



UvA-DARE (Digital Academic Repository)

Data & Democracy

Political microtargeting: A threat to electoral integrity?

Dobber, T.

Publication date

2020

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):

Dobber, T. (2020). *Data & Democracy: Political microtargeting: A threat to electoral integrity?*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 4

Measuring effects of issue-based microtargeted messages on vote likelihood: reporting a field experiment in a multi-party setting

Abstract

Political microtargeting is the subject of heated societal debate but not much is known about its effects, especially in non-US contexts. Microtargeting, used by political actors to send citizens tailored messages, could have the potential to overcome barriers that make generic political messages less effective. In this paper, we present a small-scale field experiment, which serves as a case study to illustrate how microtargeting's effects on citizens could be measured. The case study, as the first European-based empirical test of microtargeting's effectiveness, showed that receiving a microtargeted message increased vote likelihood, but this increase did not translate into actual votes.

Introduction

Political microtargeting (PMT) techniques are gaining popularity in European political campaigns (Anstead, 2017; Dobber, Trilling, Helberger, & De Vreese, 2017). The appeal of microtargeting is that people are potentially more responsive to messages that are personally relevant than to generic messages.

Critics of microtargeting warn for the potential for manipulation (Gorton, 2016; Zuiderveen Borgesius et al., 2018), which, for instance, can occur when a political message deliberately evokes a feeling of 'social pressure' or when a campaign tries to leverage feelings of fear. Another risk of microtargeting is the possibility for the infringement of citizens' (intellectual) privacy (Richards, 2015; Rubinstein, 2014). Critics also warn for the information asymmetry between the campaign and the targeted voter (Tufekci, 2014), which occurs when campaigns know a lot about the individual voter, but the voter knows (almost) nothing about the data used by the campaign to target her.

Proponents praise PMT's potential to interest voters in political issues that are relevant to them personally. Consequently, microtargeting could increase political participation (Zuiderveen Borgesius et al., 2018). For example, a student may be more inclined to respond to a political message about student housing than to a message about pensions, and the opposite may be true for the pensioner.

Microtargeting's threats and promises are quite salient, but there is only limited knowledge about effects of PMT on citizens. Present literature is US-based, and not easily generalizable to non-US contexts with different regulatory regimes, campaign budgets, and electoral systems.

Measuring effects of political microtargeting comes with challenges. For instance, how can you determine *a priori* which participant should receive which specific stimulus tailored to her personally? In this study, we explore a possible research set-up for measuring the effects of PMT, conduct a field experiment in a multi-party context, and in doing so, contribute novel empirical evidence to the microtargeting debate by answering the following key question: *'What are effects of microtargeted issue-based messages on citizens' vote likelihood and issue salience?'*

The field experiment occurred in a real election, and consisted of a pre and a post survey. The pre-survey served to discover which issue was most salient to the participant, and how likely participants deemed themselves to vote for specific political parties. This information helped to determine which participant should receive which stimulus message. The tailored stimulus was administered on March 15 and March 16: a few days before the election (March 21, 2018). The post survey was conducted right after election day to measure potential changes in vote likelihood and issue salience.

Further conceptualizing political microtargeting

Common in the description of PMT is the understanding that PMT requires personal information of citizens, the application of (big) data analysis, and tailoring messages to the receiver (Gorton, 2016; Hersh, 2015; Rubinstein, 2014; Tufekci, 2014; Zuiderveen Borgesius et al., 2018). However, the literature is not clear about when regular targeting becomes *microtargeting*. How small does the targeted audience have to be before we speak of microtargeting? We argue that to decide whether a political actor is *microtargeting*, we should not look at the size of the targeted group per se. Rather, the microtargeted group should be 1) a subgroup of a general population, 2) this segmentation should be based on two or more characteristics, 3) the subgroup should be perceived by the campaign as homogenous, and 4) the campaign should send different messages to different audiences.

Imagine the political actor is the party of French president Emanuel Macron, who is campaigning for national elections. If Macron's campaign would target law students of the Sorbonne university with a message tailored to them being Sorbonne students *and* target medicine students of the University of Paris with a different tailored message, we should speak of microtargeting. One, all the students of the Sorbonne are a subgroup of the general population, as are the University of Paris students. Two, the segmentation of Macron is based on several characteristics: *law students* who *study at the Sorbonne University*, or *medicine students* who *study at the University of Paris*. Three, Macron may deem the Sorbonne students homogeneous, and the University of Paris students too, and likely for the purpose of a national election they are. Four, the Sorbonne students as well as the University of Paris students receive a different message tailored to them belonging to their respective universities. Sending more two (or more) tailored messages, one to the Sorbonne students, and one to the University of Paris students, is crucial. Microtargeting should, when done correctly, ignore the people who are perceived as not susceptible to the message (Sorbonne students are unsusceptible toward the message tailored to University of Paris students and vice versa), and only reach the perceived susceptible people, ideally raising the average effectiveness of both messages.

Why could microtargeting work?

In this section, we explore an undertheorized element of PMT: which mechanisms could influence PMT's effectiveness?

When done correctly, microtargeted messages are important to the receiver, and congruent to her opinions, which would increase the chance of message scrutiny and, subsequently, the chance of influencing the message recipient (Petty et al. 1995; Wheeler, Petty & Bizer, 2005; Chang 2006; Petty and Cacioppo 1986; Wheeler, DeMarree, & Petty, 2008).

According to dual-processing theories, the processing of information, and subsequent decision-making, is handled by two different types (see Evans & Stanovich, 2013, for an overview of critique on dual-processing theory). These types are discussed in different terminology: ‘system 1 and system 2’ (Kahneman, 2011), ‘type 1 and type 2’ (Evans & Stanovich, 2013), ‘peripheral and central’ (Petty & Cacioppo, 1986), ‘heuristic and systemic’ (Chaiken, 1980). The consensus is that there is a fast, autonomous type of processing that does not require working memory (‘Type 1’), and a reflective type of processing that *does* require working memory and is capable of cognitive decoupling: a relative complex form of problem solving (Koichu & Leron, 2015). This is known as ‘Type 2’ processing (Evans & Stanovich, 2013). Persuasion can occur via Type 2 processing, and should result in a relatively enduring persuasion. But persuasion can also occur through Type 1 processing, which should result in temporary. Microtargeting should increase the likelihood ‘successful’ processing, Type 1 or Type 2, by making a message personally salient, easy to understand, and by introducing a powerful argument on the basis of the receiver’s personal situation, beliefs, behavior, and/or traits (e.g., Petty & Cacioppo, 1986). Especially activating Type 2 processing should increase the chance of persuasion, and, therefore, increase the likelihood to vote for the campaign that sent the microtargeted message that activated Type 2 processing.

Literature on effects of political microtargeting

Measuring PMT is challenging, but not impossible. This paper presents a field experiment that serves as a case showing how PMT could be measured. Before we discuss the case more elaborately, we briefly discuss the literature from which we derive the expectations.

The field of microtargeting research is in its infancy. While the literature is steadily growing, there is still much unknown about the actual effects of the technique. Endres and Kelly (2018) have focused on how the use of predictive scores (a measure that many microtargeting efforts are based on, which can for example be a calculation of a voter’s likelihood to vote for a certain party) have implications for specific demographic groups of the electorate. Young people, for example, are much more likely to be systematically ignored by political campaigns that use predictive scores to target advertisements. Actual effects of political microtargeting on citizens are measured exclusively in a US setting. Hersh and Schaffner (2013), warned for negative effects of wrongly targeting an ethnic group (‘mistargeting’). Haenschen and Jennings (2019) showed that online PMT ads boosted voter turnout, but only for young voters in competitive districts. Endres (2019) found that issue-congruent microtargeted messages increase cross-pressured voters’ support for the campaign that sent the microtargeted message.

A meta-analysis of the *persuasive* effects of campaigning techniques found almost no evidence for any persuasive potential of non-microtargeted political messages (Kalla & Broockman, 2017). However, Kalla and Broockman (2017) note that finding

persuadable voters can be a successful strategy, but this requires sizable effort and resources.

From a European perspective, the US focus of microtargeting research is suboptimal because the findings cannot easily be generalized to European contexts. Research has found that the originally American microtargeting techniques take on a hybrid form, dependent on contextual factors when applied in Europe (Anstead, 2017; Dobber et al., 2017). Unfortunately, as of yet, there are no empirical studies on effects of PMT in a European, multiparty, context. Despite the contextual differences between US and EU contexts, on the basis of present research we expect the following.

H1: Receiving an issue-based microtargeted message increases the likelihood to vote for the party that sent the microtargeted message.

Issue salience

A microtargeted message should be about an issue that is personally salient to the voter. Contacting someone about an issue that is not important to her, would be illogical: how could a campaign influence someone by appealing to an issue irrelevant to the voter? Previous research has indeed found the conditional effect of issue salience on issue based voting (Bélanger & Meguid, 2008). These findings are based upon a shared agenda of issues that are salient to the general public. Agenda setting theory describes a public agenda, set by the topics traditional media cover (McCombs & Shaw, 1972). However, we would argue that with the rise of personalized news and political microtargeting techniques, the ‘shared’ agenda of issues may become increasingly individualized. If a campaign infers which issue is (most) important to a potential voter, say ‘education’, it would send only political information about education to that specific voter. As a consequence, the individual voter may come disproportionately in contact with one specific issue and may therefore assign more relative weight to education, and less weight to the other issues. This leads to the second hypothesis:

H2: Receiving an issue-based microtargeted message increases the relative salience of the targeted issue.

Case and methods

Using a field experiment to study PMT comes with upsides and downsides. Different approaches are possible, and should be used in future research to get a better grasp of PMT’s effects on political behavior. As such, the case that we present is only meant to give first, but not definitive, insight into PMT’s effects in a natural, multiparty setting, *and* into possible ways to measure them.

Case

This field experiment took place in the context of the 2018 municipality elections in Utrecht. This is a one district, proportional representation electoral system, which means that every vote counts equally. There were 16 political parties running for office. Citizens did not vote for a candidate, but for a party. Moreover, citizens did not elect a mayor, but rather the members of the municipality council. Consequently, the personalities of specific candidates typically did not take on a big role in this municipality elections.

Municipality elections are second order elections, perceived by citizens as less important than national elections (Lefevere & Van Aelst, 2014). The turnout for the 2018 municipal election in Utrecht was 59%. In comparison, the turnout in Utrecht for the Dutch national election was almost 84% (Municipality of Utrecht, 2019).

Experimental design

To realize an optimal design, we used the house styles of two political parties: left-wing PvdA, and right-wing VVD. The field experiment took place around the municipality elections in the Netherlands (election day: March 21, 2018), and consisted of one pretreatment and one post treatment round of data collection. At T2, only 21 VVD-participants had completed the survey, making further analysis meaningless. Hence, we only focus on the much larger experimental group of PvdA-participants who have completed the survey at T1 as well as T2 ($N = 51$), and the participants in the control condition who also completed both surveys ($N = 35$).

Our experiment consisted of one control group and one experimental group. The participants in the experimental group received one out of 6 messages (see Appendix A), tailored to them personally. The local branch of PvdA did not send tailored messages at all during the campaign (they did not prioritize microtargeting in their campaign).⁹

Exposure to personalized message

Exposure occurred when a respondent received a political message tailored to her most salient issue. We tailored the message on the basis of the information on the respondents' scores on issue salience, collected at T1. To determine which tailored message we had to send the specific respondent, we looked at which issue was most salient to her and which party had the higher vote likelihood score. If the most salient issue was, for example, 'education', and the respondent's vote likelihood was highest for political party PvdA, we send her the PvdA-message about education. When more than one issues were equally salient, we randomly selected an issue between the highest scoring issues. We took the same approach when vote likelihood scores were equal.

⁹ The original design also included a 'Facebook' and a 'Facebook + mail' condition. However, this part of the study is not reported here as the attempt to buy and place the ads in the Facebook auction, via the custom audiences function, did not succeed. As a result, we can only compare and report the 'personalized ad via mail' vs the control condition in this manuscript.

We phrased six different political messages. After pre-testing these messages¹⁰, a professional designer crafted six designs that looked like authentic campaign material from the PvdA (see Appendix A for the messages). The statements on the stimuli were phrased in exactly the same way the PvdA phrased their political messages, and were designed in the house style of the PvdA. All stimuli started with the same words: ‘Be certain of’ and were followed by a statement relating to the specific issue (e.g., ‘Be certain of a safe neighborhood’). This way, we were able to pool the participants that were exposed to one of six different issue-based stimuli as one experimental group for our analyses. The stimuli were the size of postcards, and addressed in a white envelope with the address of the participants written on it by hand to decrease the chance that participant would regard the messages as unwanted advertising.

Sample

There were two rounds of data collection. The first round ($N = 124$) took place between January 20 and March 3. The second round of data collection ($N = 86$) took place between March 23 and May 4. Data collection was carried out by the lead author and three assistants. The manipulation was administered on March 15 and March 16: a few days before the election (March 21, 2018).

Before the first round of data collection began, potential respondents received a letter in which we announced our visit and explained to respondents what they could expect. These letters were sent to 1550 respondents living in the four neighborhoods of our interest, chosen because of their similar demographics (see paragraph ‘municipality selection’). The streets within each neighborhood were randomly chosen.

Then, we visited the potential respondents at home, and told them we were studying news consumption and political preference. While we were not actually interested in news consumption, we did not want to give away the focus of our experiment in advance (immediately after round 2 of data collection, we debriefed the participants on the true nature of the study). We then asked them to cooperate and to consent to our processing of their data for scientific purposes. After receiving informed consent, we administered a survey in which we measured issue salience, vote likelihood, and demographic variables year of birth, education level, gender. We also asked them for their e-mail addresses (see Appendix B for the questionnaire), and we announced that there would be a second round of data collection. After the first round of data collection, we had successfully surveyed 124 respondents ($M_{age} = 49$, $SD_{age} = 13.49$; 51% female; education: 8% was lower educated (high school or lower), 37% medium (community college and bachelor degree), 55% high (master’s degree or higher).

10 $N = 31$. All messages sufficiently referred to the specific issues they were meant to refer to. See appendix E.

The second round of data collection was conducted on the same group of people as in round 1, but this time via e-mail and via post, shortly after the election (see Appendix C for the survey of round 2). The participants who had provided us in round 1 with their e-mail addresses received an invitation for an online survey. Those who had not, received a survey per postal mail, an accompanying letter, and a return envelope. To increase responses, we sent two reminders via post or e-mail. To further increase response, we visited the ‘non-respondents’ at home and, if they were home, administered the survey right away. If the specific non-respondent was not home, we delivered a hand-written note in which we kindly asked them to fill in the survey (we also added the survey, together with a return envelope). In this second round of data collection, 86 people participated (retention rate 69%). The demographics of the respondents that answered our survey questions twice are similar to those in round 1 ($M_{age} = 50$, $SD_{age} = 12.92$; 55% female; education: 8% community college or lower, 92% had bachelor’s degree or higher), but do differ from the general population, especially regarding education level ($\mu_{age} = 41.5$; 50% female; 70% community college or lower, 30% hold bachelor’s degree or higher).

Issue salience

This is a dependent variable, and was measured at T1 as well as T2. We asked the following question “*Could you indicate for the following issues, which are part of the tasks of the municipality, how important those issues are for you personally? The score of 1 stands for not at all important, and the score of 10 stands for very important. You can also answer ‘I don’t know.’*” We then listed the following six issues: crime ($M = 9.33$, $SD = .82$, $N = 6$), integration ($M = 7.50$, $SD = 1.41$, $N = 8$), job market ($M = 8.67$, $SD = .82$, $N = 6$), quality of health care ($M = 6.75$, $SD = 1.04$, $N = 8$), quality of education ($M = 7.58$, $SD = 1.62$, $N = 12$), housing market ($M = 7.55$, $SD = 2.25$, $N = 11$)¹¹.

Vote likelihood

This variable was measured at T1 and T2. At T1, we asked the following question: “*Could you indicate the likelihood that you will cast your vote for the following parties for the coming municipality elections?*” At T2, after the elections, we asked: “*Could you indicate the likelihood that you would have cast your vote for the following parties for the recent municipality elections?*” At T2, we also asked for which party the participant had cast their vote. But ‘actual vote’ was a less useful measure because there were 16 parties running in the election. The chance that a participant voted for PvdA is slim, while their voting intention could have changed regardless. But this change would not be registered if we only looked at actual vote.

The vote likelihood item was measured on a 1-10 scale, where 1 stood for the lowest likelihood. To keep the survey relatively short, we did not ask the respondents to indicate their vote likelihood for all 16 competing parties, but only for the PvdA, the

¹¹ These are the scores of round 2.

VVD and also the center-Christian party CDA. This way, we could tell the respondents that we were only interested in the classic leftist (PvdA), center (CDA), and rightist (VVD) parties, without raising suspicion about our upcoming PvdA messages.

Non-included variable

At T2, we also measured campaign alignment: *To what extent did the campaigns of the following political parties align with the topics that you yourself find important?* Respondents could score the six major parties (CDA, D66, GroenLinks, PvdA, SP, VVD) on a scale of 1-10. they could also select *'I don't know'*. We do not include this variable in our analysis, because this question was likely too difficult to answer. The item non-response was quite high (20%).

Debriefing

We explained that we had sent respondents a tailored message and, in broad terms, explained the objective of the study. We also provided contact information of the first author and of the university's ethical committee.

Municipality selection

The field experiment took place in Utrecht, the Netherlands. Together with the local branch of PvdA, we determined 4 neighborhoods where electoral turnout is generally high and socio-demographics are comparable. These neighborhoods were: 'Oud Hoograven', 'Zeeheldenbuurt', 'Wittevrouwen', and 'Wilhelminabuurt'. We aimed for neighborhoods with high turnout, because we expected higher participation rates there¹². There are 270.000 eligible voters in the city of Utrecht. There were 16 parties running for a seat in the municipality council.

Manipulation check items

Whether the manipulation worked as intended was measured with two items. In the first item, we showed a picture of the specific manipulation that we had sent the respondent and asked whether the respondent had "*seen this message in the run up to the municipality elections*". Respondents could answer: yes, no, and don't know.

The second manipulation check item measured whether the respondent found the message appealing: "*Could you, apart from your opinion about the party behind the message, indicate to what extent this message appeals to you personally?*" This item was measured on a 7-point scale where 1 stands for 'absolutely not appealing' and 7 stands for 'absolutely appealing'.

12 We consider the consequences for the representativeness of our sample not problematic, as it is very difficult to achieve a representative sample when conducting a field experiment. We aimed for a larger sample instead of a slightly more representative sample.

PvdA did not send issue-based mailers to the participants in this study. The only mailer 'from the PvdA' they received was the experimental stimulus. The majority of respondents in the experimental condition did not see or did not remember seeing the experimental stimulus (57% stated they did not see, 14% did not remember, 29% stated they did see the stimulus). We cannot be sure that all participants in the experimental condition saw the stimulus. Not remembering (correctly) does not mean that the stimulus has failed: a message can make an imprint even if the person does not remember the message. There were two weeks or more between receiving the stimulus and completing the second survey. During an election campaign, citizens receive multiple political messages, and it is difficult to remember them all. Moreover, as Vavreck (2007) showed, there is bias in the response to these kinds of manipulation check items. But 57% is a high number, and we are not sure which participant has seen the stimulus and which participant has not. We devote some more attention to this ambiguity in the results section.

Finally, we measured the appeal of the messages on a 7-point scale ($M = 5.24$; $SD = 1.64$). Comparing the tailored messages of the experimental group, using an ANOVA, we found no significant differences in the extent to which they found their tailored messages appealing: $F(5, 43) = 0.62$, $p = .68$. The mean score shows the respondents found the stimuli on average quite appealing. The ANOVA indicates that the appeal does not differ significantly between the tailored messages.

Randomization check

There were no significant differences between the control group and the experimental group regarding gender ($t(83) = .28$, $p = .39$, age ($t(84) = -.55$, $p = .29$, and education ($t(83) = .26$, $p = .40$).

Procedural challenges

Studying effects of microtargeting in a field experimental setting comes with five procedural challenges: tailoring the message, matching the message, delivering the message, isolating the effects of the message, and measuring persuasion.

Tailoring the stimulus message

A PMT message should be tailored to at least two characteristics of some subgroup, or individual. These can be any type of perceived characteristic: say, personality or political issue salience. The researcher should get relevant information *before* the stimulus is constructed. This information serves to make inferences about the people in the sample, which forms the basis of crafting a message tailored to the characteristics of the specific participant. There are three viable approaches to solve this problem. First, one could get access to a specific dataset from a data broker (e.g., Hersh, 2015). An upside to this approach is that it is in line with the real-life approaches many political actors take. A downside is that getting access to a dataset can be expensive, and the researcher relies on inferences made by third parties in an opaque data-analysis process.

A second option would be to turn to platforms such as Google and Facebook (e.g., Hager, 2019; even though Hager did not microtarget, the research design can be useful for microtargeting research). These platforms are easy to use, and have an enormous range of inferred user characteristics. The upside and downside mentioned in the first option also apply to this option. A third option is to go into the field and survey participants. Apart from being able to ask participants about the characteristics that are of interest to the researchers, being able to control the data-collection process and the inferences, one could get informed consent for the experiment and the processing of political preferences at the same time. A downside is that this is very time consuming. In this paper's case study, we opted for this last option.

Matching the stimulus message

The researcher should match the right person to the right stimulus. After all, mistargeting would be problematic. Hersh and Schaffner (2013) showed that people who received a message targeted to a characteristic they do not possess, punish the sender of the message. Regular targeting could allow a researcher to geotarget a specific neighborhood with one ad targeted to the entire neighborhood, but microtargeting acknowledges that there are different people who respond differently to different messages living in one neighborhood. Consequently, people living in one neighborhood should receive different, tailored, stimuli. When using the acquired dataset (first option), or the Google/Facebook option (second option), matching could occur via email addresses, or phone numbers. In this field experiment set up however, the researchers simply wrote down the home addresses.¹³

Delivery of PMT message

The stimuli can be delivered online (e.g., via Facebook or Google), which requires an online advertisement auction. Here, the attention of the citizen is sold to the most 'attractive' bid (e.g., highest offer, most engaging content). Oftentimes, there are many actors bidding for the same set of eyes. No advertiser can really be certain in advance of 'winning' the ad auction and consequently showing the ad to *all* desired individuals. A more reliable option is to deliver the stimuli offline, which was done in the current study. Advantages of using a direct mailer are certainty that the stimulus will be delivered in a specific time frame, and the possibility to address the stimulus to a specific person or household (increasing the chances of reaching the right person). A different option is to reach out to Google for information about which groups can be targeted via Google with certainty (see Hager, 2019).

13 This should occur in a GDPR compliant manner. For example, by getting informed consent for the processing of sensitive data and by assigning each participant a code.

Isolating effects of PMT message

Field experiments focusing on political advertising sometimes compare neighborhoods or districts with each other (see Hager, 2019; Gerber, Gimpel, Green, & Shaw, 2011). This approach is not optimal when studying microtargeting, because it is difficult to maintain that individual neighborhoods are distinct homogeneous groups in the setting of a field experiment. It is more likely that there is quite some variation within each neighborhood. Comparing neighborhoods could yield information about regular targeting effects, but not about the effectiveness of microtargeting. In this study, we did not randomly assign neighborhoods to a specific condition, but rather randomly assigned each person to a specific condition.

Measuring persuasion

According to Hillygus and Shields (2009), persuasion occurs when partisans change attitudes or behavior after being contacted by a rival party on a personally relevant issue, on which the partisan disagrees with their own preferred party. This is a somewhat US-centric idea of persuasion, where there are de facto two parties that can flip voters. In the Netherlands, citizens generally have consideration sets of ideology-consonant political parties (e.g. a leftist consideration set, or a Christian set of parties). Voters do not often switch parties between consideration sets, but do more often switch parties within consideration sets (Rekker and Rosema, 2019). In the current study, we do not cross-pressure voters and thus do not measure persuasion as conceptualized by Hillygus and Shields (2009). Rather, we strictly measure changes in vote likelihood.

Results

An OLS regression analysis showed that, for the participants in the experimental group, receiving the stimulus had a significant positive effect on participants' likelihood to vote for the PvdA. While controlling for vote likelihood scores at T1, gender, year of birth and educational background, the PvdA experimental group had significantly higher vote likelihood scores than the control group. Moreover, the younger citizens were less affected by the stimulus (see Table 4.1).

There were 6 participants in the control condition who casted their vote for PvdA. There were 7 participants in the experimental group who casted their vote for PvdA ($t(77) = .45, p = .33$). The increase in vote likelihood did not significantly translate into extra votes for PvdA.

Receiving the stimulus decreased the likelihood to vote for traditional center Christian party CDA as well as for rightist party VVD, but not significantly ($p_{\text{-CDA}} = .28; p_{\text{-VVD}} = .23$; see Appendix D).

The manipulation check showed that 57% of the participants in the experimental condition did not remember seeing the stimulus, 14% did not know for sure and only 29% did remember. The analysis above included all experimental participants. But zooming in on the three groups, using an ANOVA, we find a significant difference on the three groups vote likelihood: $f(2, 48) = 4.75, p = .01$ (see Table 4.2). A Bonferroni post-hoc comparison showed that the participants in the experimental group that did remember seeing the stimulus ($N = 15$) had a significantly higher ($M = 7.87; SD = 1.73$) vote likelihood than those that did not remember ($M = 5.90; SD = 2.43; N = 29$), and those that were not sure whether they saw the stimulus ($M = 5.28; SD = 2.50; N = 7$). These three groups did not score significantly different on vote likelihood in the pre-survey at T1: $F(2, 48) = 2.39, p = .10$.

Using a t-test to compare the group that remember seeing the stimulus and the group that did not remember (ignoring the group that did not know), the difference in mean vote likelihood scores was significant as well: $t(42) = 2.79, p = .004$. The Kolmogorov-Smirnov test ($\chi^2 = .12$) as well as the Shapiro-Wilk test for normality ($W = .91, p = .10$) was non-significant for the group that remembered seeing the stimulus. The same applied to the group that did not remember seeing the stimulus ($\chi^2 = .41; W = .99; p = .95$). This means that we can assume normality and that the t-test is robust. The findings indicate that we may underestimate the effect of the stimulus on the vote likelihood of the recipient.

Turning to *issue salience*, we encountered a ‘ceiling effect’. About 10% of the scores at T1 could not increase further. To overcome this, we calculated a ‘relative issue salience score’ (score of most salient issue at T1/all issue salience scores combined). An OLS regression showed no significant differences in relative issue salience between the experimental group and the control group at T2, while controlling for their relative issue salience score at T1 (see Table 4.3).

Table 4.1 - OLS regression detailing effect stimulus on likelihood to vote for PvdA.

Vote likelihood (T2)	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>	95% CI
Vote likelihood T1	.63	.07	8.91	<.001	.49 – .77
Condition*	1.06	.41	2.62	.01	.26 – 1.87
Gender	.20	.42	.48	.63	-.63 – 1.03
Year of birth	-.04	.02	-2.64	.01	-.07 – -.01
Educational background	.21	.23	.90	.37	-.26 – .69
Constant	84.67	31.41	2.70	.01	22.12 – 147.22

$N = 82$. *Control = 0, experimental = 1; $R^2 = .57$

Table 4.2 - Means (SD) for likelihood to vote PvdA broken down per remembering seeing the stimulus.

Remember seeing manipulation	Do remember (<i>N</i> = 15)	Do not remember (<i>N</i> = 29)	Do not know (<i>N</i> = 7)
Likelihood to vote PvdA	7.87 (1.73) ^a	5.90 (2.42) ^b	5.29 (2.50) ^b

N = 51. Difference between superscript ^a and ^b is significant: *p* = .01

Table 4.3 - OLS regressing detailing effect stimulus on issue salience.

Relative issue salience score (T2)	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>	95% CI
Relative issue salience score T1	.44	.10	4.30	<.001	.24 – .65
Condition*	-.001	.003	-.35	.73	-.01 – .01
Constant	.08	.02	4.93	<.001	.05 – .11

N = 82. *Control = 0, experimental = 1; *R*² = .19

Discussion

In this paper, we conceptualized PMT, and identified challenges to measuring effects of PMT messages on political behavior. Simultaneously, we set out to show how these effects *could* be studied.

Tailoring the stimulus message proved relatively easy. Using round 1 of data collection for this purpose allowed us to quickly and easily match participants to a certain message. An upside of this approach is that you can, in theory, quickly learn anything about the participants: from their issue salience to their personality. A downside is that it is time-consuming and difficult to collect data using door-to-door surveys on a large scale. Moreover, political parties often take a different approach to microtargeting. In reality parties generally rely on inferred data and tailor their messages to perceived voters (see Hersh, 2015). But there is evidence that parties in the Netherlands and in Germany also collect data with door-to-door visits (Dobber et al., 2017; Kruschinski & Haller, 2018).

Similar to Endres (2019), who found Republican messages microtargeted to Democratic partisans increased support for the Republican candidate, this current study finds that a microtargeted message increases citizens' vote likelihood. Different from Endres (2019), who studied voters who were cross-pressured on a congruent issue, this current study focused on citizens' most salient issue. Moreover, where Endres also found a decrease in Democratic partisan support for the Democratic candidate, this current study finds no significant decrease in likelihood to vote for competing parties CDA and VVD.

Different from Hersh and Schaffner (2013), Hersh (2015), Endres (2019), and Haenschen and Jennings (2019), this current study was set in a European, multiparty context. While PMT increased participants' vote likelihood, this did not translate into actual votes for the microtargeting parties. We attribute this to the presence of 16 competing parties. Participants most likely already had a preferred political party (that was not the PvdA), or at least had settled on a set of parties for which they wanted to vote (see Rekker & Rosema, 2019). The chance that an increase in vote likelihood would lead to participants abandoning their preferred party for the PvdA is slim. As such, measuring to what extent an increase in vote likelihood translates into actual votes in a multiparty setting requires many more observations. But any political party should be happy with an increase in vote likelihood. Early in the election cycle such an increase could help a political party to be seen as a viable option. Later in the election cycle, one microtargeted message seems to move the needle, but not enough to win votes.

Participants in the experimental condition who remembered seeing the stimulus held significantly higher vote likelihood scores ($M = 7.87$; $SD = 1.73$) than the participants in the experimental condition that did not remember seeing the stimulus ($M = 5.90$; $SD = 2.42$), while this difference was not there yet at T1. A possible explanation is that a portion of the experimental group really did not see the stimulus. Consequently, the effects found in this study could be underestimated because the experimental group includes untreated participants. A different explanation is that the people who did not remember exposure found the message not personally relevant, causing them to cast aside the message without much thought.

Issue salience

We found no significant effects of microtargeting on issue salience. In part, we can attribute this to our relatively small number of observations: a consequence of our decision to conduct a field experiment rather than an online experiment. It proved very difficult, time consuming and expensive to get the same person to answer a survey twice, and therefore a larger N was unfeasible. Consequently, there could be an effect, we simply were unable to measure it.

Limitations

Apart from sample size, a limitation of our study is the composition of our sample: predominantly higher educated participants, who live in Utrecht, and who are active voters. These are typically not regarded as 'vulnerable' citizens, but the PMT message still had a modest effect on them. It would be interesting to study whether a PMT message affects less active, less educated citizens differently.

Another limitation is that participants received only one stimulus. Sending more stimuli would have increased the chance that more people saw the stimulus. Moreover, it is rare to find a relatively long-lasting effect in a field-experimental setting,

research from Gerber, Gimpel, Green and Shaw (2011), and Hill, Lo, Vavreck, and Zaller (2013) shows. But this research focuses on mass communication, rather than personalized communication. Potentially an individualized approach has the capacity for longer-lasting effects (see, e.g., Broockman and Kalla, 2016). In any case, this study's randomized experimental set-up and the difference in vote likelihood scores within the experimental condition, between those who remember seeing the stimulus and those who do not or are unsure, both provide reassurance that this relatively long-lasting effect is not due to chance.

Future research

Future research should not send only one stimulus, but should actively retarget over a longer period of time. Since digital intermediaries such as Facebook and Google provide valuable and easy-to-use infrastructure for large-scale microtargeting, future research should focus on the effectiveness of political microtargeting on these platforms. Ethical research standards are rather important, especially when studying this on a large scale and during an actual election.

This study was the first European-based field experiment to examine the effects of political microtargeting on vote likelihood in multiparty setting, and as such, adds novel empirical evidence to the field of political microtargeting research. The findings from this study can help better understand PMT and help come with ways to limit PMT's threats, while maximizing its potential.