

1 **Towards Measuring Sensitivity of Psychometrics in Crowdsourcing Tasks:**
2 **Engaging with Fact-checked Content Online**
3
4

5
6 This study attempts to explore the relationships between psychometric test outcomes and how participants rate the truthfulness of
7 online news articles. We seek to be able to help in the identification of misinformation. We attempt to explore the relationship between
8 political orientation and several other psychometrics. We then ask the user to rate social media article as well as fact-checked article
9 before finding the relationship between the variables and the difference in rating of each articles. We faced an issue in aggregating the
10 result of each participant as well as computing the distance in ordered scale (magnitude given values in limited range).
11
12
13
14
15
16
17

18 **1 INTRODUCTION**
19

20 As the popularity of social media keeps on rising and the usage of recommender algorithms becomes more rampant,
21 we can clearly see the echo chamber effect. In this research, we want to measure the sensitivity of self-reported
22 psychometric tools in relation to the participant’s bias towards misinformation. By understanding the relationship
23 between psychometric test outcomes and their bias towards misinformation, it would then be possible to make it
24 easy to identify misinformation. We are interested in looking if political orientation plays a role in the bias towards
25 misinformation as well as if other self-reported psychometric tools that are meant to detect if a person has a high
26 likelihood of falling into misinformation will have any relationship with how they rate online content when used by
27 non-experts.
28
29

30 The project aims to contribute to advancing knowledge on how to effectively present verified content on screen-based
31 interfaces. We anticipate that our outcomes will benefit fact-checking professionals and user experience (UX) designers
32 working in debunking dis/misinformation. The results may also benefit the community by making personalised verified
33 content created by fact-checking organisations such as RMIT ABC Fact Check.
34
35
36
37

38 **2 STUDY**
39

40 In our study, each participant answered a set of demographic questions, including political orientation and a series
41 of self-reported psychometric tools, followed by rating the truthfulness of five social media articles. For each article,
42 we first asked participants to rate the truthfulness of an article depicted as a social media post. Then, we showed the
43 fact-checked version of the article and asked them to re-rate the truthfulness. The study follows a between-subject user
44 study design, where each participant rates the same five articles.

45 Participants were recruited via the crowdsourcing platform Mechanical Turk and redirected to the experiment in the
46 Qualtrics platform. For quality control purposes, we only allowed crowd workers with an approval rate greater than
47 95% and who had completed more than 1000 tasks in the platform. The pilot data was collected in January 2023. For the
48 preliminary study, we recruited 20 participants and discarded the data of one participant due to an invalid Mechanical
49 Turk ID. Participants received US\$5 for completing the study.
50
51
52

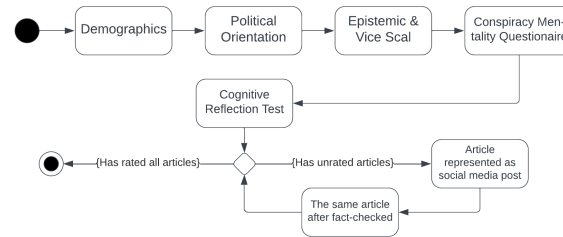


Fig. 1. **Experiment flow.** The flowchart shows how the experiment components interact with each other.

2.1 Psychometric Instruments

We selected four psychometric tests based on their relevance to research on misinformation.

- *Political Orientation*: 11-point Likert scale ranging from 0 to 10. 0 For conservative and 10 for liberal.
- *Epistemic Vice Scale (EVS)*: The result from the epistemic vice scale is two measures. The indifference and rigidity. Indifference measures the lack of motivation to find the truth, while rigidity measures the person's insensitive to evidence [3].
- *Conspiracy Mentality Questionnaire (CMQ)* [1]: Since conspiracy theories could be categorised as misinformation, the result of this set of questions could be used as an indicator that people with a high conspiracy mentality is biased towards their conspiracy believes. When given the truth as evidence, they will instead dismiss the truth and prefer the false news. The study shows that conspiracy theories still thrive in the 21st century and that if one person believes in one conspiracy theory, they will be most likely to believe in other conspiracy theories[1]. From the article, we hypothesised that those with higher CMQ score will most likely to believe in misinformation/conspiracy article.
- *Cognitive Reflection Test (CRT)* [2, 5]: The cognitive reflection test attempts to measure the decision-making characteristics in time & risk preferences of a person. Because identifying misinformation requires time, we think that this test would be a good indicator if a person would jump straight to a conclusion or spend some time figuring if something is up. We hypothesised that those that failed to override their incorrect gut response would be more susceptible to misinformation. We can also see how long it takes for the participant to complete the CRT.

Initially, we are not sure how many psychometric tools we can include in the experiment. However, when we compare a subset of the psychometric test against all of the psychometric tools, we found no significant impact on the time it takes for the participant to complete.

2.2 Social Media Posts

To quantify the participant's bias towards an article posted on social media, five news articles are used and a mock social media post is created. The article consists of two factual, two misinformation, and one in-between, which are rated and fact-checked by RMIT ABC Fact Check.

Political preferences and expertise/familiarity with the topic can potentially influence how users perceive the truthfulness of fact-checked content. To mitigate the impact of such factors, we have avoided any political content in the news articles and have included multiple topics (e.g., climate change, health, technology) in each experimental condition.

2.3 Multiple Dimensions of Veracity

To capture how participants assess the truthfulness of articles, we use the seven dimensions of veracity/truthfulness proposed by Soprano et al. [4]: Correctness, Neutrality, Comprehensiveness, Precision, Completeness, Trustworthiness, and Informativeness. For each dimension, participants had to indicate the veracity using a 7-point Likert scale. The overall rating for each article is computed by averaging the seven dimensions. The same dimensions and scale are used for both social media posts and fact-checked posts.

3 PRELIMINARY RESULT ANALYSIS & DISCUSSION

Table 1. Preliminary results for 7 generalized linear model (one for each veracity dimension). The significance of the veracity dimension is shown along with the relationship between the veracity dimension and psychometric.

Veracity	R-square	Political Orientation		Epistemic Vice Indiff.		Epistemic Vice Rigid.		CMQ Avg.	
		Coef.	P> z	Coef.	P> z	Coef.	P> z	Coef.	P> z
Correctness	0.068 07	1.2712	0.207	-2.1885	0.149	-0.3561	0.843	0.0489	0.977
Neutrality	0.022 79	0.5493	0.513	0.1914	0.880	-0.3726	0.803	1.3595	0.338
Comprehensiveness	0.060 37	-0.4546	0.639	-2.0090	0.168	0.7590	0.660	2.0336	0.215
Precision	0.1027	0.8457	0.355	-2.7804*	0.043	-0.8732	0.592	1.0806	0.484
Completeness	0.1380	1.2483	0.156	-1.7757	0.180	0.0452	0.977	2.6767	0.072
Trustworthiness	0.1927	0.4532	0.635	-3.7739*	0.009	-1.1631	0.494	3.4398*	0.033
Informativeness	0.1601	0.2729	0.770	-4.1051*	0.004	-0.8849	0.595	1.6797	0.287

The primary research question we tried to answer is “What is the relationship between psychometric and non-expert accuracy to assess veracity”. To get this preliminary result, we use the Δ of each dimension of veracity for article as the target value by using $\Delta = R_1 - R$ (where R_1 is the score of the veracity dimension of the fact-checked article and R is veracity dimension of the social media article). Score of each dimension is negated if the article is deemed to be false by RMIT fact-check. We then use the psychometric result as the variable and the Δ as the target value for each participant.

Our pilot study provided initial evidence of relationships between psychometric test outcomes and how non-experts assess veracity. However, we are finding a limitation in aggregating the result of each participant. Considering that in total there are seven models for each articles, there is around 28 models to analyse. Since the 28 models would not be able to tell the overall relationship, the data of each levels of veracity from each article must be aggregated for each participant. The problem is how to compute the distance in ordered scale (magnitude given values in limited range) as well as to generate an intermediate variable in order to correlate with other variables. We are currently omitting articles that are deemed as neutral by RMIT fact-check since we currently unsure how to quantify the participants result with regards to neutrality. For example, when a negative article is given a score of 7 and the fact-check article is given a rating of 7, the delta is 0. When a positive article is given a score of 7 and the fact-check article is given a rating of 7, the delta is also 0.

The 0 magnitude does not tell the entire story because the fact checked article has failed to convince that their view is wrong while 0 magnitude for positive article means that the participant has known that the social media article is correct.

However, the magnitude for the change in negative article should be larger compared to the positive article since the fact-checked article has convinced someone that strongly believe in the negative article that their view is wrong.

Another challenge we face in the data analysis phase is how to aggregate the results per worker without penalising the good workers. A good worker can be defined as those who provides a change in the level of veracity between the initial social media post and the fact-checked article. First is to understand if there is a change in veracity and if so, is the change positive or negative. For example, a participant who rate a negative post with a score of 7 and the fact-checked article with 7 as well will have a delta of 0 while another participant rated a positive social with a score of 7 and the fact-checked article with 7 will also have a delta of 0, which would mean they see an equal value in both negative and positive article while in fact, the participant who rated 7 in the negative article could have been convinced that their view is wrong while the participant who rated 7 in the positive article proves that the social media article greatly summarised the fact-checked article.

4 FUTURE WORK

Based on the preliminary results we obtained, for the full experiment, 1000 people will be recruited in order to include a representative sample. With regards to the data analysis, we need to find a way to combine and normalise the delta into one target value so that finding relationship between the psychometric and the bias towards online content could be easily explained. By using the current approach, our model is too complicated and determining which psychometric gives you the most signal is not easy. As we found out, there is no significant time impact between those who took the PDI and those who did not take PDI. In the next phase of this project we can add the PDI into the psychometric test set. We hope by adding PDI into the psychometric set, it would provide more insights into the data. The political orientation slider can be replaced with a clearer and more direct multiple-choices (i.e., Democratic, Republican, Liberal, Labor, Independent and Other).

This way, participant has broader options to enter their political orientation and the researcher will have a less ambiguous result. The current approach of using a slider for conservative or liberal does not capture the broader political spectrum.

REFERENCES

- [1] Martin Bruder, Peter Haffke, Nick Neave, Nina Nouripanah, and Roland Imhoff. 2013. Measuring Individual Differences in Generic Beliefs in Conspiracy Theories Across Cultures: Conspiracy Mentality Questionnaire. *Frontiers in Psychology* 4 (2013). <https://doi.org/10.3389/fpsyg.2013.00225>
- [2] Shane Frederick. 2005. Cognitive reflection and decision making. *Journal of Economic perspectives* 19, 4 (2005), 25–42.
- [3] Marco Meyer, Mark Alfano, and Boudewijn De Bruin. 2020. Epistemic vice predicts acceptance of Covid-19 misinformation. *Available at SSRN 3644356* (2020). <https://doi.org/10.2139/ssrn.3644356>
- [4] Michael Soprano, Kevin Roitero, David La Barbera, Davide Ceolin, Damiano Spina, Stefano Mizzaro, and Gianluca Demartini. 2021. The Many Dimensions of Truthfulness: Crowdsourcing Misinformation Assessments on a Multidimensional Scale. *Information Processing & Management* 58, 6 (2021), 102710. <https://doi.org/10.1016/j.ipm.2021.102710>
- [5] Maggie E Toplak, Richard F West, and Keith E Stanovich. 2011. The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition* 39, 7 (2011), 1275–1289.