148/SRM/2019.V13I3.7403

Lenzner, T. (2011). *A pscholinguistic look at survey question design and response quality* [PhD thesis University of Mannheim].

Mayerl, J. (2013). Response latency measurement in surveys: Detecting strong attitudes and response effects. *Survey Methods: Insights from the Field*. https://surveyinsights.org/?p=1063

Mayerl, J., & Urban, D. (2008). *Antwortreaktionszeiten in Survey-Analysen. Messung, Auswertung und Anwendungen*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91147-2

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, *46*(1), 69–85.

Roberts, C. (2016). Response styles in surveys: Understanding their causes and mitigating their impact on data quality. In C. Wolf, D. Joye, T. W. Smith, & Y.-c. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 579–596). Sage.

Roberts, C., Gilbert, E., Allum, N., & Eisner, L. (2019). Research synthesis. *Public Opinion Quarterly*, *83*(3), 598–626. https://doi.org/10.1093/poq/nfz035

Schuman, H., & Presser, S. (1980). Public opinion and public ignorance: The fine line between attitudes and nonattitudes. *American Journal of Sociology*, *85*(5), 1214–1225. https://doi.org/10.1086/227131

Schunck, R., & Perales, F. (2017). Within- and between-cluster effects in generalized linear mixed models: A discussion of approaches and the xthybrid command. *Stata Journal*, *17*(1), 89–115.

Shin, E., Johnson, T. P., & Rao, K. (2011). Survey mode effects on data quality: Comparison of web and mail modes in a U.S. national panel survey. *Social Science Computer Review*, *30*(2), 212–228. https://doi.org/10.1177/0894439311404508

Silber, H., Lischewski, J., & Leibold, J. (2013). Comparing different types of web surveys: Examining drop-outs, non-response and social desirability. *Advances in Methodology and Statistics*, *10*(2), 121–143.

Sturgis, P., & Smith, P. (2010). Fictitious issues revisited: Political interest, knowledge and the generation of nonattitudes. *Political Studies*, *58*(1), 66–84. https://doi.org/10.1111/j.1467-9248.2008.00773.x

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.

Vannette, D. L., & Krosnick, J. A. (2014). A comparison of survey satisficing and mindlessness. In C. T. Ngnoumen & E. J. Langer (Eds.), *The Wiley Blackwell handbook of mindfulness* (pp. 312–327). Wiley.

Wolter, F., & Junkermann, J. (2019). Antwortvalidität in Survey-Interviews: Meinungsäußerungen zu fiktiven Dingen. In N. Menold & T. Wolbring (Eds.), *Qualitätssicherung sozialwissenschaftlicher Erhebungsinstrumente* (pp. 339–368). Springer.

Wooldridge, J. (2002). *Econometric analysis of cross sectional and panel data*. MIT Press.

Appendix
Results of Multilevel Probit Regression (Correlated Random Effects)

The following table shows the “raw” results of the multilevel probit regression models with correlated random effects.

Table A1

Multilevel probit regression models

	Model 1		Model 2		Model 3	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Experiment						
Explicit DK	−1.44*	0.08	−1.68*	0.11	−2.46*	0.56
Response latency (in sec.)						
Within-unit RL	−0.03*	0.01	−0.05*	0.02	−0.07*	0.02
Between-unit RL	−0.42*	0.09	−0.49*	0.16	−0.70*	0.22
Ref.: Environmental Court						
Coastal Aid Agency	−0.70*	0.60*	−0.78*	0.01	−0.78*	0.10*
Prague Energy Transition Initiative	−1.34*	0.07	−1.54*	0.10	−1.53*	0.10
German Nuclear Forum	−1.31*	0.07	−1.54*	0.10	−1.53*	0.10
Herbert-Schmaar-Foundation	−2.36*	0.07	−2.72*	0.11	−2.71*	0.11
World Space Agency	−1.52*	0.07	−1.61*	0.10	−1.60*	0.10
Squared RL						
Within-unit RL sq.	0.00	0.00	0.00	0.00	0.00	0.00
Between-unit RL sq.	0.04*	0.01	0.04*	0.02	0.07*	0.03
Interactions						
Explicit don’t know × Within-unit RL	-	-	-	-	0.03	0.04
Explicit don’t know × Between-unit RL	-	-	-	-	0.41	0.30
Explicit don’t know × Within-unit RL sq.	-	-	-	-	−0.00	0.00
Explicit don’t know × Between-unit RL sq.	-	-	-	-	−0.04	0.04
Intercept	2.64*	0.18	2.39*	0.52	2.81*	0.59
Controls	No		Yes		Yes	
Random slope RL	No		Yes		Yes	
AIC	7244.7		3397.7		3401.2	
BIC	7238.0		3597.5		3625.9	
Log likelihood	−3610.3		−1666.9		−1664.6	
Nr. observations	7676		3800		3800	
Nr. groups (individuals)	1287		637		637	
Var.: Intercept	1.534		1.21		1.21	
Var.: Slope RL	-		0.00		0.00	

Controls not shown.

* $p < 0.05$