Title

Cue to Trust? Investigating the Impact of Political Advertising Transparency Disclaimers on Citizen Trust Evaluations

Permalink

https://escholarship.org/uc/item/8fs0x868

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Gordon, Hazel Clare
Stafford, Tom
Dommett, Katharine

Publication Date

2024

Peer reviewed

A Cue to Trust? Investigating the Impact of Political Advertising Transparency Disclaimers on the Trust Evaluations of UK Citizens

Hazel C. Gordon (hcgordon1@sheffield.ac.uk)
Department of Psychology, University of Sheffield,
219 Portobello Street, S1 4DP

Tom Stafford (t.stafford@sheffield.ac.uk)
Department of Psychology, University of Sheffield,
219 Portobello Street, S1 4DP

Katharine Dommett, (k.dommett@sheffield.ac.uk)
Department of Politics and International Relations,
University of Sheffield, 2 Whitham Road, S10 2SJ

Abstract

Citizens in a democracy must navigate an increasingly dense information landscape. Regulation can aid this navigation by mandating disclosures of the source and nature of political campaign material. In many countries, legislators are increasing transparency requirements for online advertising in particular. The current paper looks at how and if citizens use such disclaimers to infer the intent of political advertisers during the process of a trust evaluation. This paper describes a survey experiment that specifically investigates evaluations of unknown campaigners, theorising such conditions will maximise any effect disclaimers have on trust. Testing both sponsorship and micro-targeting disclaimers, no support is found for the theoretical claim that viewing a disclaimer can increase how trustworthy a political advertiser is perceived to be. There is preliminary support that, for some individuals, viewing a disclaimer increases scepticism.

Keywords: Political Advertising; Transparency Disclaimers; Trust; Evaluation; Perceived Trustworthiness

Introduction

Trust is understood as a relational phenomenon between a trustor and a trustee (Simpson & Vieth, 2022), the nature of which is characterised by risk for the trustor and freedom for the trustee (Lauharatanahirun & Aimone, 2022; Siegrist, Earle & Gutscher, 2010). Accordingly, in a representative democracy, citizens often vote for candidates whom they trust to represent their interests despite being uncertain as to if they will act as hoped. Communicating trustworthiness is therefore a key challenge faced by political candidates.

Political advertising is a primary means by which candidates - and campaign groups with vested interest in the political goals of a candidate - can communicate with citizens during election periods (Birch, 2011). Essential to a democracy is ensuring citizens are aware of their political options and engage with a range of perspectives during this time (Birch, 2011; Dahl, 1989). Online political campaigning enables more opportunity for citizens to encounter information and thus become informed (Norris, 2003;

Zuiderveen Borgesius, Möller, Kruikemeier, Ó Fathaigh, Irion, Dobber, Bodo & de Vreese, 2018). Particularly for less well-known campaigners, who do not afford the reputational and financial advantages of large political parties, online political advertising is an efficient and cost-effective means by which they can spread their reach (Zuiderveen Borgesius et al., 2018). Advertising as a form of political communication, however, can be perceived negatively by citizens (Dermody & Scullion, 2005; Kates, 1998; Thomson & Brandenburg, 2019), particularly when harnessing micro-targeting practices, the act of using stored data about online users to curate tailored messages only visible to them (Dobber, Trilling, Helberger & de Vreese, 2018). Observation of these negative perceptions have led many to theorise that increasing advertising transparency will reduce the perceived trustworthiness of advertisers who disclose such information (Binder, Stubenvoll, Hirsch & Matthes, 2022; Binford, Wojdynski, Lee, Sun & Briscoe, 2021; Dobber, Kruikemeier, Helberger, & Goodman, 2023; Jost, Kruschinski, Sülflow, Haßler & Maurer, 2022; Krukiemeier, Sezgin & Boerman, 2016). In contrast, Beckert, Koch, Viererbl and Schulz-Knappe (2021) and Wood (2022) argue that the act of being transparent in and of itself can signal trustworthiness.

This paper describes a survey experiment designed to add to this research field, drawing from theories of cognitive psychology to ask under what conditions might the effect of a disclaimer be maximised. It is theorised that the effect of disclaimers will be enhanced under conditions where the advertiser is unknown. Identifying and testing such conditions is expected to shed light on which of the competing theoretical mechanisms are empirically supported, which will increase understandings of how informational cues are used by citizens to evaluate the trustworthiness of political campaigners.

Perceived Advertiser Trustworthiness

A trust evaluation is understood to occur as a decision process (Engelmann & Hare, 2018; PytlikZillig & Kimbrough, 2016),

through which both pre-existing knowledge, beliefs and affective perceptions are integrated with the information available for processing in the external environment to form a belief about the trustworthiness of an object in a particular context (Farolfi, Chang & Engelmann, 2022). A disclaimer could theoretically impact a trust evaluation by changing the information available to a citizen in the external environment. Disclaimers make explicit the decisions made by an advertiser, for example, that an advert has been paid for or that it has been micro-targeted at specific groups. Furthermore, it is thought that disclaimers increase the saliency of the persuasive intent behind an advertisement (Jost et al., 2022). For unknown campaigners, when pre-existing knowledge is absent, citizens may rely more on environmental cues, potentially increasing the impact disclaimers have on trust evaluations.

Despite a number of studies seeking to explain the impact of disclaimers on perceptions of advertiser trustworthiness, no clear association has emerged. Although the presence of a disclaimer appears to increase awareness of a persuasive attempt (Dobber et al., 2023; Krukiemeier, Sezgin & Boerman, 2016), there is no support a disclaimer consistently impacts the perceived trustworthiness of the advertiser in a predictable direction (Binder et al., 2022; Dobber et al., 2023; Jost et al, 2022; Krukiemeier, Sezgin & Boerman, 2016; Ridout, Franz & Franklin-Fowler, 2015). For example, Dobber, Kruikemeier, Helberger, and Goodman (2023) found some support that disclaimers under some conditions can have a small positive impact on perceptions of credibility[1], while Ridout, Franz and Franklin-Fowler (2015) found the presence of a disclaimer negatively impacted perceptions of both credibility and trustworthiness. Other studies find no association (e.g., Binder et al., 2022; Jost et al., 2022; Krukiemeier, Sezgin & Boerman, 2016). One potential explanation of these conflicting findings is that all referenced studies except Ridout, Franz and Franklin-Fowler (2015) used political entities with a high likelihood of being known to participants, the implications of which are discussed below.

Pre-existing Knowledge and Trust Evaluating the environment is a time and energy costly process (Gigerenzer & Gaissmaier, 2011). The reinforcement learning model of trust posits that once perceptions of trustworthiness are established between a trustor and trustee through repeated experience, individuals will use their learned predictions about the trustworthiness of the trust object to predict the trustworthiness of their future actions (O'Doherty, Lee & McNamee, 2015). In contrast, when faced with a novel trust object where an impression of trustworthiness must be formed, individuals instead rely on cues from the environment to infer the likelihood that a trustee can be trusted (Brudner, Karousatos, Fareri & Delgado, 2022). The

relative importance of the perceptions elicited by the presence of a disclaimer in a trust evaluation are therefore hypothesised to be greater in a context where the advertiser is previously unknown. This hypothesis is supported by Ridout, Franz and Franklin-Fowler (2015) who tested fictional campaigners, finding perceived trustworthiness of an advertiser was lower in conditions when a disclaimer was present versus other forms of communication. However, the latter study tested disclaimers with adverts promoted by interest groups, conditions of which have been evidenced to evoke different evaluations of trust compared to candidate or political party campaigns (Weber, Dunaway & Johnson, 2012; Binford et al., 2021). Thus, this study further contributes to the field by testing the effect of disclaimers for an unknown political party.

Disclaimer Format and Recall Another consideration in the research field is the extent to which disclaimers are noticed and processed, considering the peripheral role they play on an advertisement (Dobber et al., 2023). Recall of the information contained on a disclaimer is often used by researchers as a proxy for noticing (Binford et al., 2021; Jost et al., 2022). This is by no means a perfect proxy, yet the methodological ways to capture if the information on a disclaimer has been processed are limited (Binford et al., 2021). For example, eye tracking studies have shown no clear association between time spent observing a disclaimer and subsequent effects, and it recognised information can be encoded peripherally (Jost et al., 2021). Measuring recall therefore provides a means by which noticing of disclaimer information can be captured (Binford et al., 2021). Furthermore, it is theoretically feasible that any associations between disclaimers and trust will only be observed in those who engage fully with the disclaimer (Krukiemeier, Sezgin & Boerman, 2016), thus comparisons between those who do and do not recall viewing the information may further illuminate which theoretical mechanisms are supported.

Method

A between-subjects experimental survey (N = 1222) was designed that randomly assigned participants into one of two conditions: control (no adverts viewed with a disclaimer) and high transparency (all adverts viewed with disclaimer). Participants viewed Facebook-style political advertisements promoting a fictional political party named 'The Reconnect Party'.

Recruitment UK citizens were recruited using the online recruitment platform Prolific. This recruitment platform was designed for academic research purposes and has been shown

---

[1] It is noted credibility and trustworthiness are distinct constructs. Credibility relates more to the perceived quality of information communicated, and trustworthiness the intent of the information

source. However, theoretical and empirical accounts support that support that they are closely related (Hilligoss & Rieh, 2008; Metzger & Flanagin, 2013).
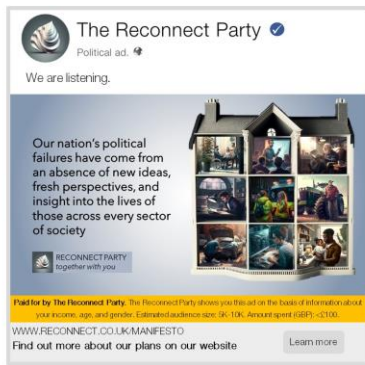
Figure 1: One example from the four advertisements shown to participants (high transparency condition). The yellow banner holds the disclaimer information.

to offer advantages regarding representative samples and participant engagement (Palan & Schitter, 2018; Litman, Moss, Rosenzweig & Robinson, 2021). To ensure data quality, participants were excluded on the following criteria: failing two (out of two) attention checks; and completing the survey in under half the median time (<4 minutes).

Materials Four advertisements were designed by the research team for the purpose of this study, an example of which can be viewed in figure 1. The advertisements were designed with the aim of obscuring ideological or partisan cues by communicating a generic campaign message; associations with existing UK parties were minimised through the party name, slogan and colour scheme. The advertiser was the same as who was advertised; the advertisements both promoted, and were promoted by, The Reconnect Party.

Disclaimer Format In the control condition, the advertisements were viewed without a digital disclaimer. In the experimental condition, the participants viewed a digital disclaimer with the following text: 'Paid for by The Reconnect Party. The Reconnect Party shows you this ad on the basis of information about your age, income and gender. Estimated audience size: xK-xK. Amount spent: £x.' The numbers replaced by an 'x' varied between advertisements and was included to increase the believability of the fictional campaign by mirroring information disclosed by the Facebook advertising library. Targeting wording was adapted from Dobber, Kruikemeier, Helberger and Goodman (2023). To further enhance conditions under which a disclaimer might have an effect, the disclaimer was designed to be more visually intrusive than those usually included on sponsored Facebook posts and was formatted as a bright yellow banner placed underneath the advertisement image. Increasing visual intrusiveness has been evidenced to boost disclaimer recall (Jost et al., 2022).

Measures

Perceptions of trustworthiness are measured using items designed to capture how the respondents evaluated the political party with regards to both trust and distrust, capturing three outcome measures: perceived trustworthiness, untrustworthiness and manipulativeness. (Un)trustworthiness item wordings were adapted from Weinberg (2022) who developed the items to capture citizen evaluations of the intentions of politicians, thus making the phrasing highly relevant to this experimental context. In line with Weinberg (2022), and in response to recent empirical support that indicates trust and distrust decisions occur through distinct neural mechanisms (for a review see Haas, 2022), trustworthiness and untrustworthiness are measured separately. Perceived manipulativeness, another distrust evaluation, is measured due to its contextual relevance in relation to evaluations of persuasive communication forms such as advertising (for details see Binder et al., 2022).

Perceived Trustworthiness Items were measured on a scale of 1 (strongly disagree) to 7 (strongly agree) across 6 items: 'The Reconnect Party seems like they would perform their job competently'; 'The Reconnect Party seems like they would be good at getting the job done'; 'The Reconnect Party seem as if they adhere to a strong moral code'; 'It seems The Reconnect Party would tell the truth'; 'The Reconnect Party seems like they would treat all groups in society fairly'; 'The Reconnect Party seems to care about other people' (Cronbach's alpha: 0.91).

Perceived Untrustworthiness Items were measured on a scale of 1 (strongly disagree) to 7 (strongly agree) across 6 items: 'The Reconnect Party seems to lack technical expertise'; 'The Reconnect Party seems like they would waste public money'; 'The Reconnect Party seems like they would distort facts to make their policies look good'; 'The Reconnect Party are happy to make promises at elections, but it seems like they would forget them afterwards'; 'It seems like The Reconnect Party don't really understand the problems facing ordinary people'; 'The Reconnect Party seems like they would look after their own interests rather than trying to help others' (Cronbach's alpha: 0.86).

Perceived Manipulativeness A single item captured this construct, measured on a scale of 1 (strongly disagree) to 7 (strongly agree): 'The Reconnect Party's adverts are manipulative'.

Recall Two binary yes/no response items were measured at the end of the survey: 'When viewing The Reconnect Party's advertisements earlier, did the adverts include information informing you who paid for them?'; 'Did the adverts include information informing you if it had been targeted towards certain groups?'.

Table 1: Outcome variable distributions by experimental condition

		No disclaimer viewed (control)				Disclaimer viewed			
	Scale	Mean	SD	Min	Max	Mean	SD	Min	Max
Perceived trustworthiness	1-7	4.94	0.93	1.17	7.00	4.94	0.86	1.00	7.00
Perceived untrustworthiness	1-7	3.17	1.00	1.00	6.67	3.20	0.95	1.00	6.67
Perceived manipulativeness	1-7	3.33	1.38	1.00	7.00	3.42	1.33	1.00	7.00

Covariate Agreement with the campaign message, a proxy for political fit, is included in each tested model as a covariate. A single item was measures on a scale of 1 (completely disagree) to 5 (completely agree): 'How much do you agree or disagree with the messages put forward in The Reconnect Parties' campaign'.

Pre-registration
The methods and some hypotheses for this research are preregistered at: https://doi.org/10.17605/OSF.IO/KNQ2B. However, note that the hypotheses and analysis in this paper are different to the preregistered plan, which focused on the effect of pre-existing beliefs as moderators of the effect of experimental condition. This paper presents main effects of experimental condition only. Additionally, analyses of the effect of recall were not pre-registered are exploratory in their nature.

Data Availability
The full dataset for the study is available at https://osf.io/gybnm/.

Results
A full sample breakdown can be viewed at https://osf.io/gybnm/ in document 'sample.pdf'.

Effect of Viewing the Disclaimer
The outcome measures and the covariate of agreement were included as numerical variables in all models. Trustworthiness and untrustworthiness items were mean averaged to create one score for each respondent. Table 1 shows the descriptive statistics for the outcome measures (trustworthiness, untrustworthiness and manipulation), shown separately across the control condition (no disclaimer viewed) and the experimental condition (disclaimer viewed). Overall, respondents skewed towards having positive perceptions of the political party and their campaign, disagreeing that the advertisements were manipulative, and agreeing that the political party were trustworthy.

An OLS linear regression modelling technique was used to compare the effect of the experimental condition (viewing all adverts with a disclaimer) against the control (no disclaimer viewed) for each of the outcome measures, including the agreement covariate. Three models were tested, finding no significant difference between scores across the control and

experimental condition for perceived trustworthiness ($B = -0.01$, $p = .71$; model fit: $F(2, 1219) = 459.9$, $p < .001$, adj. $R^2 = .43$), perceived untrustworthiness ($B = 0.04$, $p = .33$; model fit: $F(2, 1219) = 330.3$, $p < .001$, adj. $R^2 = .35$), or perceived manipulativeness ($B = 0.10$, $p = .15$; model fit: $F(2, 1219) = 114.0$, $p < .001$, adj. $R^2 = .16$).

The covariate of agreement (a proxy for political fit with the campaign message) was a significant predictor in each of the three models, with a 0.76 (SE = 0.03, $p < .001$) predicted scale increase in perceived trustworthiness, a 0.75 (SE = 0.03, $p < .001$) predicted scale decrease in perceived untrustworthiness, and a 0.70 (SE = 0.04, $p < .001$) scale decrease in perceived manipulativeness, for every one scale increase in agreement.

Effect of Recalling the Disclaimer
An exploratory analysis was conducted to investigate the effect of recalling the specific information communicated by the disclaimer on the three outcomes measures. As the disclaimer communicated multiple pieces of information, sponsorship and targeting recall were measured separately. Table 2 shows the rate of recall across the conditions. Within the experimental condition, 52.7% of participants correctly recalled that they had been shown sponsorship information (26.4% of total sample), and a lower percentage of 44.5% correctly recalled that they had been shown targeting information (22.3% of total sample).

The two recall measures were transformed into categorical predictors consisting of 4 factor levels: correctly recalling the disclaimer in the experimental condition (level 1, reference level); incorrectly not recalling the disclaimer in the

Table 2: Correct and incorrect recall

Condition	Recall	Sponsorship Information		Targeting Information	
		N	%	N	%
Control	Correct	547	44.8	467	38.2
	Incorrect	64	5.2	144	11.8
Disclaimer viewed	Correct	322	26.4	272	22.3
	Incorrect	289	23.7	339	27.7

Note. Those in the control condition were rated as correct when they answered 'no' to the recall measure. Those in the experimental condition were rated as correct when they answered 'yes' to the recall measure.

Table 3: OLS linear regression models testing the effect of recall on the three outcome measures

	Dependent variable		
	Trustworthiness B(SE)	Untrustworthiness B(SE)	Manipulativeness B(SE)
Sponsorship disclaimer recall (*ref = correct recall, disclaimer viewed*)			
No recall, disclaimer viewed	-0.06 (0.06)	**0.17 (0.06) ****	**0.27 (0.10) ****
No recall, no disclaimer viewed	-0.02 (0.05)	0.05 (0.06)	0.03 (0.09)
Incorrect recall, no disclaimer viewed	<.001 (0.09)	-0.04 (0.11)	-0.04 (0.17)
Agreement	**0.79 (0.03) *****	**-0.75 (0.03) *****	**-0.69 (0.05) *****
F-statistic	230.0 (4, 1217) ***	167.9 (4, 1217) ***	59.0 (4, 1217) ***
Adjusted R^2	0.43	0.35	0.16

Note. ***p <.001, **p < 0.01, *p < 0.05, B = unstandardised coefficient, SE = standard error, significant results shown in bold.

experimental condition (level 2); correctly not recalling the disclaimer in the control condition (level 3); incorrectly recalling the disclaimer in the control condition (level 4).

Testing the effect of recall even in those who had not been shown the disclaimer (control condition) was necessary to provide context for any effect that recalling the disclaimer had. For example, it can be seen in table 2 that some respondents who had not been shown a disclaimer recalled viewing one (sponsorship: 5.2% of total sample, targeting: 11.8% of total sample). This raised the possibility that the measure of recall not only captured attention paid to the disclaimer. This has important implications for the interpretation of any significant association found between recall and perceived trustworthiness. For example, this analysis assumes a causal relationship between engaging with the disclaimer and changes to perceptions of trustworthiness. Yet, if the measure of recall does not only reflect engagement, then an association could be interpreted in reverse: trusting the content more increases the likelihood respondents would answer 'yes' for the recall measures. Furthermore, as agreement with the campaign message was associated strongly with perceptions of trustworthiness, it could be the case that more agreement motivated respondents to look at the content for longer, thus partly accounting for any association between recall and trust. These possibilities were checked for using the following logic: if the recall measures reflect trust in the political party rather than noticing of the disclaimer then we would expect, for those in the experimental condition, recalling the disclaimer information (level 1) to be significantly associated with higher levels of agreement compared to not recalling the disclaimer information (level 2).

To check for this, the extent to which recall of the sponsorship and targeting information were associated with agreement was first modelled in two separate OLS linear regressions which included recall as a predictor and agreement as an outcome. For both recall predictor variables, the four categorical levels previously outlined were included in the models. The resulting model for sponsorship recall

supported a significant difference between level 1 and level 2 for the outcome agreement (B = -0.13, SE = 0.06, p = .04). However, the reliability of this prediction was compromised by a poor model fit (F(3, 1281) = 1.49, p = .22, adj. R^2 = .001). Accordingly, key OLS model assumptions were violated by the presence of outliers in the outcome measure. To improve the reliability of the predictions, the model was re-fitted with a robust linear regression technique, using the MASS package in R studio and fitting the model with M-estimated Huber weights which are effective at reducing the influence of outliers in model estimates (for a review see De Menezes, Prata, Secchi & Pinto, 2021). The re-fitted model did not replicate the significant effect for this comparison (B = <.001, SE = <.001, p = .09). For the second recall measure, recall of the targeting information, no difference was found between level 1 and level 2 in either an OLS regression or a robust regression. Thus, there was not strong evidence to suggest that recall was associated with agreement. This increases confidence in the claim that for the respondents in the experimental condition, any association between recall and the outcomes reflected the effect of attention paid to the disclaimer on subsequent trust evaluations, justifying a directional hypothesis.

Recall of Sponsorship Information Three separate OLS linear regression models were fitted with sponsorship recall as a predictor and agreement as a covariate against the three outcome measures (trustworthiness, untrustworthiness and manipulativeness). Table 3 describes each model, showing that there was a significant association between not recalling the sponsorship disclaimer when it had been viewed (level 2) and higher reported levels of untrustworthiness (B = 0.17, SE = 0.06, p = .01) and manipulativeness (B = 0.27, SE = 0.10, p = .01), compared to the reference level. Furthermore, the reference level did not significantly differ with respect to these outcomes from either group in the control condition (level 3 and 4).

Recall of Targeting Information These effects were not

replicated when including the targeting recall factor included as a predictor. In each comparison between the reference level and level 2, no significant difference was found for perceived trustworthiness ($B = 0.01$, $p = .91$), untrustworthiness ($B = 0.05$, $p = .40$) or manipulativeness ($B = -0.07$, $p = .50$).

Discussion

This paper described a survey experiment designed to investigate what informational cues are prioritised by citizens when evaluating the trustworthiness of an unknown political campaigner, specifically testing differences in evaluations when sponsorship and microtargeting transparency disclaimers are present and absent. Previous research has theorised conflicting mechanisms as to the effect of such disclaimers on subsequent trust evaluations. Some researchers emphasise the potential for disclaimers to draw attention to the persuasive nature of a campaign message (Jost et al., 2022), and theorise their presence will elicit scepticism as a defence against the persuasion attempt (Friestad & Wright, 1994; Dobber et al., 2023; Binder et al., 2022; Krukiemeier, Sezgin & Boerman, 2016; Jost et al., 2022; Binford et al, 2021). Other researchers alternatively emphasise the potential for transparency to act as a trust cue that communicates open and honest intentions (Beckert et al., 2021; Wood, 2022), considerations that might be expected to predict higher perceptions of trustworthiness.

It was theorised that when a campaigner is unknown, the relative use of information cues available in the environment will increase in an evaluation, as existing beliefs about the trustee that are usually prioritised in trust evaluations (O'Doherty, Lee & McNamee, 2015) are absent. The logic followed that under these conditions any effect of disclaimers on trust would be enhanced, enabling a clearer understanding of which theoretical mechanism is supported. Alternatively, investigating disclaimers under conditions of maximal impact can reveal whether transparency disclaimers are *not* perceived by UK citizens as a relevant informational cue with respect to the trustworthiness of campaigners. As well as providing theoretical clarity, this has practical implications when it comes to establishing if legislative goals are feasible. For example, the UK government expects disclaimers to increase trust in campaigns (Cabinet Office & Smith, 2021) while EU legislators hope they will enhance scepticism about political information distributed online (Dobber et al., 2023).

This experiment found no support that the presence of a disclaimer increased perceptions that the campaigner was trustworthy. In line with Binder, Stubenvoll, Hirsch and Matthes (2022), these findings support that agreement with the political message of the campaign was used as the primary means of determining trust.

However, there was some evidence that reported recall of the information communicated by the disclaimer was associated with *dis*trust perceptions (trust and distrust were measured separately, for a review see Haas, 2022). The recall measure was originally designed to capture the extent to which respondents paid attention to the disclaimer. It was assumed those in the experimental condition who were shown the disclaimer but did not recall it had processed the information less than those in the same condition who correctly recalled it. However, there was no support for a significant difference in the trustworthiness, untrustworthiness and manipulativeness perceptions between those who correctly recalled the sponsorship disclaimer who were shown it, and those in the control condition who were not shown disclaimers at all. Unexpectedly, respondents who reported not recalling the sponsorship information when they had indeed been shown it were found to have significantly higher perceptions of untrustworthiness and manipulativeness, indicating more distrust in the campaigner in this group of respondents.

This casts doubt on the claim that the recall measure captured attention paid to disclaimer information. It instead suggests the recall measure tapped into a different perception, nonetheless one that is potentially relevant to understanding how the presence of a disclaimer impacted the distrust evaluations made by this group of respondents. One speculative interpretation of this effect is that the recall measure captured individual differences in the reaction to a disclaimer. For example, increasing transparency may have, for some, elicited scepticism in the motives of the campaigner causing them to question the extent to which the campaigner was being completely open when revealing their sponsorship affiliations. This raises an important question about how much information citizens perceive themselves to need before they can feel affirmed that they have full transparency, rather than partial transparency that keeps unfavourable information hidden. If citizens perceive only partial transparency to be achieved, this may lead the disclosure of more information to increase scepticism rather than reduce it, compared to if no additional information had been provided (O'Niell, 2006). This interpretation would support the theoretical accounts that predict disclaimers will increase scepticism (Dobber et al., 2023; Binder et al., 2022; Krukiemeier, Sezgin & Boerman, 2016; Jost et al., 2022; Binford et al., 2021), but further suggests this only happens for some citizens.

Due to the exploratory nature of the recall analysis, the significance of these effects would need to be replicated to have confidence in such an interpretation. Additionally, it is unclear as to why the same effect was not found for recall of the micro-targeting part of the disclaimer on perceptions of distrust. This highlights a key limitation of this study, and it is recommended future research test different aspects of disclaimer information separately.

In conclusion, this research adds to the body of empirical work testing the effect of political advertising disclaimers, further building confidence in the claim that they do not act as a cue to trust. However, some preliminary support is provided for the claim that disclaimers can increase scepticism, further suggesting there are individual differences in reactions to disclaimer content; a deeper understanding of which may help add further clarity to theoretical accounts in this research field.

References

Beckert, J., Koch, T., Viererbl, B., & Schulz-Knappe, C. (2021). The disclosure paradox: How persuasion knowledge mediates disclosure effects in sponsored media content. *International Journal of Advertising*, *40*(7), 1160-1186.

Binder, A., Stubenvoll, M., Hirsch, M., & Matthes, J. (2022). Why am I getting this ad? How the degree of targeting disclosures and political fit affect persuasion knowledge, party evaluation, and online privacy behaviors. *Journal of Advertising*, *51*(2), 206-222.

Binford, M. T., Wojdynski, B. W., Lee, Y.-I., Sun, S., & Briscoe, A. (2021). Invisible transparency: Visual attention to disclosures and source recognition in Facebook political advertising. *Journal of Information Technology & Politics*, *18*(1), 70–83.

Birch, S. (2011). *Electoral malpractice*. USA: Oxford University Press.

Brudner, E., Karousatos, A., J., Fareri, D. S., & Delgado, M. R. (2022). Trust and reputation: How knowledge about others shapes our decisions. In F. Krueger (Eds.), *The Neurobiology of Trust*. Cambridge: Cambridge University Press.

Cabinet Office & Smith, C. (2021) Transparency in Digital Campaigning: Government Response. London. Available at:https://www.gov.uk/government/consultations/transparency-in-digital-campaigning-technical-consultation-on-digital-imprints

Dahl, R. A. (1989). *Democracy and its Critics*. London: Yale University Press.

De Menezes, D. Q. F., Prata, D. M., Secchi, A. R., & Pinto, J. C. (2021). A review on robust M-estimators for regression analysis. *Computers & Chemical Engineering*, *147*, 107254.

Dermody, J., & Scullion, R. (2005). Young people's attitudes towards British political advertising: Nurturing or impeding voter engagement?. *Journal of Nonprofit & Public Sector Marketing*, *14*(1-2), 129-149.

disinformation, and moral boundaries. *Journal of Business Ethics*, *17*, 1871-1885.

Dobber, T., Kruikemeier, S., Helberger, N., & Goodman, E. (2023). Shielding citizens? Understanding the impact of political advertisement transparency information. *New Media & Society*, 14614448231157640.

Dobber, T., Trilling, D., Helberger, N., & de Vreese, C. (2019). Spiraling downward: The reciprocal relation between attitude toward political behavioral targeting and privacy concerns. *New Media & Society*, *21*(6), 12121231.

Engelmann, J. B., & Hare, T. A. (2018). Emotions can bias decision-making processes by promoting specific behavioral tendencies. In A. S. Fox, R. C. Lapate, A. J Shackman & R. J. Davidson (Eds.), *The nature of emotion:*

fundamental questions. New York: Oxford University Press.

Farolfi, F., Chang, L., & Engelmann, J. B. (2022). Trust and emotion: The effects of incidental and integral affect. In F. Krueger (Eds.), *The Neurobiology of Trust*. Cambridge: Cambridge University Press.

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of consumer research*, *21*(1), 1-31

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, *62*, 451-482.

Haas, B. W. (2022). Trust and Distrust: Key Similarities and Differences. In F. Krueger (Eds.), *The Neurobiology of Trust*. Cambridge: Cambridge University Press.

Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, *44*(4), 1467-1484.

Jost, P., Kruschinski, S., Sülflow, M., Haßler, J., & Maurer, M. (2022). Invisible transparency: How different types of ad disclaimers on Facebook affect whether and how digital political advertising is perceived. *Policy & Internet*, *15*(2), 204-222.

Kates, S. (1998). A qualitative exploration into voters' ethical perceptions of political advertising: Discourse,

Kruikemeier, S., Sezgin, M., & Boerman, S. C. (2016). Political microtargeting: relationship between personalized advertising on Facebook and voters' responses. *Cyberpsychology, Behavior, and Social Networking*, *19*(6), 367-372.

Lau, R. R., Sigelman, L., & Rovner, I. B. (2007). The effects of negative political campaigns: A meta-analytic reassessment. *The Journal of Politics*, *69*(4), 1176-1209.

Lauharatanahirun, N., & Aimone, J. A. (2022). Trust and risk: Neuroeconomic foundations of trust based on social risk. In F. Krueger (Eds.), *The Neurobiology of Trust*. Cambridge: Cambridge University Press.

Litman, L., Moss, A., Rosenzweig, C., & Robinson, J. (2021). Reply to MTurk, Prolific or panels? Choosing the right audience for online research. SSRN.

Matthes, J., Hirsch, M., Stubenvoll, M., Binder, A., Kruikemeier, S., Lecheler, S., & Otto, L. (2022). Understanding the democratic role of perceived online political micro-targeting: longitudinal effects on trust in democracy and political interest. *Journal of Information Technology & Politics*, *19*(4), 435-448.

Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of pragmatics*, *59*, 210-220.

Norris, P. (2003) Tuned out voters? Media impact on campaign learning. *Ethical Perspectives*, *9*(4), 200-21.

O'Doherty, J. P., Lee, S. W., & McNamee, D. (2015). The structure of reinforcement-learning mechanisms in the human brain. *Current Opinion in Behavioral Sciences*, *1*, 94-100.

O'Niell, O. (2006). Transparency and the Ethics of Communication. In C. Hood & D. Heald (Eds.),

Transparency: The Key to Better Governance? New York: Oxford University Press.

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22-27.

PytlikZillig, L. M., & Kimbrough, C. D. (2016). Consensus on Conceptualizations and Definitions of Trust: Are We There Yet?. In E. Shockley, T. M. S. Neal, L. M., PytlikZillig, & B. H. Bornstein (Eds.), *Interdisciplinary Perspectives on Trust: Towards Theoretical and Methodological Integration*. Springer.

Ridout, T. N., Franz, M. M., & Franklin-Fowler, E. (2015). Sponsorship, disclosure, and donors: Limiting the impact of outside group ads. *Political Research Quarterly*, *68*(1), 154-166.

Siegrist, M., Earle, T. C., & Gutscher, H. (2010). Trust, risk perception and the TCC model of cooperation. In M. Siegrist, T. C. Earle & H. Gutscher (Eds.), *Trust in risk management: Uncertainty and scepticism in the public mind*. UK: Earthscan.

Simpson, J. A., & Vieth, G. (2022). Trust and psychology: Psychological theories and principles underlying interpersonal trust. In F. Krueger (Eds.), *The Neurobiology of Trust*. Cambridge: Cambridge University Press.

Thomson, R., & Brandenburg, H. (2019). Trust and citizens' evaluations of promise keeping by governing parties. *Political Studies*, *67*(1), 249-266.

Weber, C., Dunaway, J., & Johnson, T. (2012). It's All in the Name: Source Cue Ambiguity and the Persuasive Appeal of Campaign Ads. *Political Behavior*, *34*(3), 561–584.

Weinberg, J. (2022). Can political trust help to explain elite policy support and public behaviour in times of crisis? Evidence from the United Kingdom at the height of the 2020 coronavirus pandemic. *Political Studies*, *70*(3), 655-679.

Wood, A. K. (2022). Voters use campaign finance transparency and compliance information. *Political Behavior, 45*(4), 1553-1579.

Zuiderveen Borgesius, F., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., Bodo, B., & de Vreese, C. H. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, *14*(1), 82-96.