



UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI
SCIENZE SOCIALI E POLITICHE

spstTREND

WHAT TO DO WITH INSTRUCTIONAL MANIPULATION CHECKS IN DATA ANALYSIS?

Evidence from the ResPOnsE COVID-19 survey

Riccardo Ladini, Nicola Maggini

ESRA conference - Milan, July 18th, 2023

Session "Inattentiveness and satisficing in self-administered survey"

STARTING POINT

“if even the oldest research methods still have limitations that are often overlooked and underappreciated, then can we imagine how is it with newer, innovative methods” (Nichols and Edlund, 2020, 625)

AN ISSUE IN ONLINE SURVEYS: respondents' inattentiveness

- A possible solution to detect it: the use of attention checks
- How to deal with attention checks in substantial analysis?
Still an open question
- Need for further reflections



ATTENTION CHECKS IN SURVEY RESEARCH

Various kinds of attention checks:

- Subjective and factual manipulation checks (after experiments)
 - Bogus Items
 - Instruction Respondent Items
 - Instructional Manipulation Checks
-
- **Tools for distinguishing between attentive («workers») and inattentive respondents («shirkers»),** Berinsky et al. 2014)



THE INSTRUCTIONAL MANIPULATION CHECKS

We are now interested to know whether people take the time they need to follow carefully instructions in interviews.

To show that you've read this much, please **select only the options “Local newspaper websites” and “I do never consult websites” as your two answers**, no matter of the websites you actually visit

- Repubblica
- Il Giornale
- La Stampa
- Corriere
- Dagsopia
- Siti di agenzie
- Huffington post
- Il Fatto Quotidiano
- Libero quotidiano
- Local newspapers websites**
- Other
- I do never consult websites**



THE WORKING OF IMCs

Respondents not passing attention checks:

- higher tendency toward **straightlining**, **speeding**, **item nonresponse**
- **lower answer consistency**

(Alvarez et al. 2019; Anduiza and Galais 2016; Gummer et al. 2021; Ladini 2022; Mancosu et al. 2019)



IMCs are **effective tools in distinguishing between more attentive and less attentive respondents**

BUT IMCs' outcome is not independent from individual characteristics (e.g. education, interest in politics and survey topic...)



WHAT TO DO WITH IMCs' IN DATA ANALYSIS?

What to do in the analysis with people not passing IMCs?

- Filtering them out (Oppenheimer et al. 2009)
- Not delete them as those respondents have different socio-demographics and attitudinal characteristics, **but account for them in the analysis** (Juan Rubio and Revilla 2021; Nichols and Edlund 2020). **How?**
 - Weighting and other ex-post adjustment: issues of sample representativeness persist
 - Be transparent: run **separate analyses** (all sample, only sample of respondents passing IMCs). If substantial differences do emerge?
- A shared procedure about what to do is still lacking!



THE IMPACT ON SUBSTANTIAL ANALYSIS

With standard survey questions and research designs?

- Univariate analysis: differences in distributions of several variables by IMCs outcome (Juan Rubio and Revilla 2021: 7 out of 18 variables analysed)
- **Multivariate analysis:**
 - No substantial difference in regression results when considering the whole sample or only respondents passing IMCs (Anduiza and Galais 2016) and IRIs (Gummer et al. 2021)
 - Substantial differences in causal effect by IMCs outcome depend on the variables analyzed (Juan Rubio and Revilla 2021 on support for climate change policies)

Further empirical evidence is needed (Juan Rubio and Revilla 2021)



OUR PROPOSAL: REPLICATION ANALYSIS

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 168–176
<https://doi.org/10.17645/pag.v10i4.5766>

Article

When Believing in Divine Immanence Explains Vaccine Hesitancy: A Matter of Conspiracy Beliefs?

Riccardo Ladini * and Cristiano Vezzoni

H1: Believing in divine immanence is positively associated with vaccine hesitancy



OUR PROPOSAL: REPLICATION ANALYSIS

Data

4th wave of the **ResPOnsE COVID-19** survey project (Vezzoni et al. 2020, 2022)

Sample:

- drawn from an **opt-in panel** of an **Italian** survey research institute (Swg S.p.a.)
- Reproduces population distribution by sex, age class and geographical area

Period:

November-December 2021, n = 3005

- One IMC received by a random subsample of respondents (n = 1,767), randomly placed at the beginning, middle, and end of the questionnaire
- % respondents passing the IMC = **56%**



OUR PROPOSAL: REPLICATION ANALYSIS

Methods and Measures

DV: Propensity to get vaccinated in a hypothetical future pandemic (0: not at all; 10: total)

IV: Belief in divine immanence

Additive index (0-10) of the following 3 items (0-10 agreement scale):

1. Miraculous healings do exist
2. God intervenes directly in our lives
3. Prayer can heal physical illness

Controlling for **gender**, **age class** (three categories), level of **education** (three categories) and **geographical area** (4 categories)

Linear regression models



UNIVARIATE ANALYSIS

- IV and DV Means by IMC'S Outcome (n=1,337)

| | IMC correct (n=750) | IMC failed (n=587) | P-value (t-test) |
|---|------------------------|-----------------------|------------------|
| DV: Propensity to get vaccinated (0-10) | 8.5 | 8.3 | 0.07* |
| IV: Belief in divine immanence (0-10) | 3.9 | 4.5 | 0.001*** |



UNIVARIATE ANALYSIS

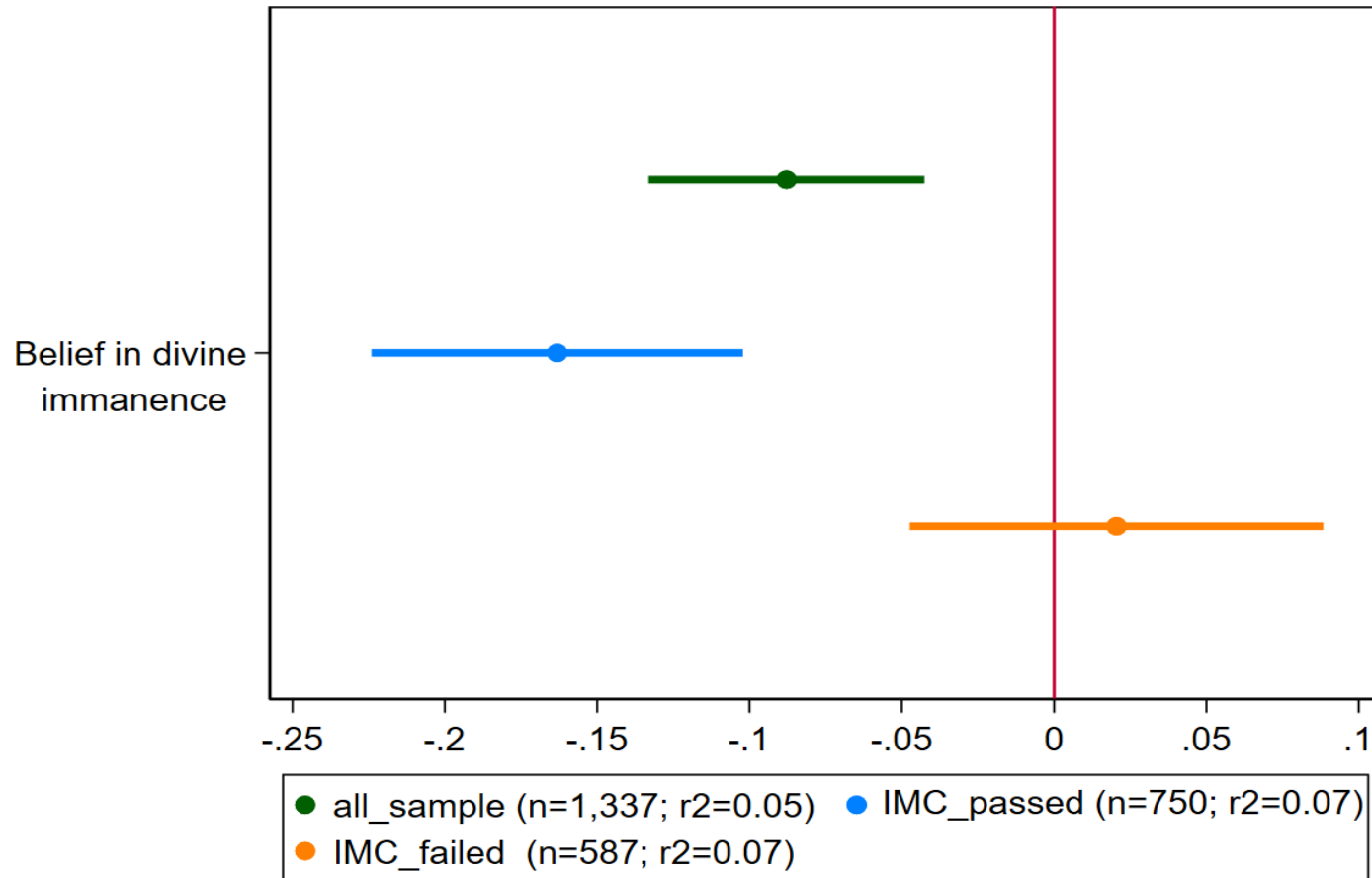
- Control variables' distribution by IMC'S Outcome

| | | IMC correct (n=750) | IMC failed (n=587) | P-value (chi2) |
|------------------|-------------|------------------------|-----------------------|----------------|
| Sex | Male | 48.7 | 55.4 | 0.015*** |
| | Female | 51.3 | 44.6 | |
| Education | Low | 7.6 | 11.8 | 0.000*** |
| | Medium | 63.6 | 69.5 | |
| | High | 28.8 | 18.7 | |
| Age | 18-34 | 16.9 | 16.0 | 0.123 |
| | 35-54 | 34.1 | 39.5 | |
| | 55 and over | 48.9 | 44.5 | |
| Geo-Area | North-West | 30.4 | 27.3 | 0.065* |
| | North-East | 21.1 | 17.4 | |
| | Centre | 21.0 | 20.0 | |
| | South | 28.5 | 34.4 | |



MULTIVARIATE ANALYSIS

Linear regression coefficients (DV: propensity to get vaccinated, 0-10)

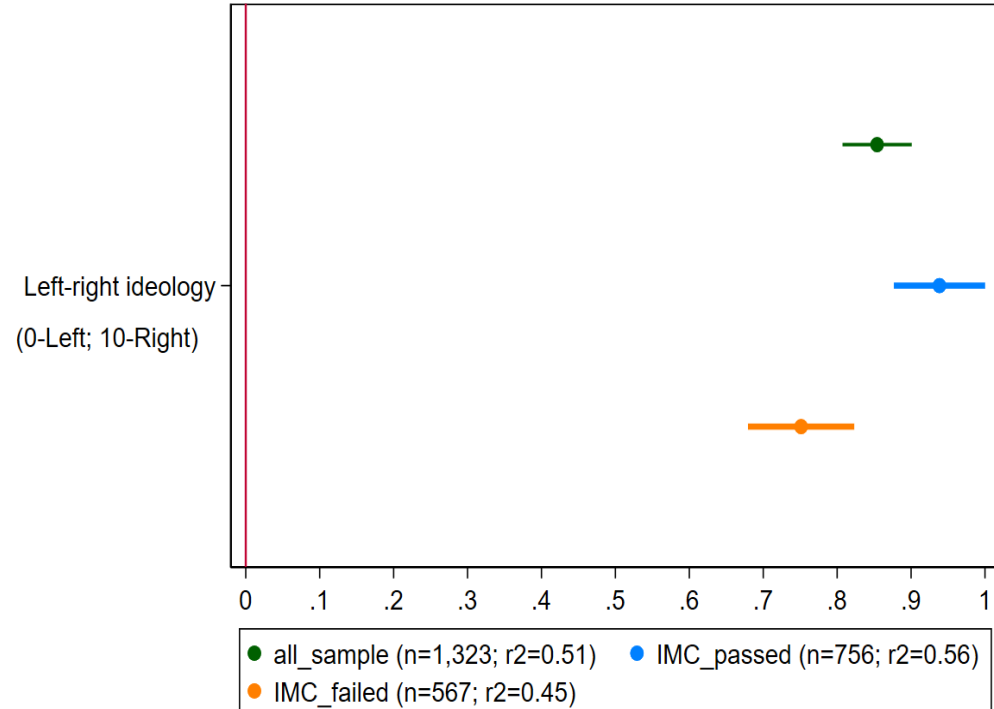
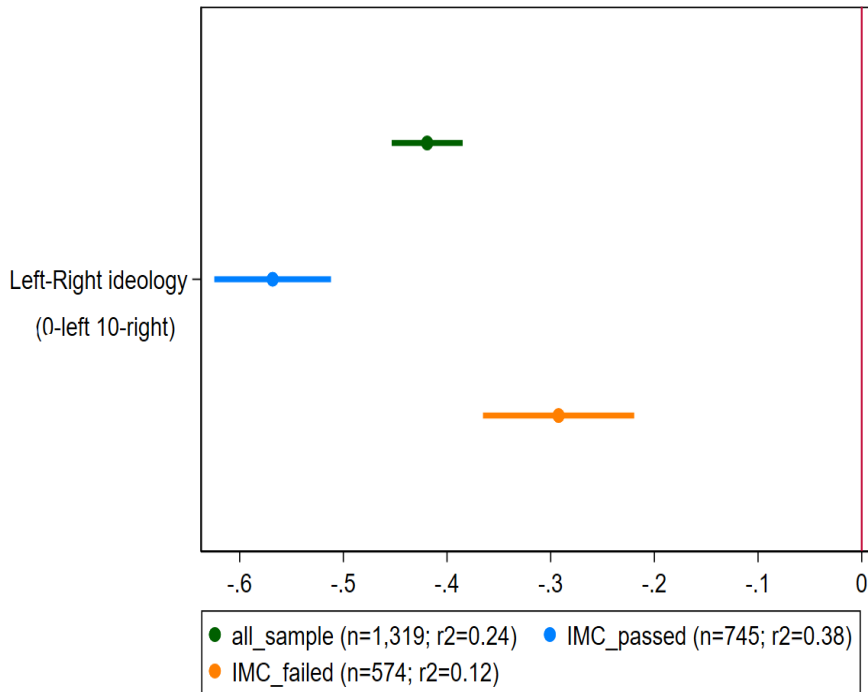


MULTIVARIATE ANALYSIS: Left-right and immigration attitudes (and Propensities to Vote)

Linear regression coefficients (controlling for socio-demographics)

DV: attitudes toward immigration
0: anti-10:pro

DV: Propensity to vote for Fratelli d'Italia (0-10)



CONCLUSIONS

- Depending on the IMCs' outcomes, even very simple and consolidated relationships between variables lead to different results
- The issue of answers' quality in online surveys is crucial and largely affect substantive analysis, and data users should be aware of it
- IMCs are likely to be too stringent in identifying inattentive respondents, even in their simplest version
- Are IMCs useful in standard research design? On their own, probably not
- Are less discriminating attention checks (e.g. Instruction Response Item, Gummer et al. 2021, despite the issue of false positives) more useful? **Maybe**



CONCLUSIONS

- But still, what to do with IMCs in data analysis?
 - Not accounting for them leads to more conservative estimates of relationships between attitudes, but is it the same with other variables?
 - Evaluating IMCs together with other paradata on survey attentiveness (Alvarez et al. 2019; Juan Rubio and Revilla 2021), but it is not a common practice among data users



Comments and questions
are welcome



Thank you!

<https://spstrend.it>

 Sps Trend

 @SpStrend



UNIVARIATE ANALYSIS

- IV and DV Means by IMC'S Outcome

| | IMC correct | IMC failed | P-value (t-test) |
|--|-------------|------------|------------------|
| DV: Propensity to get vaccinated (0-10) | 8.6 | 8.1 | 0.002*** |
| IV: Belief in divine immanence (0-10) | 4.1 | 4.6 | 0.003*** |



UNIVARIATE ANALYSIS

- Control variables' distribution by IMC'S Outcome

| | | IMC correct (n=987) | IMC failed (n=780) | P-value (chi2) |
|------------------|-------------|------------------------|-----------------------|----------------|
| Sex | Male | 46.2 | 53.6 | 0.002*** |
| | Female | 53.8 | 46.4 | |
| Education | Low | 8.3 | 12.7 | 0.000*** |
| | Medium | 65.6 | 67.3 | |
| | High | 26.1 | 20.0 | |
| Age | 18-34 | 16.2 | 16.3 | 0.072* |
| | 35-54 | 34.4 | 39.4 | |
| | 55 and over | 49.5 | 44.5 | |
| Geo-Area | North-West | 30.6 | 26.8 | 0.076* |
| | North-East | 19.6 | 19.9 | |
| | Centre | 20.6 | 18.7 | |
| | South | 29.3 | 34.6 | |

