Routledge
Taylor & Francis Group

Economic and Political Studies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/reps20

Exposure to conspiracy theories in the lab

Loukas Balafoutas , Alexander Libman , Vasileios Selamis & Björn Vollan

To cite this article: Loukas Balafoutas , Alexander Libman , Vasileios Selamis & Björn Vollan (2021): Exposure to conspiracy theories in the lab, Economic and Political Studies, DOI: 10.1080/20954816.2020.1818930

To link to this article: https://doi.org/10.1080/20954816.2020.1818930

Published online: 11 Feb 2021.

ECONOMIC AND POLITICAL STUDIES
https://doi.org/10.1080/20954816.2020.1818930

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Exposure to conspiracy theories in the lab

Loukas Balafoutas[a], Alexander Libman[b], Vasileios Selamis[a] and Björn Vollan[c]

[a]Department of Public Finance, University of Innsbruck, Innsbruck, Austria; [b]Institute of Sociology, Ludwig Maximilian University of Munich, Munich, Germany; [c]School of Business and Economics, Philipps-Universität Marburg, Marburg, Germany

ABSTRACT

Conspiracy theories are widespread in the modern information era. Being exposed to conspiracy theories may affect behaviour, for example, by spreading mistrust among people and within organisations, even if it does not necessarily generate widespread beliefs in the conspiracy narrative. Our paper investigates the effect of exposure to conspiracy theories on strategic sophistication. We present evidence from a laboratory experiment, in which we prime half of our participants with exposure to a conspiracy theory. We find that such exposure leads to increased strategic sophistication. Using a causal mediation analysis we confirm that the effect on sophistication arises independently of whether people believe in the content or not.

ARTICLE HISTORY
Received 19 August 2019
Accepted 11 September 2019

KEYWORDS
Conspiracy theory; k-level reasoning; trust; priming

Introduction

Conspiracy theories flourish in the modern information environment; watching TV or surfing the internet almost inevitably leads to exposure to various conspiracy ideas, providing simple explanations for an increasingly complex world (Radnitz and Underwood 2017). Conspiracy theories spread with ease through online social networks (Del Vicario et al. 2016). These theories are widely present in the comment sections of media websites (Wood and Douglas 2013), and are even the most likely outcome of searches on certain topics on the internet (Kata 2010). Thus, one can hardly imagine an individual in the modern world who has never encountered a conspiracy-based interpretation of social phenomena. Experimental and survey data from psychology link the belief in conspiracies to a number of phenomena with substantial costs for individuals and for the society in general: lower levels of voting, donating and volunteering (Uscinski and Parent 2014), stronger support of fringe political groups (Sunstein and Vermeule 2009) and self-damaging behaviour, for example, rejection of medical treatments (Kalichman 2009), medical tests (Ford et al. 2013) and inoculations

CONTACT Alexander Libman alexander.libman@soziologie.uni-muenchen.de Institute of Sociology, Ludwig Maximilian University of Munich, Munich, Germany

(Jolley and Douglas 2014b). Conspiracies refuting climate change reduce the willingness of individuals to engage in environmentally friendly behaviour. Beliefs in conspiracies within an organisation reduce employee loyalty and increase personnel turnover (which is likely to have a negative impact on organisational performance). Conspiracy beliefs can even trigger wars and repressions (e.g. Rubin 2011).[1] How does, however, exposure to conspiracy theories change human behaviour even if people do not start believing in them? Does this exposure matter for the decisions made by individuals outside the context in which a conspiracy theory was formulated?

A conspiracy theory can be defined as a narrative about 'hidden, malevolent groups secretly perpetuating political plots and social calamities to further their own nefarious goals' (Oliver and Wood 2014, 952). Thus, an explanation of the social phenomena based on conspiracy theories should combine three elements: belief in (a) intentional and (b) hidden activities of certain forces, (c) for the purpose of harming the common good and an individual in particular. Many conspiracy theories are obviously implausible (which does not stop people from believing in them),[2] but there are also cases when the conspiracy-based explanations are actually true, like in numerous military coups or conspiracies against the incumbent leaders throughout the human history (Gray 2010). Still, in many cases, conspiracy theories are part of or linked to the 'fake news' phenomenon, or the growing spread of misinformation and misinterpretation of events observed in many countries (Allcott and Gentzkow 2017).[3] Overall, the spread of conspiracy theories resembles a paranoid way of interpreting events and is believed to be a major challenge for many societies (Oliver and Wood 2014). Our main focus in this study is not on the epistemological value of conspiracy theories; we rather view them as a particular way of explaining the world and want to understand how people react upon encountering them. While conspiracy theories have been addressed in psychology, political science and philosophy, presenting correlational evidence from public opinion data or using survey experiments, we are – to our knowledge – among the first to study this phenomenon using the toolbox of experimental economics with a randomised assignment of treatment.[4]

Our paper devises an incentivised laboratory experiment with random assignment of a conspiracy priming, in order to examine how exposure to conspiracy theories affects strategic sophistication, defined as 'the extent to which players' behaviour reflects attempts to predict others' decisions' (Costa-Gomes, Crawford, and Broseta 2001, 1193). Our focus is not on immediate consequences of beliefs in the content of conspiracy theories (e.g. those believing in vaccinations being used by nefarious forces may be less willing to subject themselves to inoculations), but rather on how subjects experiencing an encounter with conspiracy theories change their behaviour in more generic situations, possibly not covered by conspiracy theories (e.g. standardised games played in economic experiments). Evidence of such behavioural changes would indicate a societal impact of conspiracy theories even beyond polarisation (McHoskey 1995), social disengagement (Jolley and Douglas 2014a), or disbeliefs in facts (Lewandowsky, Oberauer, and Gignac 2013). There is substantial evidence that strategic sophistication, while influenced by cognitive abilities and education, also varies depending on the social situation an individual finds his/herself in or (in experiments) on the design of the game (Georganas, Healy, and Weber 2015; Penczynski 2016; Allred, Duffy, and

Smith 2016). Hence, strategic sophistication appears to be malleable (even) in the short run to some extent.

We find that being exposed to a conspiracy theory changes the behaviour of subjects in our experiment. In particular, the conspiracy treatment increases strategic sophistication among the subjects, that is, it makes them think more carefully about the reaction of other subjects they interact with. We link this finding to several features of conspiracy narratives: their emphasis on intentional actions upon carefully designed plans and, somewhat paradoxically, their self-representation as cases of critical independent thinking. At the same time, the effect is not driven by the increasing *belief* in conspiracy theories (although we document an increase in beliefs as a consequence of the conspiracy treatment as well); rather, it is *exposure* as such, which for most subjects does not elicit changes in beliefs in conspiracy theories but is enough to activate higher strategic sophistication.

The remaining part of the paper is organised as follows. The next section explains the arguments behind our expectations on how exposure to conspiracy theories should affect behaviour. The following two sections present our experimental design and report the results. Then, the paper extends the main analysis by studying another possible effect of exposure to conspiracy thinking (trust levels). The last section concludes.

Strategic sophistication and exposure to conspiracy theories

While the research on conspiracy theories in social sciences and psychology has been growing (Douglas et al. 2019), the link between strategic sophistication and exposure to conspiracy theories has received little attention in the scholarly literature. In order to formulate our expectations concerning this variable, we start by refining the discussion about the possible effects of exposure to conspiracy narratives. While a certain subgroup of subjects is likely to be immediately convinced by the validity of a given story and embraces the corresponding conspiracy theory, for many subjects such an increase in beliefs is likely to occur only after prolonged exposure (and possibly conditional on their prior beliefs or personalities[5]) or not occur at all. A randomised experiment is an appropriate tool for investigating these short-term effects of encounters with conspiracy narratives given the short-term exposure to a conspiracy narrative. The question is, however, whether this short-term exposure matters for human behaviour. We hypothesise that being exposed to conspiracy theories even for a brief period of time could, among other things, increase the level of strategic sophistication in decision-making. There are several potential reasons for this effect.

First, conspiracy theories commonly relate to *intent:* malicious forces are consciously trying to induce harm (rather than harm other individuals by accident or as a by-product of their action). In fact, a typical feature of many conspiracy narratives is a focus on the possible motives of the conspirators rather than on actual empirical facts. Facts can be manipulated (and, in any consistent conspiracy narrative, are seen as being manipulated); thus, the only way to uncover the truth is to ask the 'cui bono' question – who is benefitting from it (Blaskiewicz 2013; Bost and Prunier 2013)? For those individuals who fully embrace a conspiracy narrative, the answer to this question is also given: it is part of the narrative, which tells exactly who is benefitting from the

conspiracy, why the conspiracy is perpetuated and how the facts should be interpreted. Individuals who are subjected only to a short-term encounter with conspiracy thinking might not yet share the argument of the conspiracy narrative, but still, start thinking more about other people's intentions.

Second, somewhat paradoxically, many conspiracy theories eagerly present themselves precisely as outcomes of critical unbiased thinking, free from the boundaries and stereotypes of the mainstream (Brotherton 2015). Texts and movies introducing the neophytes to conspiracy thinking are structured as those calling for open discussion of the predominant point of view, for questioning every argument and every statement with pure logic, rejecting the rigorous boundaries of conventional knowledge. A brief encounter with the conspiracy narrative could thus encourage individuals to invest larger effort into trying to understand the phenomena they encounter, including other people's actions. This could increase strategic sophistication. Again, for individuals who unconditionally believe in a conspiracy narrative, this effect is unlikely to be present: the narrative dictates a specific way of how one has to interpret the events. But in the case of short-term encounters, individuals may not have yet internalised the narrative but still start thinking about the actions of the others more carefully.

Third, conspiracy narratives also assign a high level of strategic sophistication to the conspirators, who carefully think about how the public reacts to their actions and are able to manipulate the society to their favour. Subjects' strategic sophistication can be endogenous to their beliefs about how 'deep' the reasoning of those they interact with is (Alaoui and Penta 2016). For our study, one could hypothesise that exposure to conspiracy narratives could make subjects more sensitive to how well calculated human action could be, in turn incurring higher strategic sophistication.[6]

While the three arguments presented above presuppose a positive link between exposure to conspiracy narratives and strategic sophistication, it is also possible to suggest an opposite hypothesis: conspiracy narratives could decrease strategic sophistication. First, even short-term encounters with conspiracy theories, especially if the subjects have no previous experience of those, can distract people's attention, directing it away from thinking about the interaction with other individuals in an everyday situation (or playing the game in case of our experimental study) towards thinking about the unusual narrative. In short, conspiracy theories could overload an individual's cognitive capacity. Second, contrary to our argument presented above, even a short-term encounter with conspiracy theories could make people more naive, essentially accepting any explanation or argument they are presented with. This would again lead to a decline in strategic sophistication.

The conjecture that even short-term exposure to conspiracy theories by itself can change human behaviour has potentially important implications, given the frequent appearance of conspiracy theories in the media and the internet. Even if people are not convinced by these reports, it may already be enough to change how they perceive the world and how they act. The hypothetical effect of the conspiracy priming on strategic sophistication is not necessarily triggered by an increasing belief in conspiracy theories since an encounter with conspiracy theories makes individuals more likely to carefully think about other people's actions and motives without necessarily embracing a

particular narrative explaining these motives. In the next section, we present an experiment devised to test our hypothesis.

Experiment

The experiment was conducted at the EconLab of the University of Innsbruck. Subjects (n = 144) were randomly assigned either to a control or a treatment group, which was primed to induce conspiracy thinking. The treatment group watched a three-minute newscast presenting a sympathetic account of the moon landing conspiracy (the idea that the US astronauts never went to the moon and that the moon landing was filmed on the earth[7]). The experimental group watched an equally long documentary on the space shuttle program.[8] Both movies were thus devoted to the same general topic (space exploration). Both were taken from the US television, which should be less familiar to our subjects (given that the experiment was conducted in Austria), who are also unlikely to experience any specific biases with respect to particular channels or even familiarity with the typical narratives of the US broadcast.

To increase attention to the movies and thereby the likelihood of successful priming, at the end of the experiment we asked subjects to respond to six factual questions on the content of the movies, rewarding correct answers with €0.5 each. There was no significant difference in the number of questions answered correctly between the treatment and the control group, confirming that the movies were of a roughly similar level of difficulty. This is important if our goal is to analyse the implications of the conspiracy treatment on strategic sophistication. As mentioned, differences in observed strategic sophistication across subjects can in principle be driven by two factors: the cognitive abilities of the subjects and their willingness to invest effort in thinking about the possible behaviour of the opponent. Cognitive limitations are primarily individual-specific and thus, given the random assignment of treatment, should not affect our results. While one could argue that encountering a more complex narrative could simply force individuals to reason more intensively (regardless of the content of this narrative), this is not the case in our data. Our findings also refute the alternative hypothesis that the effects we observe in our study are driven by the cognitive overload caused by the conspiracy priming. For our subjects, both the conspiracy priming and the control treatment were equally challenging; hence, the reported effects cannot be driven by the experimental group experiencing cognitive overload.

The main part of the experiment was the money request game (Arad and Rubinstein 2012), a variation of the beauty contest game intended to elicit strategic sophistication.[9] In this game, players were divided into pairs. Each player was asked to bid an integer amount between €5 and €14. The participant with the smaller bid received the amount of this bid plus €10; the participant with the larger bid received the amount of the bid. In the case of a tie, both participants received exactly their bid. The best response to any bid of the other participant larger than €5 in this game is to bid exactly one euro less. If the other participant bids €5, the best response is to bid €14. We refer to the extent of the strategic sophistication of the participant as k-level reasoning (Bosch-Domènech et al. 2002).[10] Participants of level-0 reasoning would bid €14 (not taking the opponent's response into account). Level-1 reasoning implies a bid

Table 1. Differences between the experimental and control groups and corresponding *t*-tests.

	Primed group	Control group	Δ
Priming cheques			
Moon landing	2.736 (0.148)	1.972 (0.140)	0.764*** (0.204)
Princess Diana	1.958 (0.123)	1.861 (0.118)	0.097 (0.170)
Pharmaceuticals	3.361 (0.146)	3.125 (0.149)	0.236 (0.209)
Roswell	1.444 (0.103)	1.361 (0.089)	0.083 (0.136)
9/11	2.472 (0.163)	2.333 (0.147)	0.139 (0.219)
Outcome			
k-level reasoning	1.721 (0.115)	1.281 (0.136)	0.440** (0.179)

Notes: The table presents mean values for all variables, with standard deviations in parentheses. ***Significant at 1% and **5% levels. The variables in the first five rows are defined as the extent to which participants indicate that they find the corresponding conspiracy theory plausible, on a scale from 1 ('not plausible at all') to 5 ('entirely plausible'). The five theories are the following: (1) The moon landing by Neil Armstrong in 1969 was a hoax. (2) The death of Princess Diana was set up by British intelligence on behalf of the Royal Family. (3) Big pharmaceutical companies often do not disclose cures that have already been found (e.g. for AIDS) in order to keep earning millions of dollars. (4) UFOs landed in Area 51 in Roswell, New Mexico in 1947 and the US government has lied about it ever since. (5) The 9/11 attacks on the World Trade Centre were orchestrated by the US government.

of €13, assuming that the opponent will naively bid €14. Level-2 reasoning implies that the opponent will bid one euro less than the bid of a level-1 participant (trying to outbid him) and hence the bid would be €12, and so on. In addition to eliciting bids, we asked subjects to explain their choice. If an explanation revealed that a participant made a bid in line with the logic above, we attributed a particular level of reasoning to him/her (zero, one, two, and so forth). If an explanation revealed a purely irrational choice (e.g. random draw) or a lack of understanding of the game, the observation was dropped.[11] We refrained from eliciting subjects' beliefs about their partners' choices because the very act of doing so could have a direct additional effect on strategic sophistication (e.g. forcing subjects to think more about how their partners would act).

As a manipulation check of our priming method, we asked participants at the end of the experiment to respond to a questionnaire containing, among others, a number of items on some widely known conspiracy theories. Respondents had to evaluate the extent to which they agreed with these statements on a five-point Likert scale. Einstein and Glick (2015) argue that asking questions about whether people actually believe in conspiracies can change the effects of the conspiracy priming; we avoided this problem by administering the questionnaire after both the money request game and the trust game (see the section 'Extension: trust game') were played. The list of questions is shown in the footnote to Table 1. In addition, we collected information on personality traits by administering a ten-item version of the Big Five inventory (Gosling, Rentfrow, and Swann 2003). We observe no significant differences between the treatment and control groups in any of the five personality traits ($p > 0.331$, Mann–Whitney U tests).

Results

Effects of the conspiracy priming

Table 1 compares subjects in the treatment and control groups along several dimensions. First, we look at their responses to the question of whether they believed the moon landing was actually staged (in order to confirm that the priming resulted in more widespread conspiracy thinking), as well as to questions related to other

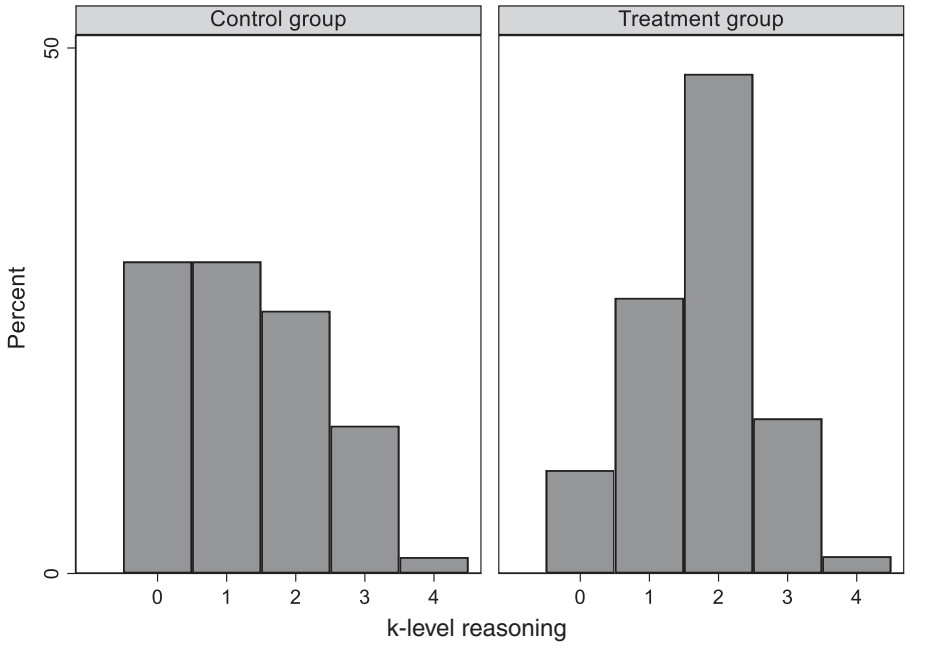

Figure 1. K-level reasoning of subjects in the treatment and control groups.

conspiracy theories. Second, we look at their k-level of reasoning. Subjects in the treatment group were indeed more likely to say that they believed in the moon landing conspiracy, but beliefs in all other conspiracy items do not vary across treatments. This supports the interpretation of the results of our experiment as causally driven by our priming rather than an outcome of unobserved heterogeneity of subjects. More importantly, the conspiracy priming does have an impact on behaviour: we find that subjects in the primed group exhibit on average higher levels of strategic sophistication than those in the control group (1.72 vs. 1.28, $p < 0.05$, t-test). Thus, in line with our hypothesis, being exposed to conspiracy theories seems to make subjects act more strategically and also view other people they interact with as more sophisticated.

We perform two further tests on our data. First, Figure 1 plots the distribution of subjects according to k-level reasoning in the treatment and control groups. One can see that differences between subjects in the two groups extend beyond the difference in the mean responses. While in the control group the most frequent k-level reasoning levels are $k = 0$ and $k = 1$ (19 subjects each), in the treatment group only six subjects exhibit the k-level reasoning $k = 0$, with the median being $k = 2$ (28 subjects). A Kolmogorov–Smirnov test for equality of distributions weakly rejects the hypothesis that k-level reasoning across treatment and control groups has the same distribution function ($p = 0.067$). The most successful strategy in playing the game in the experimental group is therefore to bid €11 (this would outbid the median opponent), while in the control group it is €12 or even €13. Second, we test whether the number of irrational choices or cases when subjects did not understand the game (i.e. observations, which were dropped from the analysis) was different across the experimental and control groups. Overall, 11% in the control group and 15% in the treatment group were dropped because of the irrational justification of the choice; these shares are not significantly different from each other ($p = 0.464$). This again

Table 2. Equilibrium and choices of the subjects.

	5	6	7	8	9	10	11	12	13	14
Equilibrium (%)	0	0	0	0	0	0	40	30	20	10
Treatment group, actual choices (%)	0	0	3	1	3	1	14	44	25	8
Control group, actual choices (%)	1	1	0	1	1	1	17	24	25	28

reassures us about our interpretation of the findings: the frequency of irrational choices is balanced across treatments.

Table 2 reports the symmetric Nash equilibrium distribution of the game along with the actual distribution of choices made by subjects in the experimental and control groups.[12] Both treatment and control groups strongly deviate from the equilibrium distribution. In the treatment group, we observe the prevalence of the choice of bidding €12, that is, level-2 reasoning.[13] In the control group, level-0 reasoning is the most prominently present in our data.[14] From this point of view, we need to add an important caveat to our analysis. Li and Rong (2018) show that the results of the money request game are significantly biased by risk aversion. Therefore, what appears to be an increase in k-level reasoning could actually be driven by the greater willingness to risk in order to win the game. The conspiracy priming could hypothetically produce this result: those who are not convinced could be more eager to risk to show that they are more intelligent than the conspiracy thinkers, and those who are convinced could be more eager to risk to show that they are more intelligent than those who accept the official narrative. The problem with this interpretation is that our subjects do not know anything about the conspiracy beliefs of other subjects so that there is no reason for them to actually expect other participants of the experiment to be more or less intelligent (there were no references to this type of logic in the explanations of individual choices we collect). However, we mention this interpretation of our results for completeness.

In Appendices B and C we report two extensions of our analysis. First, while so far we have reported average treatment effects, in Appendix B we search for heterogeneous treatment effects by gender, age, and religiosity. We report evidence of a significant positive association between the strength of religious beliefs and the strength of the priming effects on strategic sophistication. Religiosity is an interesting variable in the context of our study since there is a substantial discussion about its relation to conspiracy thinking (Franks, Bangerter, and Bauer 2013; Van der Tempel and Alcock 2015; Jasinskaja, Lahti, and Jetten 2019). Religiosity could be a substitute to conspiracy thinking, presenting an alternative set of beliefs explaining the observed regularities in the world.[15] It can, however, also encourage conspiracy thinking and strengthen its effects.[16] Second, to better understand the behaviour of individuals subjected to the conspiracy priming, we conducted a survey where we asked the respondents about the feelings or impressions they experienced while watching the conspiracy video (Appendix C). While this survey is unable to causally identify the effect of conspiracy thinking and the self-assessment of the experienced impressions can be difficult to analyse, we still find the results of the survey an interesting extension to our findings since they offer insights into the possible mechanisms explaining how the conspiracy priming is perceived. The survey gives some indication that many subjects interpreted the conspiracy video as a factor making them think more deeply about the behaviour of others.

Explaining the treatment effect: beliefs in conspiracy?

Our hypothesis regarding the effect to the conspiracy priming on strategic sophistication relies on the assumption that individuals, after encountering a conspiracy theory narrative, start thinking more carefully about the intentions of others and are likely to change their behaviour even if they do not immediately change their beliefs. Our experimental results reported above show that priming has an effect on both levels: individuals change their beliefs to some extent, and they increase their level of strategic sophistication. This begs the question of identifying the driving forces behind the observed treatment effect: is this effect driven by the change in beliefs, and if so, to which extent? Or does the conspiracy priming affect behaviour directly, and independently of beliefs in conspiracies? To explicitly address this question, we use the toolbox of the causal mediation analysis (CME) (Imai et al. 2011).

The CME allows us to decompose the effect of treatment into a direct and an indirect effect, working through a mediator variable – another characteristic, which is affected by the treatment and in turn affects the outcome. The mediator in our case is the belief in the moon landing conspiracy; the treatment is whether one was in the primed or the control group; and the outcome is the level of strategic sophistication. For each of the effects, one is able to test whether it is significantly different from zero; thus, we will be able to understand the extent to which the increase in strategic sophistication is mediated by the increase in beliefs in the moon landing conspiracy. The algorithm applied to estimate the effects is the following. First, it estimates two linear models: the mediator model, where the mediator (beliefs in moon landing conspiracy) is regressed on the treatment; and the outcome model, where the outcome (k-level thinking) is regressed on the mediator and treatment (detailed regression output for both models is reported in Appendix D). Second, it simulates model parameters from their distribution and, for each draw, computes the value of the mediator, the value of the outcome for this value of the mediator, and the direct, indirect and total effects. We apply the algorithm with 1,000 parameter draws and use the implementation of the algorithm in Stata by Hicks and Tingley (2011).

The results of the CME reported in Table 3 are unequivocal. The indirect effect (i.e. the portion of the effect of the treatment mediated through the beliefs of subjects) is very small in size and not significantly different from zero at the 5% level.[17] The total effect is positive and significant (indicating that the level of strategic sophistication indeed goes up when subjects are exposed to the conspiracy priming), and is almost entirely driven by the direct effect (the effect of treatment *per se*, independently of the effect on beliefs). Conscious beliefs in conspiracies are thus not necessary for conspiracy thinking exposure to affect behaviour. Simple exposure to the conspiracy narrative,

Table 3. Causal mediation analysis, disaggregating direct and indirect treatment effects.

	Effect	(95% confidence interval)	
Indirect effect	−0.002	−0.132	0.129
Direct effect	0.447**	0.079	0.811
Total effect	0.445**	0.081	0.805

Note: **Significant at the 5% level.

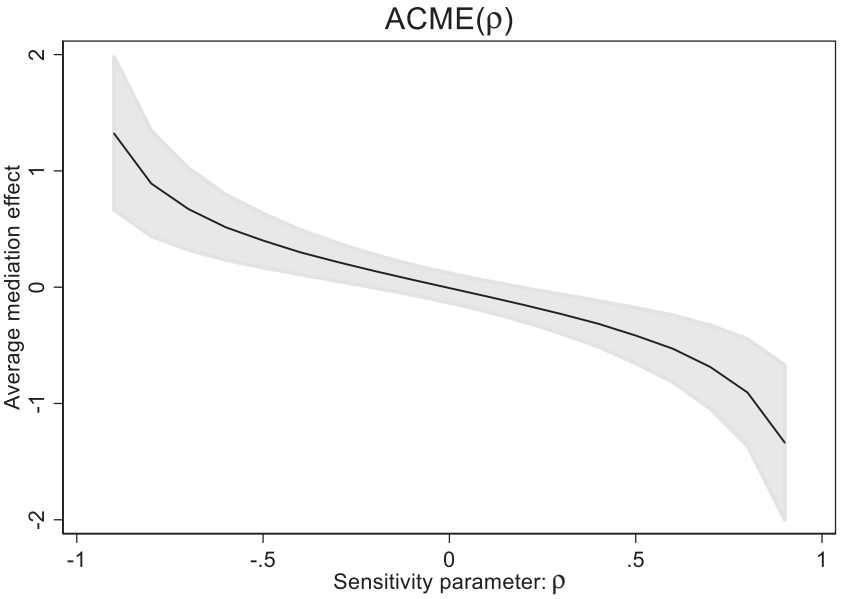

Figure 2. Indirect effect of the treatment on strategic sophistication for various values of the sensitivity parameter ρ.

even if it does not change the beliefs of the subjects, increases their level of strategic sophistication.

The causal mediation analysis relies on the sequential ignorability assumption.[18] Hence, in line with the literature, we also perform a sensitivity test to the violation of this assumption, which happens if there is an unobserved characteristic affecting the relationship between the mediator variable and the outcome variable (e.g. a personal characteristic affecting the belief in the moon landing conspiracy and k-level thinking). Figure 2 shows how the estimated indirect effect would change depending on the value of the parameter ρ, measuring the correlation in error terms of the outcome and mediator models. Under the sequential ignorability assumption, $\rho = 0$. Figure 2 shows that, in order for the indirect effect to be significantly different from zero, ρ should reach the value of about 0.25 or more in absolute terms, which would imply the existence of a rather strong confounder. This reassures us about the validity of the results of the CME. Note that, while we present the sensitivity analysis for completeness, the main analysis of the paper is based on the random assignment of treatment and hence we should not be concerned about any confounding unobserved characteristics.

Extension: trust game

As an extension of our analysis, we perform a different test, eliciting the effects below conspiracy priming on trust among subjects. This test is motivated by two considerations. First, one of the widely reported features of conspiracy beliefs in the existing research is that they are associated with low trust (Goertzel 1994, Abalakina-Paap et al. 1999; Brotherton, French, and Pickering 2013). Mistrust is even sometimes seen as the fundamental feature underlying conspiracy thinking as such (Aupers 2012).[19]

Therefore it appears to be plausible to test the effect of the conspiracy priming on trust in our experiment as well. Second, the analysis of the effect of the conspiracy priming on trust can also help us verify the claim made in the previous sections. Both empirical literature and theoretical arguments clearly link the decline of trust to beliefs in conspiracies. There is no reason to expect that short-term *exposure* to conspiracy narratives if it did not lead to the formation of beliefs in conspiracies or at least made people more inclined to believe in them, should have produced an effect on trust. Thus, if we observe no effect of the conspiracy priming on trust, we can be to a larger extent reassured that changes in beliefs are not the main factor driving the behavioural change we documented in the subsection 'Effects of the conspiracy priming'.

To perform the test in question, we let our subjects play a trust game (Berg, Dickhaut, and McCabe 1995) intended to check whether the conspiracy priming leads to an erosion of trust (this game and the money request game were played in a random order). In this game, players were divided into pairs. In each pair, both players received an endowment of €5. Then, one of the players (A) could decide to invest (part of) this endowment. The invested amount was multiplied by three and passed on to player B, who could then transfer some of the money back to player A – but was not obliged to do so. Larger invested amounts by A in this game correspond to higher levels of trust. Our results suggest that the conspiracy priming did not have any effect on the level of trust, as measured by first-mover transfers in the trust game (3.28 vs. 2.94, $p = 0.481$, t-test).[20] If we exclude all observations that were also excluded in the analysis of k-level reasoning (irrational justification of the decisions made), we still find no significant difference between the treatment and control groups in the trust game (3.14 vs. 3.00, $p = 0.785$, t-test). We also find no correlation between the money transfer and k-level reasoning among our subjects (Spearman's $\rho = -0.0114$ for the entire sample and 0.1078 in the treatment, $p = 0.930$, $p = 0.585$, respectively).

The result should be treated with caution for a number of reasons. First, because we can extract only one observation from each pair of players (the one deciding to invest money), the number of observations in this analysis is almost half as large as in the analysis of the money request game (72 vs. 125). This reduces the power of the test, and it can lead to different outcomes in the analysis of the trust game than in the money request game.[21] Second, we have no information on the beliefs of the subjects about the partners' behaviour, which limits our ability to interpret the results of the trust game. Third, the results of the trust game can be affected by framing (see Appendix A) and, in particular, by the fact that the action of the subject A is referred to as an 'investment'. Thus, we urge to consider these results as auxiliary findings to the main results of the paper.

Conclusions

The widespread presence of conspiracy theories in the public discourse, media and the internet poses the question as to how exposure to these theories affects human behaviour. The goal of this paper is to study certain effects of conspiracy theories using an incentivised laboratory experiment, and to test whether exposure to conspiracy theories changes human behaviour beyond the specific narrative of a given theory. Our design

allows us to understand how conspiracy theories affect behaviour after a brief exposure, which can change the actions of individuals without necessarily convincing them of a theory's validity. We find that conspiracy theories affect human behaviour, by increasing the level of strategic sophistication. The effect of conspiracy exposure on k-level reasoning does not require people to actually start believing in conspiracies: the causal mediation analysis has shown that simple contact with such theories may be enough to change behaviour. We hypothesise that the effect of exposure to conspiracy narratives can be driven by three factors: the focus on intent in the conspiracy theories; the self-representation of conspiracy theories as critical accounts of an official narrative; and the high level of strategic manipulation and planning assigned to conspirators in the conspiracy theories. Our design does not allow us to identify which of these factors specifically are triggered by the exposure to conspiracy theories, but it is plausible that all of them operate jointly in producing the effect we observed in the study.

As a caveat, we note that all of our subjects were university students. While the reliance on student samples is very widespread in economic experiments, we have to point out its limitations for the generalisation of our findings. To begin with, for groups with a lower level of education, conspiracy theories could, first, appear to be more convincing – that is, more subjects might actually believe that the narrative is true (Van Prooijen 2017). Alternatively, a cognitive overload from encountering conspiracy theories becomes more probable. In addition, other groups of subjects could simply pay less attention to the conspiracy narrative than students, and in this case the effect of our priming should be weaker. As always, extending our study to other groups of subjects constitutes a promising avenue of further research.

Importantly, our study does *not* suggest that conspiracy theories have a beneficial effect on human behaviour or on society as a whole: the harmful effects of these beliefs are very well documented in the literature (see the introduction to this paper).[22] It is important to keep in mind that the observed changes in behaviour are most likely driven by the *short-term* exposure to conspiracy theories and are not associated with people believing in conspiracies. For the true believers, the results could be entirely different. While in the experiment we artificially limited the encounter of our subjects with a conspiracy narrative to one brief episode, in the real world the exposure to conspiracy theories and fake news could be much more prolonged and thus make more people believe in such theories. At the same time, our study shows that one should also avoid an alarmist sentiment where every encounter with a conspiracy narrative in the media discourse or in the mass culture is seen as an immediate risk.[23] At least for some audiences (like the well-educated subjects of our experiment), occasional discussion of conspiracy theories may stimulate a higher strategic sophistication of reasoning.[24] From a policy perspective, the question is then whether and how it is possible to prevent people from *over-exposure* to conspiracy theories in the modern free media environment; finding solutions to this challenge remains beyond the scope of the study.

Notes

1. For a review of the literature on the consequences of conspiracy thinking see Douglas and Sutton (2018).

2. In a survey of the US public conducted in 2014, 4% of the respondents claimed to believe in the conspiracy of reptiloids (lizard-like aliens) ruling the world, with another 7% being unsure about it, see Public Policy Polling (2013).
3. Conspiracy theories are particularly suitable for the purpose of spreading or supporting fake news, because they are by design difficult to falsify – any evidence to the opposite can be seen as part of the effort of the conspirators to hide the truth.
4. Libman and Vollan (2019) use an experimental game to study conspiracy thinking in Russia and China; but they do not apply a randomised assignment of treatment and infer beliefs in conspiracies from a survey.
5. Research on the effects of propaganda, for example, shows that the same content generates opposite behavioural effects depending on what people believed in in the first place (Peisakhin and Rozenas 2018).
6. Importantly, all the effects we describe are not driven by the fact that the subject exposed to a conspiracy narrative interprets other experimental subjects s/he interacts with as *hostile*. In the conspiracy narrative, 'hostility' is assigned to the malicious forces, which instigate a conspiracy. A brief encounter with the narrative is extremely unlikely to make the subject believe that other participants of the game are hostile. Rather, it triggers more careful reasoning about other subjects' actions and intentions.
7. See https://en.wikipedia.org/wiki/Moon_landing_conspiracy_theories.
8. The two short movies are available from the authors upon request.
9. The game was computerised using z-tree (Fischbacher 2007). For experimental instructions (translated from German) see Appendix A.
10. K-level models in economics are used to describe the hierarchy of strategic sophistication, 'smartness', or 'naivety' of subjects. Level-0 actors choose their action naively and without any regard to the choice of other actors; level-1 actors assume that the population consists of level-0 actors and respond to the behaviour of these actors; level-2 actors assume that the population consists of level-1 actors; and so on (see Stahl 1993; Stahl and Wilson 1994).
11. In one case, the individual chose 11, but gave a justification for the choice clearly consistent with level-1 reasoning (the individual referred to Example 1 in Appendix A and explained that s/he decided to bid one euro less to gain €10). We assign this individual to level-1 reasoning as well. The results of Table 1 hold if one excludes this observation from the analysis.
12. Note that, while the game we use is substantially similar to Arad and Rubinstein (2012), we use somewhat different choice options and payoffs.
13. Arad and Rubinstein show the prevalence of the level-2 and level-3 reasoning among the subjects of their experiment.
14. This result has to be highlighted, since we actually observe a higher prevalence of the level-0 reasoning than in case of Arad and Rubinstein (2012); this may be the case, for example, because of a different payoff structure, or differences in the degree of risk aversion of our subjects. Unfortunately, we collect no information regarding risk attitudes.
15. Similarly to conspiracy theories, it rejects randomness and assumes that observed phenomena are an outcome of a hidden intent – for example, God's will – making beliefs in conspiracies unnecessary.
16. This can be the case, for example, if religiosity goes hand in hand with a critical attitude towards science.
17. In fact, if we correlate the responses to the question concerning beliefs in the moon landing conspiracy and the level of strategic sophistication, we find no significant correlation between these two variables (in the primed group the correlation coefficient is 0.082, and in the control group it is −0.089).
18. This assumption implies that '(a) conditional on the observed pretreatment covariates, the treatment is independent of all potential values of the outcome and mediating

variables, and (b) the observed mediator is independent of all potential outcomes given the observed treatment and pretreatment covariates' (Imai, Keele, and Tingley 2010, 310).

19. There are several reasons for this correlation. First, conspiracy thinking, ultimately, is about hidden forces at work aiming to harm an individual and manipulating others for their benefit, which clearly runs contrary to trusting people (in particular people one does not know). Second, conspiracy narratives always portrait the dominant scientific or political interpretation of events as a lie, which should also contribute to a decrease in trust. Conspiracy theories frequently directly encourage their disciples not to trust anybody (except the fellow conspiracy believers).
20. In addition to trust, our data allow us to compare trustworthiness across groups, measured by means of back-transfers of second movers in the trust game. However, trustworthiness is not part of our hypotheses and was only elicited in order to have a balanced between-subjects design in the trust game. Thus, unsurprisingly, no treatment differences are found in this dimension (4.39 vs. 3.64, $p=0.42$, t-test).
21. A simple way of doubling the size of the sample would have been to let subjects play both roles in the trust game. However, we refrain from doing so since this approach could bias decisions. Regardless of whether subjects' payoffs are determined by how they played in both roles or (randomly) in one of the two roles, conditional reciprocity could arise if the expectations about how others will treat a subject in a particular role influence how this subject oneself behaves in this role. Previous literature has indeed shown that role uncertainty can have a strong impact on behaviour (see Iriberri and Rey-Biel 2011).
22. Furthermore, from a normative point of view, an increase in strategic sophistication is not necessarily associated with a welfare improvement.
23. A related result by Nera, Pantazi, and Klein (2018) shows that being exposed to conspiracy fiction does not increase beliefs in conspiracy theories.
24. Here, our study can be linked to some recent research on heterogeneous effects of conspiracy thinking, for example, Kim (2019) who shows that under certain conditions beliefs in conspiracies increase engagement in politics.
25. We surveyed our subjects on the use of social media but did not find any substantial differences.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Loukas Balafoutas https://orcid.org/0000-0002-3477-5096
Björn Vollan http://orcid.org/0000-0002-5592-4185

References

Abalakina-Paap, Marina, Walter G. Stephan, Traci Craig, and W. Larry Gregory. 1999. "Beliefs in Conspiracies." *Political Psychology* 20 (3): 637–647.

Alaoui, Larbi, and Antonio Penta. 2016. "Endogenous Depth of Reasoning." *The Review of Economic Studies* 83 (4): 1297–1333.

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31 (2): 211–236.

Allred, Sarah, Sean Duffy, and John Smith. 2016. "Cognitive Load and Strategic Sophistication." *Journal of Economic Behavior & Organization* 125: 162–178.

Arad, Ayala, and Ariel Rubinstein. 2012. "The 11–20 Money Request Game: A Level-k Reasoning Study." *American Economic Review* 102 (7): 3561–3573.
Aupers, Stef. 2012. "Trust No One: Modernization, Paranoia and Conspiracy Culture." *European Journal of Communication* 27 (1): 22–34.
Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (1): 122–142.
Blaskiewicz, Robert. 2013. "The Big Pharma Conspiracy Theory." *Medical Writing* 22 (4): 259–261.
Bosch-Domènech, Antoni, José G. Montalvo, Rosemarie Nagel, Albert Satorra. 2002. "One, Two, (Three), Infinity… Newspaper and Lab Beauty-contest Experiments." *American Economic Review* 92 (5): 1687–1701.
Bost, Preston R., and Stephen G. Prunier. 2013. "Rationality in Conspiracy Beliefs: The Role of Perceived Motive." *Psychological Reports* 113 (1): 118–128.
Brotherton, Robert. 2015. *Suspicious Minds: Why We Believe Conspiracy Theories*. New York, NY: Bloomsbury.
Brotherton, Robert, Christopher C. French, and Alan D. Pickering. 2013. "Measuring Belief in Conspiracy Theories: The Generic Conspiracist Beliefs Scale." *Frontiers in Psychology* 4: 279.
Costa-Gomes, Miguel, Vincent P. Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-form Games: An Experimental Study." *Econometrica* 69 (5): 1193–1235.
Del Vicario, Michela, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, Eugene H. Stanley, and Walter Quattrociocchi. 2016. "The Spreading of Misinformation Online." *Proceedings of the National Academy of Sciences of the United States of America* 113 (3): 554–559.
Douglas, Karen M., and Robbie M. Sutton. 2018. "Why Conspiracy Theories Matter: A Social Psychological Analysis." *European Review of Social Psychology* 29 (1): 256–298.
Douglas, Karen M., Joseph E. Uscinski, Robbie M. Sutton, Aleksandra Cichocka, Turkay Nefes, Chee Siang Ang, and Farzin Deravi. 2019. "Understanding Conspiracy Theories." *Political Psychology* 40 (S1): 3–35.
Einstein, Katherine Levine, and David M. Glick. 2015. "Do I Think BLS Data Are BS? The Consequences of Conspiracy Theories." *Political Behavior* 37 (3): 679–701.
Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10 (2): 171–178.
Ford, Chandra L., Steven P. Wallace, Peter A. Newman, Sung-Jae Lee, and William E. Cunningham. 2013. "Belief in AIDS-related Conspiracy Theories and Mistrust in the Government: Relationship with HIV Testing among at-Risk Older Adults." *The Gerontologist* 53 (6): 973–984.
Franks, Bradley, Adrian Bangerter, and Martin W. Bauer. 2013. "Conspiracy Theories as Quasi-religious Mentality: An Integrated Account from Cognitive Science, Social Representations Theory, and Frame Theory." *Frontiers in Psychology* 4: 424.
Georganas, Sotiris, Paul J. Healy, and Roberto A. Weber. 2015. "On the Persistence of Strategic Sophistication." *Journal of Economic Theory* 159 (A): 369–400.
Goertzel, Ted. 1994. "Belief in Conspiracy Theories." *Political Psychology* 15 (4): 731–742.
Gosling, Samuel D., Peter J. Rentfrow, and William B. Swann, Jr. 2003. "A Very Brief Measure of the Big Five Personality Domains." *Journal of Research in Personality* 37 (6): 504–528.
Gray, Matthew. 2010. "Revisiting Saddam Hussein's Political Language: The Sources and Roles of Conspiracy Theories." *Arab Studies Quarterly* 32 (1): 28–46.
Hicks, Raymond, and Dustin Tingley. 2011. "Causal Mediation Analysis." *The Stata Journal: Promoting Communications on Statistics and Stata* 11 (4): 605–619.
Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15 (4): 309–334.
Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality." *American Political Science Review* 105 (4): 765–789.
Iriberri, Nagore, and Pedro Rey-Biel. 2011. "The Role of Role Uncertainty in Modified Dictator Games." *Experimental Economics* 14 (2): 160–180.

Jasinskaja-Lahti, Inga, and Jolanda Jetten. 2019. "Unpacking the Relationship between Religiosity and Conspiracy Beliefs in Australia." *The British Journal of Social Psychology* 58 (4): 938–954.

Jolley, Daniel, and Karen M. Douglas. 2014a. "The Social Consequences of Conspiracism: Exposure to Conspiracy Theories Decreases Intentions to Engage in Politics and to Reduce One's Carbon Footprint." *British Journal of Psychology* 105 (1): 35–56.

Jolley, Daniel, and Karen M. Douglas. 2014b. "The Effects of Anti-vaccine Conspiracy Theories on Vaccination Intentions." *PLoS One* 9 (2): e89177.

Kalichman, Seth C. 2009. *Denying AIDS: Conspiracy Theories, Pseudoscience and Human Tragedy*. Göttingen, Germany: Copernicus.

Kata, Anna. 2010. "A Postmodern Pandora's Box: Anti-vaccination Misinformation on the Internet." *Vaccine* 28 (7): 1709–1716.

Kim, Yongkwang. 2019. "How Conspiracy Theories Can Stimulate Political Engagement." *Journal of Elections, Public Opinion and Parties*. Advance online publication.

Lewandowsky, Stephan, Klaus Oberauer, and Gilles E. Gignac. 2013. "NASA Faked the Moon Landing – Therefore, (Climate) Science Is a Hoax: An Anatomy of the Motivated Rejection of Science." *Psychological Science* 24 (5): 622–633.

Li, King King, and Kang Rong. 2018. "Choices in the 11–20 Game: The Role of Risk Aversion." *Games* 9 (3): 53.

Libman, Alexander, and Björn Vollan. 2019. "Anti-Western Conspiracy Thinking in China and Russia: Empirical Evidence and Its Link to Expectations of Collusion." *Homo Oeconomicus* 36 (3–4): 135–163.

McHoskey, John W. 1995. "Case Closed? On the John F. Kennedy Assassination: Biased Assimilation of Evidence and Attitude Polarization." *Basic and Applied Social Psychology* 17 (3): 395–409.

Nadeau, Francois. 2013. "Explaining Public Support for Space Exploration Funding in America." *Acta Astronautica* 86: 158–166.

Nera, Kenzo, Myrto Pantazi, and Olivier Klein. 2018. "'These Are Just Stories, Mulder': Exposure to Conspiracist Fiction Does Not Produce Narrative Persuasion." *Frontiers in Psychology* 9: 684.

Oliver, J. Eric, and Thomas J. Wood. 2014. "Conspiracy Theories and the Paranoid Style(s) of Mass Opinion." *American Journal of Political Science* 58 (4): 952–966.

Peisakhin, Leonid, and Arturas Rozenas. 2018. "Electoral Effects of Biased Media: Russian Television in Ukraine." *American Journal of Political Science* 62 (3): 535–550.

Penczynski, Stefan P. 2016. "Strategic Thinking: The Influence of the Game." *Journal of Economic Behavior & Organization* 128: 72–84.

Public Policy Polling. 2013. "Democrats and Republicans Differ on Conspiracy Theory Beliefs." https://www.publicpolicypolling.com/wp-content/uploads/2017/09/PPP_Release_National_ConspiracyTheories_040213.pdf

Radnitz, Scott, and Patrick Underwood. 2017. "Is Belief in Conspiracy Thinking Pathological?" *British Journal of Political Science* 47 (1): 113–129.

Rubin, Lawrence. 2011. "Documenting Saddam Hussein's Iraq." *Contemporary Security Policy* 32 (2): 458–466.

Stahl, Dale O. 1993. "Evolution of $Smart_n$ Players." *Games and Economic Behavior* 5 (4): 604–614.

Stahl, Dale O., and Paul W. Wilson. 1994. "Experimental Evidence on Players' Models of Other Players." *Journal of Economic Behavior & Organization* 25 (3): 309–327.

Stempel, Carl, Thomas Hargrove, and Guido H. Stempel. 2007. "Media Use, Social Structure and Belief in 9/11 Conspiracy Theories." *Journalism and Mass Communication Quarterly* 84 (2): 363–372.

Sunstein, Cass R., and Adrian Vermeule. 2009. "Conspiracy Theories: Causes and Cures." *Journal of Political Philosophy* 17 (2): 202–227.

Uscinski, Joseph E., and Joseph M. Parent. 2014. *American Conspiracy Theories*. Oxford, UK: Oxford University Press.

Van der Tempel, Jan, and James E. Alcock. 2015. "Relationship between Conspiracy Mentality, Hyperactive Agency Detection, and Schizotypy: Supernational Forces at Work?" *Personality and Individual Differences* 82: 136–141.
Van Prooijen, Jan-Willem. 2017. "Why Education Predicts Decreased Belief in Conspiracy Theories." *Applied Cognitive Psychology* 31 (1): 50–58.
Wood, Michael J., and Karen M. Douglas. 2013. "'What about Building 7?' A Social Psychological Study of Online Discussion of 9/11 Conspiracy Theories." *Frontiers in Psychology* 4: 409.

Appendix A. Experimental instructions.

Welcome to an experiment about decision-making procedures!

Thank you for your participation!

During the experiment, you and other participants will be asked to make decisions. Your decisions, as well as the decisions of other participants will determine your payoff according to the rules explained in what follows.

The experiment will be implemented using PC. You will make your decisions on the computer screen. All your decisions and response will be kept anonymous and confidential.

The experiment consists of the following steps:

- Video
- Decision 1
- Decision 2
- Questions to the video and other questions

Your overall income from the experiment is the sum of

- Payment for participation of €3
- Randomly determined payoff from the decision-making situations 1 and 2
- Payoff for each correct response to the questions on the video you will watch (for each correct answer you will receive €0.5)

Your payoff from the decisions 1 and 2 depends on your decisions and the decisions of another person, who will be randomly assigned to you. In each decision-making situation, you will be paired with a different person.

In what follows, you will receive the directions to the decisions. We will read the directions out loud and then give you time to read them and to ask questions. If anything is unclear, please raise your hand: we will come to you and respond to the question.

Please do not talk with each other during the experiment and switch off the mobile phones.

Video

You will now watch a six-minute-long video. You will be later asked a couple of questions about the content of this video. Therefore, pay attention to its content!

[HERE THE VIDEO IS SHOWN TO THE PARTICIPANTS]

Decision-making situation 1

In this game, you will be randomly and anonymously paired with another player (from this room) and play a game, in which each of you requests a certain amount of money. This amount of money should be between €5 and €14. Each player receives exactly the sum he

requested. One of the players receives additional €10, if he requests exactly one euro less than the other player.

You and the other player make decisions simultaneously and only once, without knowing what the decision of the other player was. The identity of the other player will not be revealed to you, and your identity will not be revealed to him or her as well. You will learn the decision of the other player only at the end of the experiment.

Example 1

If both players request €12, both players receive €12.

Example 2

If the first player requests €7 and the second €6, then the first player receives €7 and the second 6 + 10=€16

[after the decisions were made, the participants are asked to respond to the following question:]

Which reasons did you have to make the particular decision you made?

Decision-making situation 2

In this game, you are either player A or player B. First both players receive the starting capital of €5. Player A can decide to invest a part of it (from €0 to €5, you can choose any number up to the first decimal place). The invested amount will be multiplied by three and transferred to player B. Player B can then decide which part of the overall (tripled) amount of money he will transfer back to player A. Player B is not under obligation to transfer any amount of money back.

The payoffs of both players will then be determined using the following formula:
Payoff of player A = €5 – Investment + Transfer from player B
Payoff of player B = €5 + Investment of player A × 3 – Transfer to player A

Examples:

If the player A invests €5, player B receives €15. The overall sum player B then holds is 15 + 5 = €20, and player A has €0. If player B, for example, now transfers €10 back to player A, both will be paid €10.

If the player A invests €0, he will receive the payoff of €5, and player B will receive €5.

If the player A invests €2.50, player B will receive €2.50 *3=€7.50+€5 = €12.5. If player B transfers €6 back to player A, player A will receive 6 + 2.50= €8.50, and player B will receive 12.50–2.50 = €10.

If player A transfers €4 to player B, player B receives 3*4 = 12 + 5 = €17. If player B transfers €1 back to player A, player A receives 1 + 1= €2 and player B receives 17 − 1 = €16.

Appendix B. Heterogeneous treatment effects.

Here we analyse whether there is any evidence of heterogeneity in treatment effects, that is, whether particular groups of subjects react to the priming differently than other groups. We focus on the following three characteristics, which could potentially trigger differences in response to the treatment: gender, age and religiosity. This choice is driven by the following concerns:

- *Gender* could serve as a proxy for the extent of information people had on the topics of moon landing conspiracies before the experiment. The literature shows that men systematically exhibit stronger interest in space exploration than women (Nadeau 2013) and thus may be better informed on the topic. Less informed subjects are more likely to be susceptible to a stronger impact of priming compared to subjects who have already been exposed to this information and could have formed an opinion on it. Our sample is gender-balanced, with female subjects accounting for 53%.

Table B1. Regressions of the main experimental outcomes on treatment and gender.

Dep. var.	Belief in moon landing conspiracy	k-Level thinking
Female	0.276	−0.278
	(0.278)	(0.270)
Female × *Treatment*	0.472	0.145
	(0.409)	(0.353)
Treatment	−0.014	0.243
	(0.690)	(0.369)
Constant	1.567***	1.690***
	(0.447)	(0.438)
Observations	144	125
R-squared	0.138	0.058

Notes: OLS regressions. Robust standard errors are given in parentheses. ***Significant at the 1%, **5% and *10% levels.

- *Age* is also a frequent correlate of belief in conspiracies for several reasons: there are age-specific differences in media consumption[25] and in the sense of empowerment, which can move people to be more willing to believe in conspiracies (Stempel, Hargrove, and Stempel 2007). While the age differences in our sample are not very large (given that our participants are students), some variation exists between the ages of 18 and 34 (the mean and median age of 23).

- There are two possible effects of religiosity on conspiracy beliefs. On the one hand, religiosity could discourage conspiracy thinking, since it constitutes an alternative set of beliefs bringing order to chaos. But, on the other hand, it could also encourage it (through the interpretation of the observed world as an outcome of intentional design); it is not uncommon to study conspiracies as quasi-religious beliefs. In both cases, an encounter with conspiracy thinking narratives could have different effects for religious and non-religious subjects: religious subjects could be either more prone to internalising it or be more resistant to it. Religiosity is defined based on responses in the post-experimental survey question of whether subjects consider themselves religious on a scale from 1 (strongly disagree) to 5 (strongly agree).

We estimate linear regressions of the level of beliefs in the moon landing conspiracy and the estimated k-level reasoning on gender (Table B1), age (Table B2) and self-reported religiosity (Table B3), a treatment dummy and interaction terms between the treatment dummy and the respective characteristic. For age and gender differences we find no effects. For religiosity, however, we do establish heterogeneous treatment effects. Therefore, in this regression we also compute marginal effects of the priming dummy for different levels of religiosity with the associated significance levels. The regressions in Table B3 reveal that the effects of the priming on k-level thinking are significant for all subjects except those identifying themselves as definitively non-religious (about 40% of the sample), and that they become significantly stronger with more pronounced religiosity (as indicated by the interaction term *Religiosity x Treatment* in the last column of Table B3).

Appendix C. Online survey.

C1: description and analysis of the survey

To validate our interpretation that the conspiracy narrative caused an increase in strategic sophistication, we applied an additional tool. From December 2017 to January 2018, we constructed an anonymous online survey (using SurveyMonkey), which was sent to students of the University of Innsbruck through the official e-mail address used to circulate surveys addressed to the students. The survey was explicitly provided to a set of participants while

Table B2. Regressions of the main experimental outcomes on treatment and age.

Dep. var.	Belief in moon landing conspiracy	K-level thinking
Age	−0.014	−0.055
	(0.046)	(0.043)
Age × *Treatment*	0.108	0.073
	(0.067)	(0.071)
Treatment	−1.727	−1.248
	(1.569)	(1.619)
Constant	2.295**	2.566**
	(1.094)	(0.998)
Observations	144	125
R-squared	0.109	0.060

Notes: OLS regressions. Robust standard errors in parentheses. ***Significant at the 1%, **5% and *10% levels.

Table B3. Regressions of the main experimental outcomes on treatment and religiosity.

Dep. var.	Belief in moon landing conspiracy	K-level thinking
Religiosity	0.149	−0.172
	(0.124)	(0.108)
Religiosity × *Treatment*	−0.253	0.317*
	(0.172)	(0.178)
Treatment	1.310***	−0.222
	(0.412)	(0.400)
Constant	1.636***	1.668***
	(0.294)	(0.305)
Observations	144	125
R-squared	0.106	0.080
Marginal effects of *Treatment*		
Religiosity = 1 (not religious)	1.057***	0.095
	(0.275)	(0.256)
Religiosity = 2	0.805***	0.411**
	(0.203)	(0.186)
Religiosity = 3	0.552**	0.728***
	(0.256)	(0.259)
Religiosity = 4	0.299	1.044**
	(0.386)	(0.403)
Religiosity = 5 (highly religious)	0.046	1.361**
	(0.540)	(0.566)

Notes: OLS regressions. Robust standard errors in parentheses. ***Significant at the 1%, **5% and *10% levels.

coming from the same pool as our experimental subjects were recruited (Innsbruck students) but not belonging to the group of experimental participants – otherwise we would face the already described problem that asking survey questions before the game is played we may bias the way subjects behave in the game, and asking questions after the game is played we may bias the responses to the survey. From this point of view, looking at a different, but still comparable group of subjects represents an attractive solution.

The survey asked the students to watch the conspiracy priming video and then report to us typical reactions the video caused. The main objective of the survey is not to establish a causal effect of the conspiracy priming (as this is ensured through the randomised design of the experiment), but rather to obtain some information about how exactly subjects could have interpreted the conspiracy priming. Since responses to the survey were not incentivised, we treat its outcomes with caution; however, we believe that the survey contains some useful additional information helping us interpret behaviour in our experiment.

Overall, we received 186 responses. The average age of the respondents was 24.4 years old (almost identical to that of the experimental subjects; if one removes six senior students older than 40 years, the average age becomes 23.7 years), and the sample was almost gender balanced (43% female). Our questionnaire included five main questions concerning several possible

Table C1. Online survey responses.

	Full sample		Know little of moon landing conspiracy		Respondents finding the video silly		Respondents not finding the video silly		Difference in means, respondents finding the video silly and not finding the video silly
	Average response	Share of positive responses (%)	Average response	Share of positive responses (%)	Average response	Share of positive responses (%)	Average response	Share of positive responses (%)	
Question true goals of social institutions	2.421	57	2.379	60	2.966	28	2.171	72	0.795***
Question the motives of people I deal with	2.765	40	2.707	41	3.034	29	2.642	45	0.392***
Question the things I believe	2.402	56	2.356	56	2.966	33	2.146	67	0.819***
More difficult to convince	2.821	33	2.780	36	3.172	17	2.667	40	0.506***
Think more about the motives of others	2.614	48	2.576	49	2.828	41	2.536	51	0.291*

Notes: The average response is computed on a four-point scale, 1 = fully agree, 2 = agree, 3 = disagree, 4 =fully disagree. Respondents finding the video silly are defined as those who fully agree or agree with this statement; and respondents not finding the video silly are those who fully disagree or disagree with this statement. ***Significant at the 1%, **5% and *10% levels.

reactions to the conspiracy video, which respondents had to respond to using a four-point scale (fully agree, agree, disagree, fully disagree; for the subsequent analysis we coded the responses using a 1 to 4 scale, with 1 being 'fully agree'). The five questions were the following:

- The video makes me question the true goals of social institutions;
- The video makes me question the motives of people I have to deal with;
- The video makes me question whether things I believe in are actually true;
- After the video it is more difficult to convince me that something is true;
- After this Video I think that if you have to deal with other people, you have to think how they will react on your actions.

The first three responses explicitly focus on different types of trust: trust in the social institutions and establishments; in other people in general; and in common knowledge and beliefs. The fourth question is about increasing scepticism and doubt about information people receive. The last question can, with certain caveats, be interpreted as capturing strategic sophistication (although of course in a rather generic way). In Table C1 we report the average responses to these questions. We report both the average response and the share of those who agreed or fully agreed with each statement (share of positive responses).

Looking at the full sample, the situation appears to be as follows. More than half of the respondents agree that exposure to the conspiracy narrative made them question the true goals of social institutions and the things they believe in. This is not surprising: these are the main two components of the moon landing conspiracy narrative, which is based on the idea that the government staged the moon landing and that the common knowledge about this episode is not true. The third highest response is that of strategic sophistication, indicating that this reaction is quite widespread among respondents. Thus the survey indicates that, while the main effects of exposure to the conspiracy priming are associated with the content of the conspiracy narrative, an increase in strategic sophistication becomes very likely as well. Moreover, the relatively low share of positive responses to the question regarding trust in other people's motives echoes well with our experimental finding that the conspiracy priming has no impact on trust.

We also asked the question how much respondents knew about the moon landing conspiracy before watching the video; if we include only those who disagree or fully disagree with the statement that they knew a lot about this topic (32% of respondents), we find a similar pattern to the general population (though the share of positive responses is higher, that is, those who did not know a lot about the topic and did not have a chance to form an opinion were a bit more likely to react on the conspiracy priming (columns 3 and 4 in the Table C1).

The results become more interesting if we utilise another question of our survey: we asked respondents to indicate whether they find that the newsreel they saw was 'just silly'. Through this question we attempt to identify the group of respondents who clearly were not convinced by the content of the conspiracy narrative, and to understand whether they changed their behaviour in any way. Thirty two percent of respondents agree or fully agree that the video is silly. The fifth and sixth columns in Table C1 show that the majority of these individuals do not think that after watching the movie they started to be less trusting of governmental institutions, other people they interact with, the information they received or that they are now more difficult to convince that something is true. However, 40% of respondents who clearly disagree with the content of the conspiracy priming still indicate that after watching the video they believe they have to think more about the motives of other people they interact with. If we compare the group of those who find the video silly to those who do not share this opinion (columns 7 and 8 in Table C1), one can see that the latter are significantly more likely to agree with all the items (as expected), but for the strategic sophistication item the difference in means is the smallest and least significant ($p = 0.051$). This fits the main argument of our study: exposure to conspiracy theories results in higher strategic sophistication even if it does not convince people altogether.

C2: full text of the survey (administered in German and in English)

Dear participant

Thank you very much for your kind willingness to help us! The survey below is part of a joint project implemented by [THE AUTHORS NAMES AND AFFILIATIONS].

We would like to ask you to watch the video located at the link below

[LINK TO THE VIDEO] and to respond to some questions to this video. You will need about 6 minutes to watch the video and another 5-10 minutes at most to respond to the survey. Thank you very much once again!

Q1: This video makes me question the true goals of social institutions
Fully agree/agree/disagree/fully disagree
Q2: This video makes me question the motives of people I have to deal with
Fully agree/agree/disagree/fully disagree
Q3: This video makes me question whether things I believe in are actually true
Fully agree/agree/disagree/fully disagree
Q4: After this video it is more difficult to convince me that something is true
Fully agree/agree/disagree/fully disagree
Q5: After this video I think that if you have to deal with other people you have to think how they will react on your actions
Fully agree/agree/disagree/fully disagree
Q6: I simply find the video silly
Fully agree/agree/disagree/fully disagree
Q7: This video makes me think though I am not convinced by it
Fully agree/agree/disagree/fully disagree
Q8: I was already familiar with the main arguments of the video
Fully agree/agree/disagree/fully disagree
Q9: Other impressions (if none, insert 'None' in the field below)
Q10: And a couple of questions about you
Age (in years)
Gender (1 = male, 2 = female)
Would you describe yourselves as religious (on a scale from 1 to 5, 1 = highly religious, 5 = not religious)
Your major (1 = social sciences (e.g. sociology or political science), 2 = business administration or economics, 3 = law, 4 = humanities (e.g. history or linguistics), 5 = medicine, 6 = engineering, 7 = sciences or mathematics, 8 = other field)

APPENDIX D. Mediation analysis: baseline regressions.

Table D1. Linear regressions on mediator and outcome.

Dependent variable	Belief in moon landing conspiracy	K-level thinking
Treatment	0.844***	0.447**
	(0.212)	(0.185)
Belief in moon landing conspiracy		−0.009
		(0.076)
Constant	1.828***	1.297***
	(0.141)	(0.202)
Observations	125	125
R-squared	0.114	0.047

Note: Regressions estimated using OLS. Robust standard errors in parentheses. ***Significant at the 1%, **5% and *10% levels.