An experimental study of countermeasures against threats: real-world effects meet treatment effects

Roman Chytilek[1] · Miroslav Mareš[1] · Jakub Drmola[1] · Lenka Hrbková[1] ·
Petra Mlejnková[1] · Zuzana Špačková[2] · Michal Tóth[1]

Abstract

The experimental study of positions on policies and measures against various new types of threat is fast becoming a mainstream research practice. In this article we argue as follows: in security studies in particular, there is a risk that the experimental treatment is contaminated by subjects' previous experience of the real world ('contamination'), and this may substantially complicate the assessment of the size of the experimental treatment's causal effect. We discuss ways to decrease the risk of uncontrolled contamination. Using two experimental case studies we show two typical cases of contamination in security studies (one, where the contamination of all treatments was extremely high, and another, where the level of contamination was unknown and might have varied across the experimental groups) and consider what this implies for the substantive results of the experiments. An analysis of contamination should become a routine, especially when reporting security experiments.

Keywords Security · Experiments · Contamination · Pre-treatment · Treatment effect

1 Introduction

The challenges of the past decade, including the migration crisis, conventional and cyber terrorism and the Covid-19 pandemic, have pushed the policy for managing dynamically evolving threats to the forefront of national security. In an environment of political competition that is increasingly based on valence (Curini 2018)—that is, the emphasis is placed more and more on competence in managing political issues as well as on blame attribution, to the detriment of ideological differentiation (Sulitzeanu-Kenan and Zolhhöfer 2019)—politicians must be ever more sure of avoiding unenforced mistakes. In an environment full of 'new species of troubles' (Slovic 2002), this is an exceptional challenge.

✉ Roman Chytilek
    chytilek@fss.muni.cz

1   Department of Political Science, Faculty of Social Studies, Masaryk University, Joštova 10,
    602 00 Brno, Czech Republic

2   Department of Quantitative Methods, Faculty of Military Leadership, University of Defence,
    Kounicova 65, 662 10 Brno, Czech Republic

                                                                    ⚫ Springer

What the risks described have in common is that they are able to incite intense negative emotional reactions among the public, such as fear, anger and disgust, and to influence the life styles of the populations affected (Huddy et al. 2002). In response to this, studies that investigate the responses caused by threats to individuals, as well as their coping strategies, are becoming increasingly relevant (Rogers 1975, Lerner et al. 2003). The aim is to understand not just the connection between threat appraisal and coping appraisal at the individual level (the field of human security) but also what individuals, on the basis of their appraisal processes, expect from those who exercise political authority (and here we see the overlap with national or homeland security). There are theories about how threat perception influences the support for variously militant retaliation policies or restrictions on privacy (Gross et al. 2016), what social norms should inform the use of face recognition technologies in policing (Bromberg et al. 2020), how coping appraisal affects the willingness to give up civil freedoms (Garcia and Geva 2016) and more generally how a large measure of coping appraisal is a necessary condition of accepting restrictive policies (Krueger et al. 2020).

Experimental methods assert themselves increasingly in the conduct of such studies, and this is linked with the greater emphasis placed in security studies on causality (cf. Baele and Thompson 2017). In this respect, experiments offer a number of advantages. The most important is that there are greater possibilities of manipulating the independent variable than is the case in an observational study. If this variable is the content, source or channel of information about a threat or risk, researchers have much better options for isolating the constituent elements, and of adjusting its values to the question under study (Iyengar 2011). This is of exceptional importance, because the threats are often complex and the experimenter's research task is to identify and understand how their various components influence threat appraisal. A second, no less important, aspect is the substantial spatial and temporal diffusion of the effects of the (potential) independent variable on the research subjects in their real life. In observational research, this creates difficulties with identifying these effects and data collection.

Systemically, the 'experimental state' (see Jones and Whitehead 2018) is becoming increasingly prominent and is characterised by, first, an increasing willingness on the part of governments and states to use innovative policy procedures as part of governance, and second, by using experimental methods to introduce these innovations. The domain in which experimental methods are employed most often as part of this process is the study of the factors that influence human behaviour, in particular cognitive biases or unconscious heuristics. This is while randomised experiments serve as a research tool (see Haynes et al. 2012; John 2013) in resolving issues at the core of which are questions such as 'what (which political measure) works (best) to address a particular issue' and this is often followed by the practical application of specific policy measures. Jones and Whitehead (2018) cite examples of such policy preparation and implementation (including in homeland security) from a number of countries, including the US, UK and Singapore. They note that this is increasingly concerned with very divisive issues and, in the spirit of 'libertarian paternalism' (Thaler and Sunstein 2008), is often aimed at discreetly pushing individuals towards decisions that are beneficial to the state (latest examples: Goel et al. 2021, Lunz Trujillo et al. 2021). They also point out the possible risks of this approach—in particular, the possible artificiality of the experimental environment, the risks implied by the fact that only a small part of the real world can ever be brought into the laboratory and the potential implications for the deterioration of the relationship between the state and its citizens—if any such study is ever used to justify the intensive deployment of policy innovations.

In this paper we seek to show that the use of experimental methods is not without other issues, in particular in the crucial domain of assessing the type and size of the causal effect of the independent variable. We purposely use the term 'effect of the independent variable', because our study is concerned not just with the experimental treatment, but all the effects of the independent variable, including in the real world. We first present the problem of 'contamination'[1] of the experimental treatment and its consequences for causal inferences, and discuss the ways of reducing this contamination. We then present two experimental studies, remote from each other in terms of the phenomenon examined (one where the contamination is evidently high, and the other where the measure of contamination is not apparent). The aim is (not just) to report the results of these studies, but in particular to assess the impact of contamination on the estimate of the causal effect, while offering procedures that should become a regular part of reporting laboratory and survey experiments, where contamination represents a potentially serious problem.

2 The problems of interactions between the real world and the experimental treatment

In their critical appraisal of the opportunities and problems offered by survey experiments, Gaines et al. (2007) point out a paradox that is inherent in all types of experiment (including laboratory and field experiments). The experiment (i.e. the active intervention of the researcher in the 'data generating' process) is usually justified as an effort to uncover causal relations, which are potentially highly socially relevant. But this poses a serious problem: the experimental treatment, having as its topic, let's say, disturbing news of terrorism or (dis-)information about the possible risks of vaccines or 5G networks, is not necessarily first experienced by the experimental subjects in the laboratory—they might have already encountered it in real life. In a classic experiment with a control and an experimental group, the control group includes subjects who have already undergone treatment in the real world. In an extreme case, where the treatment or its analogue have been experienced in the real world by all subjects, we are comparing causal effects in two groups, where the subjects in the experimental group had *one more treatment* than their counterparts in the control group—but this does not necessarily mean that the subjects in the control group have not been exposed to the treatment.

This phenomenon poses another potential problem, both for experiments with a control group and for those without. It is not apparent whether the causal effect in subjects exposed to repeated treatment is of the same size as with subjects recruited from a virgin population, who undergo exposure in the laboratory or survey for the first time. Theoretically, we can envisage both those variants where the marginal utility of the treatment decreases with repeated exposure, and those where the treatment effect in the real world is short term, disappears entirely and is reactivated by a new treatment. Said more generally, the previous treatment in the real world affects the experimental treatment in some way.

The two mechanisms by means of which the real world impacts on the effect of the experimental treatment, usually lead to the average treatment effect measured being

---

[1] By 'contamination' we broadly mean the previous impacts in the real world of the variable that is being manipulated in the course of the experimental treatment, while the term 'pre-treatment' is reserved specifically for those situations where subjects previously experienced in the real world the same treatment to which they are exposed in the experiment.

smaller than if we were to investigate an entirely virgin population. This does not necessarily mean that we should always consider the estimates of treatment effects as conservative and as underestimates. The 'distraction-free' factor of the experimental treatment works in the opposite direction. An experimental session or an on-line survey seldom overwhelm the experimental subjects with a number of intense, unconnected stimuli, as happens in the real world. The stimulus examined then has too strong an effect and the resulting effect measured is an overestimate.

The contamination problem is in no way disrupted by the assumption that, on the basis of a random assignment procedure, the groups are not different from each other in any observable or non-observable variables apart from the (non-)existence of the experimental treatment, which is at the heart of the methodology. Unfortunately, the contamination problem negates some of the advantages that would normally follow, because it renders problematic the assumption that the measured difference in the value of the dependent variable between the control and experimental groups is fully accounted for by the experimental treatment.[2]

The study of countermeasures against threats is an area where the risk of contamination by the external world is very strong. Seldom is the experiment so abstract that we could justifiably argue that pre-treatment is minor or non-existent.[3] If the variables most often manipulated are the nature of the threat, its intensity, salience and source, or the source of the authority seeking to eliminate the threat, the content of the message intended to elicit a certain response among the recipients and/or the communication channel (cf. Baele and Thomson 2017), various contamination-related issues arise. If, for instance, a national (expert) institute has the main say in formulating counter-epidemic policy in a country, and all subjects know this, there is not much sense in experimentally investigating trust in various potential channels conveying epidemic information, because in all groups (including the control) this trust will be contaminated by the subjects' experience of the national institute from the real world. By contrast, if the position of such an institute in the real world were such that most of the subjects would not associate it with epidemic policy, an experiment manipulating the source of counter-epidemic authority could be greenlighted. These contrasting examples suggest that contamination is something that needs to be considered both in designing an experiment, and in deliberations about how much the effect measured is real.

How to resolve this problem? Perhaps the worst option is to focus solely on investigating issues where we do not assume pre-treatment or those where we do not believe the effect of the treatment to be long-lasting. Such situations are mostly irrelevant to the study of security-related behaviours (cf. Gaines et al. 2007), nor do they have great theoretical value. Some authors (e.g. Fesenfeld et al. 2021) believe that if the experimental treatment

---

[2] Let us imagine an experiment involving two groups, where the argument that vaccination is painless (independent variable) is expected to influence the subjects' willingness to be vaccinated (dependent variable). The measured willingness to be vaccinated will be 51% in the experimental group and 49% in the control group (not exposed to the painlessness argument). An obvious problem in interpreting this is that many subjects in both groups previously participated in the 'experiment' in the real world *in the experimental group*, where they would have been exposed to this argument, and this may have affected their willingness to be vaccinated. The size of the manipulation effect could therefore be bigger than the two percentage points measured in the laboratory.

[3] A rare example is provided by Yan et al. (2020), who studied the effect of emotions on responses to terrorism. The dependent variable, representing trust in public resources versus reliance on a private solution, was the type of insurance that subjects bought in the laboratory game.

is not preceded by any items in the questionnaire that would direct the subjects' attention to the problem examined, this provides a sufficient defence against contamination. They see this strategy as particularly appropriate for framing experiments. Although we do not deny the correctness of such an approach in terms of the experimental script, it seems to us that this conception assumes too much about the treatment effects (both in the real world and in the experiment) being short-lived—that they will disappear quickly—and hence it fails to address sufficiently those situations where the real-world effects are persistent in the subjects. We do not believe this to be an ironclad defence against contamination. This consideration also points to a possible tension between the contamination of the size of the effect by real-world events and items that precede the treatment in the experiment, and the subsequent measurement of the dependent variable. These too may cause contamination (so-called prior questions or accidental spillover effects, cf. Gaines et al. 2007). Attempting to establish how much a subject is contaminated by the real world might not be a reliable or available strategy in every experiment.

By contrast, repeated experimental treatments, and repeated measurements, provide a promising, but unfortunately only an auxiliary, path. This strategy is based on the premise that most of the truly theoretically relevant treatments are repeated in the real world too, and if this property is translated into the experimental treatment, the investigators at the very least obtain the information as to whether the repeated treatment actually has a deflating influence on the effect measured or whether the pattern of the effect's repeated manifestation is more complicated (see Lecheler and de Vreese 2013 for a discussion of this problem). In priming experiments, which increase the saliency of some stimulus, the repeated measurements also provide other advantages in terms of building up the theory (cf. Rivers and Sherman 2018). Correspondingly, the disadvantages are evident: they are related to the technical aspects of the experiment, such as greater financial outlay (and in the case of laboratory experiments, greater need for resources and laboratory personnel) and experimental mortality (or experimental attrition).

The experiments in which investigators are interested in comparing either multiple experimental groups among each other, or where the experiment is constructed as an international study and the effect of experimental treatment is compared across multiple countries, seemingly present a specific situation. But in terms of eliminating pre-treatment, the advantage of such cases is chimeric. The comparison of many experimental groups multiplies rather than decreases the contamination problem, and the inflationary and deflationary factors of the experimental effect must be considered for each individual group separately. In international studies, the context of the various countries with their different measures of contamination, or different contamination potentials, is in fact necessary to explain the variance between nations (for example: Makkonen et al. 2020).

Thus, several recommendations by Gaines et al. (2007) continue to hold true. First, investigators should carefully consider whether the presence or absence of treatment in subjects is a binary or a continuous variable, and prefer the latter by default (i.e. assume that at least some subjects have been exposed to some measure of contamination). They should also assess the quality of the previous treatment(s) (e.g. its clarity and the way it is framed) and how closely its content matched the experimental treatment (cf. Linos-Twist 2018). They need to declare their consideration about the measure of contamination explicitly. If they do so, what should follow is a qualitative assessment of all three of the effects outlined (i.e. the two deflationary effects, caused by contamination of some or all subjects, and potentially also by the decreasing effect of repeated treatment; and one inflationary, associated with the greater power of the experimental stimulus in a laboratory or in a survey); and this should be accompanied by a consideration of what this ultimately implies for

the experimental effect measured. Investigators might also want to consider whether the effect of the treatment to which the subjects were exposed in the real world was uniform or heterogeneous (cf. Druckman and Leeper 2012).

3 Data and methods

In the Results section we present two experimental studies from a project dedicated to an experimental analysis of emotional responses to threats.[4] We briefly present the basic parameters of the experimental design (objectives, experimental manipulation, environment and conditions under which the experiment was conducted and results) and subsequently discuss these in terms of the relationship between inflationary and deflationary factors and the estimated causal effect. This is a much less rigorous procedure than a statistical analysis of experimental results, and the conclusion is less rigorous too; but it should at the very least suggest the possible direction of the bias in describing the size of the effect.

Similarly to Gaines et al. (2007) we distinguish in terms of pre-treatment four basic analytical situations that may arise as a result of an interaction between previous treatment in the real world and that undertaken in the experiment. The first is that the subject could not have been exposed to treatment in the real world. The second, that there was treatment in the real world, but like the experimental treatment it had no effect. In the third, the ultimate effect has already been caused by the manipulation in the real world, and the experiment does not increase the effect in any way. And finally, in the fourth, real-world treatment has an effect, but it is short-lived and exhausts itself. The experimental treatment then works in the same way—we measure an effect, but it does not endure much beyond our study. These, of course, are ideal types. Another deflationary factor—the decreasing effect of repeated treatment—is a subtype of the last-mentioned situation (there is pre-treatment but it does not persist). This is assessed qualitatively in consideration of the context in which the experiment took place. We seek above all to estimate the frequency of subjects' exposure to treatment in the real world, and to assess its impact on the possible insensitivity of subjects to experimental treatment.

Conceptualisation is more difficult with potentially inflationary factors. Gaines et al. (2007) assume merely that experimental treatment is too strong and hence it overestimates the effect; but there are multiple factors at play, related to how well the situation and the real-world effect are represented. This is connected with three types of 'realism': mundane, experimental and psychological (cf. Wilson et al. 2010; Druckman and Kam 2011). Mundane realism assumes that experimental contents agree with real-world contents—in other worlds, the treatments are similar. In terms of inflation, if the experiment's measure of mundane realism is appropriate, it should have a neutral effect, and in security experiments mundane realism is usually easy to achieve. Experimental realism is concerned with how seriously the subjects take the experimental treatment, whether they afford it sufficient attention and believe in its veridicality. This type of realism potentially contributes not just to experimental effect inflation (in situations where investigators make an excessive effort to have a strong treatment) but also to its deflation (in situation where experimental realism is low overall and the subjects do not believe the treatment under laboratory conditions). Psychological realism (Aronson et al. 1994) is also a potential inflation factor. It is

concerned with whether the nature and intensity of psychological processes, occurring during the experiment (for example, the activation of stereotypes) are similar to those occurring during real-life treatments.

4 "Results"- case studies

4.1 Experiment 1: framing of information about Covid-19

4.1.1 Aim of the experiment

The experiment examined whether it were possible, using the framing of information about the Covid-19 pandemic, to increase people's resilience to the negative impacts of the pandemic, and whether it were possible to influence their willingness to comply with government measures to limit the spread of the coronavirus.

4.1.2 Experimental manipulation

The experiment examined how the framing of information (independent variable) influenced the feeling of personal and social threat from the pandemic, the evaluation of pandemic risks, emotional experiences, behavioural changes in response to the pandemic and agreement with the introduction of security measures to combat the pandemic (dependent variables). The experiment used three psychological theories that seek to explain the processes linked with people's emotional responses. Each of these theories approaches the issue of eliciting attitudinal change, or of activating socially desirable values in a situation of threat, in a different way, which allow the theories to be isolated as various types of framing and assign one to each experimental group.

The first, lay-epistemic theory (LET), assumes that a change of judgment or the adoption of a new stance is primarily linked with an individual's willingness to make a new judgment. Thus a certain measure of epistemic motivation is typical of the theory. If subjects are motivated towards rational deliberation, it is possible to dampen their 'reflexive emotional responses and panic'. Thus the theory assumes that we can decrease the feeling of threat and risk by means of a certain framing. The susceptibility to behavioural change and support for stronger security measures should likewise decrease (Kruglanski 2001).

A second way of framing that may influence human behaviour is by reporting about events in terms of their dreadfulness and unexpectedness. Slovic (2002) argues with regard to this that emotional responses are strongly conditioned by how much the event creates feelings of dread in people, and how new the stimulus actually is. If people feel that the issue is close to them, this will influence their responses. For instance, it matters whether an individual knows people who are at risk. To measure this effect, the study used a scale of 11 characteristics, which, according to Slovic (2002), influence the response to the threat. These are concerned with whether the threat is uncontrollable, causes dread, is a catastrophe, is global in scope, has fatal consequences, is not equitable, poses high risk to future generations, is not easily reduced, has risks that are increasing, is involuntary, and affects the subjects. Proceeding from Slovic's study, it was therefore assumed that a realistic portrayal of a threat (in this case, the Covid-19 disease) would decrease people's fears of being exposed to the threat.

The third theory is based on the concept of 'emotional contagion', according to which people adjust their behaviour and experiences to those of people around them, or to the experiences of people about whom they receive information. Schachter (1959) emphasises that individuals do not have to be in immediate contact or have personal communication with people whose emotional responses they potentially adopt. Thus, if the responses of others exposed to the pandemic are framed as calm, deliberate and not succumbing to fear, the expectation is that those responses will be reflected in the results of the experiment, or in the causal nexus examined.

4.1.3 Population, sample, fee, experimental environment and ethics

The survey experiment took place in May 2020, and was conducted on-line due to the pandemic measures in force. This was roughly at a time when the first wave of the Covid-19 pandemic had ebbed in the Czech Republic. The study participants were recruited from the population of subjects registered in the ORSEE database managed by the project's investigators (a non-representative sample of the Czech population). For this study 949 subjects were invited, of whom 301 agreed to participate. Only Czech citizens were invited. All registered participants were sent a link to the on-line questionnaire. Each respondent who completed the questionnaire in full was awarded a fee of 300 Kč (ca. 15 USD). The questionnaire itself took about 30 min to complete. Participants filed informed consent forms and were given a short on-line debriefing upon completion. There was no deception in this experiment.

4.1.4 Experimental procedure

The respondents were randomly distributed, uniformly into four (three experimental and one control) groups according to the experimental conditions, with 75 or 76 in each group. The random assignment was performed using the software algorithm in the Qualtrics system that was used to implement and distribute the questionnaire. In the questionnaire script, the subjects first answered questions concerned with their demographics (gender, age and education) and life style prior to the outbreak of the Covid-19 pandemic, and also whether they or someone close to them were in a 'Covid-19 at risk group'. Their personal characteristics were measured (big five scale and authoritarian scale). Subjects were then exposed to experimental manipulation (they read a piece of about 350 words containing one of the framings; the control group did not read a text about Covid-19). In the next steps the subjects first assessed the emotions elicited in them by reading the text (fear, anger, sadness, calmness, joy, surprise or disgust) and then the emotions elicited in them by the Covid-19 pandemic. They also answered questions on the characteristics of the pandemic (the categories of dread, controllability, catastrophic nature, extent of damage, prevention, risk escalation and prospective threat), and whether they thought the exposure to the experience of the pandemic would lead to significant life-style changes. Following that, they assessed the government's handling of the pandemic (in healthcare and the economy) and the meaningfulness of present and planned countermeasures to the pandemic (24 in total), as well as the experimental text itself and the role played by the media during the pandemic. The questionnaire was completed with a short debriefing.

4.1.5 Results

The various framings did not cause differences between the groups, in terms of the emotions declared, or in the emotions they declared generally concerning Covid-19. Nor did they influence the assessments of the role played by government in dealing with the pandemic, the role of the media, or support for planned countermeasures to cope with the pandemic. Together with these null results, statistically significant differences were measured among the groups, concerned with their assessments of the experimental texts themselves. The subjects in the Kruglanski and Schachter groups assessed the experimental material presented to them as significantly more manipulative than did the subjects in the control and Slovic groups. The text in the Slovic group was assessed as less manipulative than that in the Schachter group, and information therein presented was assessed as more relevant than the Kruglanski and Schachter, and even the control, groups. The texts in the Kruglanski and Schachter groups were seen as more manipulative than that in the control group. These effects can be interpreted as follows: the subjects saw a text based on an honest and empirically justified conceptualisation of the risk and its contextualisation using statistical arguments as more valuable in terms of information provided, and less manipulative, than the other framings.

4.2 Experimental and real-world treatment

4.2.1 Pre-treatment issues

The Covid-19 experiment was conducted in May 2020, at a time when the first wave of the pandemic had substantially ebbed in Czech Republic. Of course, as the country tried to manage the pandemic, the topic of Covid-19 became dominant in the public discourse. Studies of Czech government communications (Kabrhelová 2021; Šenk 2020) show that Czech leaders did not stick to just one of the framings here presented, but used all three in various contexts. The efforts to conceptualise the risks (Kabrhelová 2021) were somewhat weaker than abroad, but still very strong; and those framings concerned with emotional contagion, or with the resulting conformity of emotions, as well as those positively influencing the motivation to comply with pandemic measures, were also frequently used.[5] The pandemic was pervasive in the news: between a quarter and a third of all news reports in March to May 2020 were concerned with it (Anopress 2021; Vaverka 2021). Thus it is fair to assume that all participants in the experiment were exposed to pre-treatment, and that each of them was exposed (probably multiple times) to treatments (framings) that were used in *all* experimental groups. These framings were typically linked with the release of government regulations seeking to control the pandemic, and the justification of these measures in the mass media. At the same time, the Covid-19 pandemic was by far the most salient social issue. Thus it may be assumed that in terms of pre-treatment the experimental situation was very close to that where the full effect of treatment occurred in the real

---

[5] The speech by Prime Minister Andrej Babiš on 9 April 2020, before Easter, can serve as an example (IRozhlas 2020). Babiš repeatedly appealed to people, arguing that we must all face this together by behaving uniformly (framing by the conformity of emotion argument), while rationally justifying the pandemic measures (we need to help senior citizens and the health system) and justified the strictness of the measures by the seriousness of the situation, which was absolutely exceptional and unprecedented in the 30-year history of an independent Czech Republic (contextualisation of risk).

world and persisted in the laboratory, and that it was neither increased nor decreased by the experimental treatment. Even with subjects among whom the full effect of treatment had not occurred in the real world, we cannot with certainty rule out the hypothesis of very strong deflationary factors, caused by multiple contaminations of treatments. One of the consequences of this was that the greatest possible effect of the experimental treatment in every subject was smaller than would have been the case in a virgin population.

We believe the inflationary factors were neutral. The experimental material had a solid mundane realism; the news report presented to subjects in its length or content did not deviate from contemporary reporting of the pandemic in the real world. Experimental realism was facilitated by the fact that the situation was not too artificial (on-line survey experiment). Mechanisms were built into the experimental instructions to ensure that subjects did actually become acquainted with the text containing the relevant framing (a notification of upcoming manipulation check), and this increased the salience of the topic slightly compared to the real world. It is questionable, however, whether this simply compensated for the circumstance that in the real world the treatments framing the pandemic in various ways were also very salient. In this case, psychological realism is very closely linked with experimental realism—as long as it is true that the subjects took the treatment seriously, it should elicit in them psychological processes similar to those occurring in the real world.

On the basis of this reasoning, we note that the contamination by the real world caused a potentially substantial deflation of the effect of the variable manipulated (the framing of information about the pandemic). Indeed, the results suggest as much: whereas concerning the dependent variables, such as immediate emotional responses, assessments of the character of the pandemic, its influences on people's life styles and agreement with countermeasures, the differences between the groups were not statistically significant, when assessing the treatment itself, there were substantial differences in how the subjects saw the three framings presented: risk contextualisation was seen as least manipulative and most informationally valuable, while the remaining two framings were seen as more manipulative, even compared to the control group. Thus the result of the experiment was that framing matters and that this needs to be taken into account particularly in those situations where political leaders communicate with a population that is still virgin in terms of threat perception.

4.3 Experiment 2: differences between emotional responses to conventional and cyber terrorism

4.3.1 Aim of the experiment

The experiment sought to ascertain whether people respond differently to stimuli featuring cyber and conventional terrorism. Areas of study: emotional responses, individual experiences, impact on people's life styles and their stances on security measures.

4.3.2 Experimental manipulation

One control and two experimental groups. Subjects in the groups were exposed to different stimuli of the independent variable. One of the groups saw in a laboratory experiment a commercial television news report (created by the investigators) about the possible danger of a conventional terrorist attack in the Czech Republic, on the railway, using a derailing device, while the other watched a clip about the same type of threat in which

the mechanism of the attack was different (a cyber-attack on the systems of the Railways Administration, controlling railway traffic). The control group saw no clip about terrorism.

4.3.3 Population, sample, fees, experimental environment and ethics

Subjects were recruited from two populations: university students and security professionals (police and army officers). The experiment took place in November and December 2018, with additional sessions in January and February 2019. There were 284 participants in total. The experiment was conducted at the computer labs of Masaryk University in the Czech Republic. A total of 20 sessions were held, each with 10–20 participants. The experiment was distributed in the Qualtrics system and lasted for about 40 min. Subjects were awarded a show-up fee of 350 Kč (about 17 USD). Participants filed informed consent forms before the experiment. The experiment contained a deception (fake news report), which was revealed during debriefing and the purpose of the deception was made entirely clear to the subjects.

4.3.4 Experimental procedure

After the subjects had provided demographic information (gender, age, some personal characteristics, professional experience of the area of security and sophistication in security issues), the experimental manipulation followed. Subjects in both experimental groups saw a news clip about terrorism, in a format typical of commercial television broadcasts, about two minutes long. The structures of both clips are the same: they report that the police have apprehended a suspicious foreigner coming from Syria and that plans have been found on their person for planning a terrorist attack. A security expert comments on the situation, noting that the threat is real, pointing out similar attacks throughout the world and inferring that a successfully perpetrated attack would almost certainly cause a massive loss of human lives. This is supported by visual materials in the background. A Railways Administration spokesperson then addresses the issue, saying that such attacks cannot be reliably prevented. The clips end by reporting that the police are taking further measures and that an information embargo has been imposed on the case for the time being. This is to encourage subjects to think that organised terrorism is involved. The two clips differ in only one aspect: in one the threat described is conventional and in the other, cyber.

Having exposed the experimental subjects to the stimulus, the investigators measured the dependent variable: response to threat. This is by its very nature multi-dimensional. The measurements included items ascertaining immediate self-assessments by respondents of their emotional states, their feeling of threat from conventional and cyber terrorism in their personal lives as well as in society, changes to behaviour or life style in connection with terrorist threats and support for political measures connected with the fight against conventional and cyber terrorism.

4.3.5 Results

Exposure to threat of lethal cyberterrorism is not innocuous in terms of people's emotional responses. It elicits negative emotions (anger, fear, disgust and anxiety) only slightly weaker than conventional terrorism. This pattern was replicated in virtually every domain of response to the experimental stimulus described (general feelings regarding terrorism in Europe; a sense of being personally under threat; considerations of life-style changes

in connection with terrorism; index of intolerance of terrorism and preventive measures against it, and selected offensive and defensive counter-terrorism measures). The subjects exposed to the cyberterrorism stimulus generally exhibited stronger reactions than those in the control group (i.e. they declared more concern, willingness to change their life styles and support for various types of countermeasure), but these responses were somewhat weaker than those recorded in the group exposed to a conventional terrorism stimulus.

4.4 Experimental and real-world treatment

4.4.1 Pre-treatment issues

In its main outlines, the experimental manipulation replicated the study by Gross et al. (2016). Conducted in Israel and featuring lethal and non-lethal threats from the Hamas movement, the study found that the negative psychological effects of cyberterrorism were only slightly weaker than those of conventional terrorism. The key difference between the original study and its replication here presented is the way subjects interacted with treatments in the real world. Israelis have substantial experience of terrorist attacks, they know their adversary, exhibit a substantial measure of resilience and are able to consider countermeasures in terms of costs and benefits. Czechs are very different: Czech Republic belongs to the Euro-American civilisation, but unlike other European countries has had no direct experience of lethal Islamic terrorism at the time the experiment was conducted.[6] Despite this it shared with other European countries the uncertainty as to how to confront threats posed by the Islamic State, and perhaps it was the uncertainty of these threats that caused them to be perceived as very urgent in the Czech Republic at the time the experiment was conducted (CVVM 2021; Aktuálně 2016).[7]

In terms of the discussion of contamination, at first glance it seems that the Czech subjects had not been exposed to the same or similar treatment in the real world as their Israeli counterparts. Czechs could not watch news reports about foreigners suspected of terrorism apprehended in the country,[8] or about the terrorist tactic of derailing trains by a physical or cyber attack. Thus, with due caution, the experiment can be ranked as the ideal type of no pre-treatment. On the other hand, it can be argued that the treatment had as its topic only one of a number of completed or foiled Islamist attacks in the world. This is opposed to some extent by the character of some of the dependent variables measured in the domain of countermeasures, which were closely related to the Czech situation, and particularly linked to the specific character of the experimental treatment. Without knowing the results, it would seem that the study was more likely to have been conducted on a virgin population. Depending on which interpretation we choose, we need either to exclude or to consider the deflationary factors linked with repeated treatment. Furthermore, the two experimental

---

[6] Potentially the closest to such experience was an incident in December 2017, when an unknown perpetrator twice felled full-grown trees on a railway and left Islamic leaflets at the site. However, police investigation later uncovered that the perpetrator was a Czech senior citizen, a right-wing radical and member of the anti-immigration party, Freedom and Direct Democracy, who wanted to pin the crime on Islamists. He was later given four years in prison for a terrorist attack. At the time when the experiment was conducted, the general public were not yet well acquainted with the case.

[7] Fear of terrorism apparently increased in connection with the refugee crisis. In 2016, a terrorist attack was seen as the most urgent threat, and 81 percent of Czechs feared one.

[8] There were several cases where terrorist attacks were thwarted—all subsequent to the experiment.

groups were not necessarily contaminated to the same extent. While conventional terrorism, in which means of transport were involved and which claimed lives, was relatively frequent in Europe, cyber attacks were not linked with either transport or fatalities. Thus we would expect the group exposed to the stimulus of conventional terrorism to be somewhat more contaminated.

In terms of inflationary factors: in the pilot study, the mundane realism was evaluated as high by both security and media professionals, because they thought that a real-world treatment (commercial TV news clip about a foiled attack) would look very similar to that in the experiment. Experimental realism may be an Achilles heel of laboratory experiments, but mostly as a deflationary factor. In this particular case, realism was supported by the experimental script[9]; at the same time, it cannot be argued that the treatment was exaggerated, compared to the real world (it can be assumed that a foiled terrorist attack would make the main headline on Czech news). And likewise, the self-reports of emotions after watching the clip confirm that the experiment's measure of psychological realism was satisfactory, with treatment activating more negative emotions or feelings of threat in both experimental groups than in the control group, though the effect was not always statistically significant.

How, then, should we interpret the material result of the experiment—ultimately very similar to that achieved by the Gross et al. study that was replicated—with regards to the effects of interaction with the real world discussed above? The treatment effect in the group exposed to conventional terrorism stimulus presents the greatest uncertainty. It is not clear whether experimental subjects' prior experience of terrorist attacks on public transport and deep-rooted psychological processes, re-activated by the experimental treatment, were involved, or whether a virgin population not exposed to pre-treatment was examined. The differences between the conventional terrorism and control groups were, however, produced or reproduced by the effect of experimental treatment, because it can be assumed that not many subjects in the control group had been previously treated in such a way in the real world, and the questions used to measure the dependent variable had nothing to activate among these subjects, and nor did any other items in the experiment. This is one of the reasons that the effect measured in the two groups was probably not subject to a major deflation. The same conclusion can be made with the cyber terrorism group, where deflation can be ruled out with even greater certainty, given that exposure to this or similar treatment in the real world was even less likely. The differences between the two experimental groups can then be interpreted in two ways: either they are the result of the effect of treatment on a virgin population, and cyber terrorism really is seen as less disturbing than conventional terrorism; or, if we assume pre-treatment in the conventional terrorism group and the activation of an effect that had already been present in the real world, we can infer that experimental realism caused a slight deflation of the effect in the cyber terrorism group, which was less affected by pre-treatment. But cyber terrorism was not seen by subjects as harmless, and that is not questioned by either of these interpretations.

---

[9] For example, because the subjects saw not just the clip about the attack, but a section of an entire news broadcast.

5 Discussion and conclusion

As Baele and Thomson (2017) note, security studies should not trifle with people's responses and their concerns. If not considered, the contamination of experimental treatment by real-world treatment in security studies is a serious problem. Hence it needs to be addressed in the experimental design and in interpreting the results. The two case studies of the issue presented here may be remote from each other, yet they provide typical examples encountered in the study of countermeasures against threats. The Covid-19 study exhibited an extremely serious and multiple contaminations of subjects with real-world treatments. These, together with a number of other variables not included in the experiment, caused the subjects to enter the study with fixed values of the dependent variables, values that the experimental treatment was in no position to change. This degree of contamination makes it very difficult to isolate the independent variable as subjected to experimental manipulation. Paradoxically, when not just the occurrence but also the effect of the manipulation are significant in the real world, investigators then find nearly nil effect in the experiment (cf. Slothuus 2016). If, when designing the experiment, there are reasons to assume that such a situation is likely to arise, it is appropriate to include among the dependent variables such variables that would evaluate at least some characteristics of the treatment as if it were conducted on a virgin population. In the study presented, these were a direct evaluation by subjects of the informational value and manipulative nature of the treatment itself. Virtually all of the most politically and socially salient issues suffer from extreme contamination, and this poses a serious problem for study, which needs to allow for this aspect in its design.

The second of the studies presented investigated a situation where the countermeasures were concerned with a threat that was relatively new in terms of its substance and location. In this case, consideration of the effect of experimental manipulation is particularly complicated by the uncertainty about the measure of contamination from the real world (both in terms of the frequency and the longevity of the effect), and this is not necessarily the same for all the experimental groups. Furthermore, the problem concerns not just the deflationary factors, but also those linked with experimental realism and the possible inflation of the effect. A major complication is that virtually none of these effects can be measured exactly; they can only be qualitatively estimated and the resulting assessment of the size of the effect of the experimental treatment can be ordinal at best.

In this paper we presented a broader understanding of the factors linked with realism than Gaines et al. (2007), and we did not limit ourselves to inflation but also considered the deflation of the experimental treatment effect. Their qualitative assessment should become standard in experimental protocols. The summary of all deflationary and inflationary factors not only increases certainty regarding the size of the experimental treatment effect, but can be put to broader use. The logic informing this thinking can be used, for example, in international studies or to evaluate replication experiments.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Aktualne.cz: Strach z terorismu mezi Čechy roste, útoku se bojí 81 procent lidí. Economia. https://zpravy. aktualne.cz/domaci/obavy-z-terorismu-mezi-cechy-rostou-odhalil-vyzkum-verejneho/r~5e7df856b8 5811e590ec0025900fea04/ (2016). Accessed 26 June 2021

ANOPRESS–Databáze českých médií. NEWTON Media. https://www.anopress.cz/ (2021)

Aronson, E., Wilson, T.D., Akert, R.M.: Social Psychology: The Heart and the Mind. Harpercollins, New York (1994)

Baele, S.J., Thomson, C.P.: An experimental agenda for securitization theory. Int. Stud. Rev. **19**, 646–666 (2017). https://doi.org/10.1093/isr/vix014

Bromberg, D.E., Charbonneau, É., Smith, A.: Public support for facial recognition via police body-worn cameras: findings from a list experiment. Gov. Inf. Q. **37**, 158 (2020). https://doi.org/10.1016/j.giq. 2019.101415

Curini, L.: Corruption, Ideology, and Populism: The Rise of Valence Political Campaigning. . Palgrave Macmillan, Basingstoke (2018)

CVVM: Naše společnost, v21–06. Sociologický ústav AV ČR. https://cvvm.soc.cas.cz/media/com_form2 content/documents/c2/a5401/f9/ob210715.pdf (2021). Accessed 26 July 2021

Dokument: Druhý projev premiéra Babiše v době koronaviru. Přečtěte si, co řekl před Velikonocemi. iRozhlas. https://www.irozhlas.cz/zpravy-domov/koronavirus-v-cesku-andrej-babis-chytra-karantena-pro-jev_2004092018_pj (2020). Accessed 26 June 2021

Druckman, J., Kam, C.: Students as experimental participants. In: Druckman, J., Green, D.P., Kuklinski, J.H., Lupia, A. (eds.) Handbook of Experimental Political Science, pp. 41–57. Cambridge University Press, New York (2011)

Druckman, J., Leeper, T.: Learning more from political communication experiments: pretreatment and its effects. Am. J. Polit. Sci. **56**, 875–896 (2012). https://doi.org/10.1111/j.1540-5907.2012.00582.x

Fesenfeld, L.P., Sun, Y., Wicki, M., Bernauer, T.: The role and limits of strategic framing for promoting sustainable consumption and policy. Glob. Environ. Change (2021). https://doi.org/10.1016/j.gloenvcha. 2021.102266

Gaines, B., Kuklinski, J., Quirk, P.: The logic of the survey experiment reexamined. Polit. Anal. **15**, 1–20 (2007). https://doi.org/10.1093/pan/mpl008

Garcia, B.E., Geva, N.: Security versus liberty in the context of counterterrorism: An experimental approach. Terrorism and Political Violence. **28**, 30–48 (2016). https://doi.org/10.1080/09546553.2013. 878704

Goel, S., Williams, K.J., Huang, J., Warkentin, M.: Can financial incentives help with the struggle for security policy compliance? Inf. Manag. (2021). https://doi.org/10.1016/j.im.2021.103447

Gross, M.L., Canetti, D., Vashdi, D.R.: The psychological effects of cyber terrorism. Bull. at. Sci. **72**, 284–291 (2016). https://doi.org/10.1080/00963402.2016.1216502

Haynes, L., Service, O., Goldacre, B.: Test, Learn, Adapt: Developing Public Policy with Randomized Controlled Trials. Behavioural Insights Team, London (2012)

Huddy, L., Feldman, S., Capelos, T., Provost, C.: The consequences of terrorism: disentangling the effects of personal and national threat. Polit. Psychol. **23**, 485–509 (2002). https://doi.org/10.1111/0162-895X.00295

Iyengar, S.: Laboratory experiments in political science. In: Druckman, J., Green, D.P., Kuklinski, J.H., Lupia, A. (eds.) Handbook of Experimental Political Science, pp. 73–88. Cambridge University Press, New York (2011)

John, P.: Policy entrepreneurship in UK central government: the behavioural insights team and the use of randomized controlled trials. Public Policy Adm. **29**, 257–267 (2013). https://doi.org/10.1177/09520 76713509297

Jones, R., Whitehead, M.: 'Politics done like science': critical perspectives on psychological governance and the experimental state. Environ. Plan. D Soc. Space **36**, 313–330 (2018). https://doi.org/10.1177/ 0263775817748330

Kabrhelová, L.: Antropolog Samek: Komunikace české vlády oproti té německé je jako rozbouřené moře. iROZHLAS. https://www.irozhlas.cz/zpravy-domov/podcast-vinohradska-12-lingvista-samek-koron avirus-cesko_2102160600_cen (2021). Accessed 26 June 2021

Krueger, B.S., Best, S.J., Johnson, K.: Assessing dimensions of the security-liberty trade-off in the United States. Surveill. Soc. **18**, 104–120 (2020). https://doi.org/10.24908/ss.v18i1.10419

Kruglanski, A.W.: Motivation and social cognition: Enemies or a love story? Int. J. Psychol. Psychol. Ther. **1**, 33–45 (2001)

Lecheler, S., de Vreese, C.: What a Difference a Day Makes? The effects of repetitive and competitive news framing over time. Commun. Res. **40**, 147–175 (2013). https://doi.org/10.1177/0093650212470688

Lerner, J.S., Gonzalez, R.M., Small, D.A., Fischhoff, B.: Effects of fear and anger on perceived risks of terrorism: a national field experiment. Psychol. Sci. **14**, 144–150 (2003). https://doi.org/10.1111/1467-9280.01433

Linos, K., Twist, K.: Diverse pre-treatment effects in survey experiments. J. Exp. Polit. Sci. **5**, 148–158 (2018). https://doi.org/10.1017/XPS.2017.29

Lunz, T.K., Motta, M., Callaghan, T., Sylvester, S.: Correcting misperceptions about the MMR vaccine: using psychological risk factors to inform targeted communication strategies. Polit. Res. q. **74**, 464–478 (2021). https://doi.org/10.1177/1065912920907695

Makkonen, A., Oksanen, A., Kushner, G.S., Herreros, F., Winsvold, S., Bugge, S.Ø., Enjolras, B., Steen-Johnsen, K.: Fear-triggering effects of terrorism threats: cross-country comparison in a terrorism news scenario experiment. Personal. Individ. Differ. (2020). https://doi.org/10.1016/j.paid.2020.109992

Rivers AM, Sherman J: Experimental design and the reliability of priming effects: reconsidering the "Train Wreck". PsyArXiv. (2018). https://doi.org/10.31234/osf.io/r7pd3

Rogers, R.W.: A protection motivation theory of fear appeals and attitude change. J. Psychol. **91**, 93–114 (1975). https://doi.org/10.1080/00223980.1975.9915803

Schachter, S.: The Psychology of Affiliation: Experimental Studies of the Sources of Gregariousness. Stanford Univer Press, Redwood City (1959)

Šenk, M.: Expert na krizovou komunikaci Charvát: Vládě, za to, jak v této krizi informuje, bych dal kuli. E15.cz. https://www.e15.cz/koronahelpdesk-e15/expert-na-krizovou-komunikaci-charvat-vlade-za-to-jak-v-teto-krizi-informuje-bych-dal-kuli-1374023 (2020). Accessed 26 June 2021

Slothuus, R.: Assessing the Influence of political parties on public opinion: the challenge from pretreatment effects. Polit. Commun. **33**, 302–327 (2016). https://doi.org/10.1080/10584609.2015.1052892

Slovic, P.: Terrorism as a hazard: a new species of trouble. Risk Anal. **22**, 425–426 (2002). https://doi.org/10.1111/0272-4332.00053

Sulitzeanu-Kenan, R., Zohlnhöfer, R.: Policy and blame attribution: citizens' preferences, policy reputations, and policy surprises. Polit. Behavavior **41**, 53–77 (2019). https://doi.org/10.1007/s11109-017-9441-5

Thaler, R., Sunstein, C.: Nudge: Improving Decisions About Health, Wealth and Happiness, Newhaven. Yale University Press, CN (2008)

Vaverka, J.: Analýza politiky České republiky proti nemoci COVID-19. Unpublished manuscript of master's thesis. https://is.muni.cz/auth/th/w35ut/Vaverka_Diplomova_prace_COVID_19.pdf (2021). Accessed 26 June 2021

Wilson, T.D., Aronson, E., Carlsmith, K.: The art of laboratory experimentation. In: Fiske, S.T., Gilbert, D.T., Lindzey, G. (eds.) Handbook of Social Psychology, pp. 51–81. Wiley, Hoboken (2010)

Yan, J., Kniffin, K.M., Kunreuther, H.C., Schulze, W.D.: The roles of reason and emotion in private and public responses to terrorism. J. Econ. Behav. Organ. **180**, 778–796 (2020). https://doi.org/10.1016/j.jebo.2019.05.024

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.