# Political Solidarities in Novaland:

# Can we Simulate the Experience of States, Economies and Public Policies in a Virtual Online State?

This Version 17 January 2023

Achim Goerres

Professor of Empirical Political Science, www.achimgoerres.de

Jan Karem Höhne

Incoming Assistant Professor University of Hannover, www.jkhoehne.eu

Jakob Kemper

Research officer and PhD student of POLITSOLID, www.jakob-kemper.eu


University of Duisburg-Essen, Department of Political Science, project homepage https://bit.ly/politsolid

Abstract

What if we could experimentally manipulate all characteristics of states, economies and public policies and estimate their effects on citizens? This paper puts forward the first evidence from a pilot study of Novaland. Novaland is a virtual liberal democracy that only exists online and that has characteristics realistically drawn from German and Romanian welfare states. The pilot consists of an experimental online platform based on text and images in which volunteers are surveyed after they have exited the experience, are randomly assigned to different experiences, such as defined by income, corruption or unemployment, interact with each other simultaneously and thereby co-create collective decisions, such as elections or donation pools, that then determine the course of Novaland.

Our main results are: (1) The newly programmed Novaland Experience worked technically well in that over 300 participants could simultaneously interact online at the same time. (2) Participants behaved in an internally valid manner, even though there are some signs of inattentiveness. (3) The political ideology of volunteers was reflected in how they behaved in Novaland, whereas socio-demographic variables showed no systematic effects. (4) The assigned income and unemployment treatments worked in that they causally changed the way in which the participants behaved towards others and towards politics and policies in Novaland with unclear evidence as to the corruption experience. Overall, this data collection exercise is encouraging as it allows us to explore the potential for relevant social science research in virtual worlds in the 2020s.

## 1. Introduction

What if we could experimentally manipulate all characteristics of states, economies and public policies and estimate their effects on citizens in a valid manner? In this paper, we present a pilot study on the artificial state "Novaland" that was conducted with more than 300 participants of the German adult population interacting online on 9 November 2022 for one hour.

In this pilot, volunteers were exposed to a virtual state that had features of the German and the Romanian welfare sates. Novaland is described to be a functioning liberal democracy with an extensive welfare state in which instances of bad governance and corruption can materialise. Participants did not experience one Novaland, but different versions of Novaland depending on their assigned treatment groups. Participants had to assess given situations, express opinions and decide on multiple occasions during the experience, from which we draw the main outcome variables. Participants co-created two collective outcomes that then affect all of Novaland: they voted in a four-party election of artificial parties and they donated virtual money when a natural catastrophe hit, which was then redistributed across all participants.

The treatments in Novaland deal potentially with government structures (e.g., form of democracy and election), government performance (e.g., degree of corruption), social composition (e.g., share of demographic and income groups), and exogenous events (e.g., environmental catastrophes and crisis). For the very first time, we can experimentally test causal chains that could not be tested in political science research before.

The pilot was driven by two major motivations. **Methodologically**, we want to test whether the simultaneous interaction of several hundred people in an online environment programmed in oTree and Python worked. This is not trivial because interactions mean technically that information between participants and the server had to be transmitted many times in a reliable manner. Also, we wanted to assess to what extent our measurements in a purely artificial world are reliable, internally and externally valid. **Theoretically**, we are interested in political solidarities, i.e. people's multiple levels of willingness to shoulder costs by public redistribution that does not benefit themselves. The pilot (experimentally) investigates governmental, social, and financial factors influencing people's political solidarities (Goerres 2021), concentrating on income, corruption experience and unemployment as assigned treatments.

Overall, this data collection exercise worked well despite some programming error. It is encouraging as it allows us to explore the potential for relevant social science research in virtual worlds in the 2020s.

The rest of the paper is organised as follows. Section 2 gives an overview of the Novaland Experience. Section 3 lays out our methodological motivation. Section 4 describes the substantive theoretical motivation. Section 5 presents the data collection. Section 6 has the empirical results, before we conclude in section 7.

## 2. The Novaland Experience: An Overview

We have created a fully online environment in which volunteers interacted anonymously for about one hour. In this virtual world, volunteers are citizens of Novaland, a non-existing state that has features drawn from German and Romanian welfare states. Volunteers experience several, uncorrelated randomly assigned experiences spread out across the one-hour experience: income (low-medium-high), corruption in the healthcare sector after having a minor accident (yes, no), unemployment with partial loss of net income (yes, no) and exposure to a natural disaster leading to a loss of all assets (yes, no). Participants co-produce several collective outcomes: two donation pools during the experience, one before they are affected while already knowing about the disaster and one after the assignment of the natural disaster, to be distributed among the volunteers who are victims of the natural disaster. The other

collective outcome is a parliamentary election with two Lager, four parties (socioeconomically: two left, two right; one more pro-immigration than the other) and pre-election coalition pledges.

Participants spend Novas, the currency of Novaland, on various activities, such as housing, food and leisure. These activities serve two purposes: to make it easier for participants to immerse themselves in the Novaland Experience and for us to validate that the experimental manipulation of virtual income works. Three treatments will be the focus of the paper here: that of net income (low income: 1400 Novas, medium income: 2000 Novas, and high income: 3500 Novas), that of corruption and that of unemployment. The corruption experience is assigned after the experience of an accident when half of the sample have the chance to bribe the hospital doctor to receive higher quality medical treatment. The unemployment experience is assigned to half of the sample and



*Figure 1: Screenshot of the Novaland election*

results in immediate wage replacement by 50 % of the last wage as unemployment benefit that is taxed again by 30 %. The identity of the participant is kept constant as he or she is a 51-year old employee.
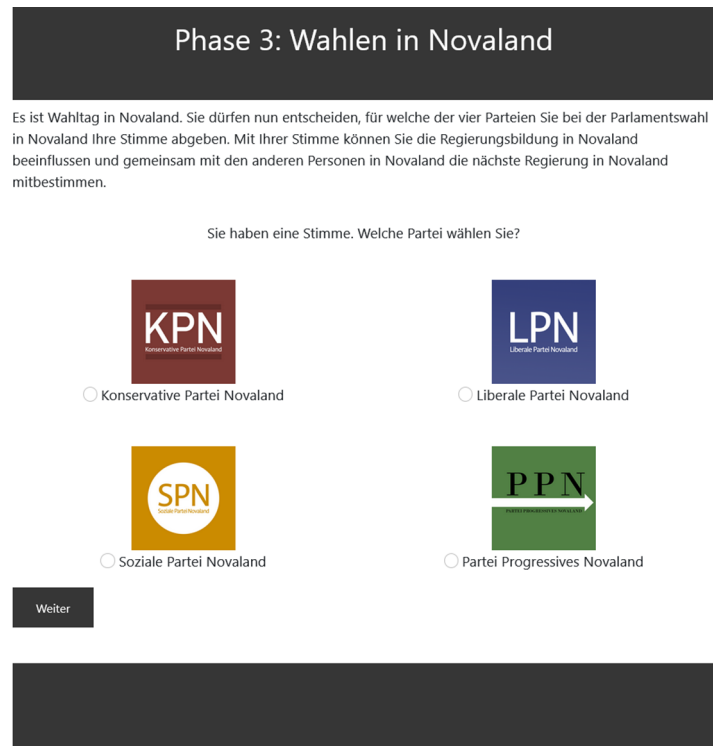
Novaland is a "low-immersive virtual environment" in which volunteers get overview kind of information. Information is mostly conveyed in the form of text (e.g., instructions are provided in written language), but it could also be conveyed in the form of visual (e.g., images) and audio inputs (e.g., speeches) and even avatar-based interviewing. The online interface of the Novaland pilot is programmed via the open-source platform oTree that is built in the Python framework. oTree has been developed by behavioural economists to allow the implementation of interactive economic games in the online world (Chen, Schonger, and Wickens 2016). Novaland works on any internet browser (e.g., Google Chrome and Mozilla Firefox) and device type (e.g., PC/laptop, tablet, and smartphone). Participants can be anywhere on the world as long as they have internet access. Info transfer between server and device is minimal.

We have two runs of the experience. One purely technical with 31 volunteers in May 2022 and one substantive with 324 volunteers on 09 November 2022. In this paper, we mainly report evidence for the latter. Participants were recruited through a social media campaign. Participants had to register for participation in advance and received 12 € compensation through PayPal and had a chance to donate the money for a national food bank organisation. Both data collections were pre-approved by an ethics review board at the University of Duisburg-Essen.

Figure 1 shows a screenshot from the Novaland data collection. This shows the election. The look of the page was designed in a way that it would work on all participant devices and platforms, most importantly mobiles, tablets and laptops. The commissioned programmer is a web designer and impleme nted colour and font schemes that captured a lean and professional

look. Figures 2 and 3 (left and right) show the three treatment conditions low-income, corruption and unemployment

*Figure 2: Treatment condition low-income*

## Phase 2: Ihre persönliche Situation in Novaland

Herzlich willkommen zurück in Novaland. Nun geht es weiter um Ihre persönliche Situation in Novaland.

Sie haben ein monatliches Brutto-Einkommen vor Steuern und Sozialabgaben von **2000 Novas**. Damit ist ihr Einkommen **unterdurchschnittlich** hoch.
Auf dieses Einkommen müssen Sie Steuern und Sozialabgaben in Höhe von 30 % zahlen, sodass Ihnen danach ein monatliches Netto-Einkommen von **1400,0 Novas** bleibt. Etwa **20 %** der Bevölkerung von Novaland hat ein geringeres Netto-Einkommen als Sie.

Von Ihrem Einkommen müssen Sie Ihre monatlichen Ausgaben bezahlen und können gegebenenfalls etwas sparen. Derzeit haben Sie noch keine Ersparnisse.

Weiter

*Figure 3: Screenshots of two treatment conditions (corruption = yes on the left; unemployment= yes on the right)*

## Phase 2: Alltagssituationen in Novaland

Auf dem Heimweg von der Arbeit stürzen sie unglücklich und verletzen sich am Fuß.
Vor der Behandlung im Krankenhaus fordert Sie ein Arzt dazu auf, eine „persönliche Bearbeitungsgebühr" in Höhe von **500 Novas** in bar an ihn zu zahlen.
Ansonsten könne die optimale Versorgung nicht garantiert werden.

Zahlen Sie dem Arzt die „persönliche Bearbeitungsgebühr" in Höhe von **500 Novas**?

○ Ja
○ Nein

## Phase 2: Arbeitslosigkeit

Die Firma, für die Sie arbeiten, muss leider Insolvenz anmelden und kündigt Ihnen. Dadurch werden Sie **arbeitslos.**
In Novaland bekommen Sie die Hälfte Ihres letzten Bruttoeinkommens als Ersatzleistung während Ihrer Arbeitslosigkeit. Diesen Betrag müssen Sie dann noch mit **30 %** versteuern.

Ihr letztes Bruttoeinkommen betrug **2850,0 Novas**.

Deswegen beträgt die Ersatzleistung brutto **1425,0 Novas**.

Wovon Ihnen nach Abzug der Steuern **997,5 Novas** bleiben.

Etwa **20 %** der Bevölkerung von Novaland hat ein geringeres Netto-Einkommen als Sie.

Weiter

## 3. Methodological Motivation

The problem that we face in existing empirical analyses of political solidarities is often that measurement is not very thorough or under-complex (Goerres and Höhne forthcoming [2023]). Simple measures of welfare state attitudes or measures of political trust are often used to approximate such ideas of political solidarities. One line for improvement is to explore the possibilities for more causal leverage by using experimental techniques in an innovative way by varying characteristics in an artificial world that cannot be varied in the real world. This is what we attempt in this paper.

There are three bodies of knowledge that we can build on: hypothetical decisions in surveys and experience sampling, text-based experimental designs and experiments in virtual environments.

### *Hypothetical Decisions and Experience Sampling in Surveys*

Novaland is a primarily text-based virtual environment. It is thus situated between surveys and survey-type measurements and immersive virtual environments. Asking individuals to assess something hypothetical is, for instance, the bread-and-butter approach in vignette analysis where individuals are confronted with a deck of N vignettes in which vignette variables of interest are varied. Most of the realised values of these variables are new to the individuals. For instance, Busemeyer and Goerres (2020) ask respondents to assess the fairness of concrete vignettes of families with a kindergarten child. Respondents are exposed to situations that they have never encountered or encountered only in a subconscious way. Vignette analysis are thus, to a limited extent, artificial with regard to the situation even though the context does not get manipulated. The validity of vignettes in surveys to predict real-life behaviour is thus unclear (Eifler and Petzold 2019).

The survey world also knows the technique of experience sampling or the experience sampling method (ESM). ESM is a frequently used measurement technique in psychology as well as behavioural and health-related research that is gaining importance in social science research. It allows researchers to measure people's behaviours, feelings, and thoughts during day-to-day activities by signalling participants following specific sampling protocols (Höhne 2020). For example, participants receive study invitations or information via email or SMS (Short Message Service) at random times during the day. Such a "signal-contingent" sampling protocol has proven its worth when investigating ongoing experiences (Christensen et al. 2003) and thus it meets the purposes of Novaland. ESM is open to different field periods ranging from few days to several weeks (van Berkel, Ferreira, and Kostakos 2018). Even though ESM comes with some methodological merits (e.g., studying people in-situ), it suffers from low response rates and high break-off rates. This can reduce the generalizability of the results. However, research has shown that the provision of incentives to participants substantially increases response rates and decreases break-off rates (Gabriel et al. 2019).

### *Text-based Experimental Applications in the Social Sciences*

We have come across one political science application in which a text-based online adventure is operationalised experimentally as a psychological intervention in which participants are subjected to changing their perspective of a minority group by going through a role-play adventure as a member of that group (Simonovits, Kézdi, and Kardos 2018). The outcome variables measure prejudice towards that group.

Jetten et al. (2017) conducted an experiment with MTurk workers by introducing them to a simple world of virtual Mambiza with a certain kind of social stratification. They then randomly assigned the volunteers to one out of three income groups, had them spend money on various items and then confronted them with different newspaper clippings describing different levels of economic security in Mambiza. After that, various dependent variables were measured to

see whether the (income group X level of insecurity) treatments had an effect on various attitudes (see also Tanjitpiyanond, Jetten, and Peters 2022).

Text-based studies from political psychology on information-processing employ the Dynamic Process Tracing Environment. Experimental subjects had to follow a moving online screen of self-scrolling text of info items to discover relevant information in an artificial campaign. Participants do immerse themselves in the world and take an interest in the outcome of the artificial campaign, even long after their involvement has ended (Lau & Redlawsk, 2006). The underlying Dynamic Process Tracing Approach had been used in other areas, such as for studying consumer behaviour in complex information environments (Groenland, Kuylen, and Bloem 1996).

All of these studies report estimates that are plausible and that in the case of the first study cited actually has a real-life impact.

### *Social Science Experiments in Virtual Environments*

Novaland draws on commercial, government simulation games (e.g., The Political Machine 2016 or Democracy 3), applications of offline role-play teaching in which participants are citizens (Shellman 2001), and on online simulations in international relations teaching in which participants are the policy-makers (Stover 2005). Principles of early text-based computer games in the 1980s and 1990s – also known as "text adventures" – and online political science experiments (Del Ponte and DeScioli 2019) are applied.

In 2007, Bainbridge described the potential of virtual world research (Bainbridge 2007). There have been attempts with using virtual worlds before, but a systematic field has not been established in the social sciences[1] so far.

Behavioural economists have explored the possibility of conducting experiments in virtual worlds. They saw the advantage of the controlled lab-like environment with the possibility of diversifying the subject pool (Atlas 2008). Virtual reality experiments are "framed field experiments" that allows the introduction of controlled context for economic decision-making. (Innocenti 2017). Virtual reality is, for instance, used to manipulate time perception and thus future discounting (Faralla et al. 2021).

Second Life, a commercial platform with 3D-virtual realities that can be used for social and commercial activities, promised to have great potential for the social sciences in the nought years. It seemed to make it easier to create environments that imitate real environments in which individuals need to act by adapting rationally to the environment (Innocenti 2017). Cooperation is higher in Second-Life settings of a dictator game than in the real world. An overview of five replicable standard instruments yielded similar results in Second Life to those in the lab (Chesney, Chuah, and Hoffmann 2009; Duffy 2011). Second Life, however, was and still is strong at allowing participants to be exactly what they are not in the real world. Outside experimental economics, there are hardly any applications in the social sciences to be found on Web of Science (search conducted on 20 November 2022).

Psychologists demonstrated that even complex social behaviours like helping and the bystander-effect can be emulated in a virtual environment (Kozlov and Johansen 2010). Another application is the moral trolley problem where participants have to decide to push a trolley and kill one person to save many people (Navarrete et al. 2012). Along similar veins, "moral machine" is a public platform created by Edmond Awad at MIT on which millions of volunteers decide, based on text and two-dimensional pictures, on moral decisions that arise

---

[1] There is a vibrant virtual reality field at the various intersections between computer science and many sciences, see e.g. Slater and Sanchez-Vives (2016).

around new AI developments in technology, such as autonomous driving (Awad et al. 2018). Self-recruited volunteers make a pairwise comparison like in an online conjoint experiment.

What is surprising is that the use of simple environments based on text and simple graphics does not seem to impact the results and that the platform that seems most easy to access for researchers, Second Life, did not take off. This is conjecture, but it may be that volunteers are used to such high-quality graphics in computer games and online animations that even the only slightly cruder Second Life is not good enough to compete. Thus, the back-to-the-basics approach of simple text and graphics may make it easier to participate for individuals with 24/7 access to high-performance graphics everywhere.

### Deriving an Assessment Framework for Novaland

Given that political science and sociology are meagre in virtual world studies, one might take a sceptical stance here and critically ask whether this kind of research does not produce artefacts or non-sensical results. It may be helpful to look at some major points one by one.

*Internal validity 1 (The perception of the platform varies between volunteers)*

Optimally, the participants should interpret all stimuli from the Novaland Experience in the same way, plus or minus some random effect that has a mean of zero. The information, the layout and other factors should be the same on average. As a first measure, we extensively pretested the Experience in countless rounds.

*Internal validity 2 (Volunteers do not take the experience seriously)*

Participants know that they interact with real people in a fully artificial setting. In the current version, the pay-out is not tied to in-experience behaviour and we will experimentally test whether changing that will make a difference in the future. Thus currently, participants only have to comply with answering all our requests for decisions and questions until the end, and will then receive the pay-out, no matter how they behaved. There is the danger of volunteers not behaving sincerely. This can take various forms. Most importantly, we fear declining attention during the online experience.

*Replicability*

The same Novaland setup needs to reproduce similar causal estimates on any set of volunteers. This means that the expectations about the quantities of interest can be pre-registered, and they can be verified in any sample of volunteers. This is not something that we can test with the current level of available evidence yet as we have not replicated the same version of Novaland so far.

*External validity 1 (Demand effects)*

It is a common problem of experiments in which volunteers fully consented to participation (different from field experiments where informed consent is rarely sought) that participants can behave in a way that they think their experimenters expect them to behave. This arises from unconscious shifts in behaviour triggered by the knowledge of being in an experiment. Demand effects are not restricted to experiments, but take shape in interviews and surveys whenever there is a real or imagined perception of expectation by the participants (Orne 2017 [1970]). These demand effects are in general modest (for political science experiments see Mummolo and Peterson 2019; Quidt, Haushofer, and Roth 2018). We cannot distinguish whether, for instance, low-income-treatment individuals behave differently because they assume they must behave like that in the experience or because this is their "natural reaction". However, it is implausible that besides money-spending activities that we mainly employ for engaging participants, participants can anticipate what our expectations were (e.g. on voting in the virtual elections).

*External validity 2 (Real-life attitudes and behaviour)*

In a perfect research world, some of the findings from Novaland should translate into real-world patterns of how people think and behave. One thing that we can look at is to compare how participants self-rate on out-of-experience variables, such as attitudes towards public redistribution and their in-experience behaviour. Especially those characteristics that are deeply engrained in humans, such as their political ideology, should also be reflected in the way that people behave in Novaland.

This sub-chapter has given an overview of existing lines of research in the survey world, text-based experimental applications and social science experiments in the virtual world. Then, we put forward an assessment framework that we will use in order to quality-check to what extent the behaviour in the Novaland experience is meaningful. Let us now turn to our theoretical motivation for developing Novaland.

## 4. Substantive Theoretical Motivation

Even though the approach allows all kinds of experimental manipulations and substantive applications, we are interested in explaining political solidarities in advanced industrial democracies. Political solidarities are multiple individual-level kinds of willingness to support public redistribution for other individuals than oneself (Goerres 2021). The concept in the plural reflects the observation that the willingness to shoulder public costs by the same individual varies dependent on two very different systems.

On the one hand, support for redistribution is contingent on the level of the polity on which redistribution takes place. Citizens in multiple-level systems can differentiate between the different levels of the political system and can assess whether they trust the system to do the job and whether they concur with that level being the appropriate level for redistribution.

On the other hand, political solidarity is shaped by perceptions towards the potential recipients of a certain kind of redistribution. In a context where citizens are increasingly socially diverse by descent, the likelihood of recipients being constructed to be like oneself decreases (Cavaillé and Trump 2015). We focus on the three lines of identification age, income and ethnicity/descent, but only vary income in the current version of Novaland.

Political solidarities are thus a multi-dimensional phenomenon that is worthy of our attention. Some people argue that such solidarities are not necessary in a functioning liberal, representative democracy as long as people pay their taxes and elections are free and fair (Levy 2017). Others would argue that political solidarities are important, but only arise at certain neuralgic junctures in history, like after wars or natural catastrophes (Bauer et al. 2016; Titmuss 2019 [1958]). These higher levels of political solidarities can then be enshrined in new or altered institutions that retain their underlying function of solidarity even when the collective effort to support one another in a publicly legitimised structure has subsided.

There is an almost endless number of determinants that influence political solidarities. Some reflect learned behaviour, personality traits or even genetic predisposition interacting with the social environment. The problem with identifying the causal impact with many of them is that they are confounded with third factors. For example, the economic situation of a person is influenced by job experience, occupation, education, economic opportunities and other factors. Thus, if we want to know the impact of income on a person's political solidarities, observational studies make it hard to separate these different aspects. Therefore, we turn towards other research designs that can help us to get better leverage on potential causal effects.

In the current version, we focus on three determinants that we can randomly assign: income, the experience of corruption (Rothstein and Eek 2009) and the experience of unemployment

(Naumann, Buss, and Bähr 2016). Both are two broad classes of experiences that have been studied extensively and that we can import into the virtual world of Novaland to assess the validity and plausibility of our findings there. In order to estimate the causal impact of these variables on measures of political solidarities and related concepts, we use a number of outcome variables in the Novaland Experience that we are interested in. In the current set-up, we will look at donations to others in the expectation of wildfire damage, voting in a four-party election, satisfaction with the election outcome, satisfaction with governmental programme and satisfaction with the personal situation. The broad expectations are that higher-income people move to the right, become more prosocial towards other people and are more satisfied. The experience of corruption and unemployment make people less dissatisfied and less prosocial towards other people.

## 5. Data and Methods

### Data

Novaland is programmed on the oTree platform that is based on Python. We started by developing storyboards for the development of the Novaland Experience and then commissioned its programming to a professional programmer.

The main data collection that we present here was conducted online on 09 November 2022 with a convenience sample that was mainly recruited via a commercial campaign through social media platforms, such as Instagram, Facebook and Twitter. Participant recruitment started on 31 October 2022. Participants were required to register for the study until 6 pm on 8 November 2022. Participation took about one hour and participants received an incentive in the amount of 12€ that was paid via PayPal (about 8% of the participants donated their incentive to "Tafel Deutschland"). On data collection day, an email with individualised access tokens was sent out to all registered participants. The data collection is based on a data management plan agreed upon with the university's data protection officer. An ethics approval was granted by an ethics review board of the University of Duisburg-Essen.

### The Basic Structure of the Novaland Experience

Participants experienced their own, unique version of Novaland as its citizens. Even though Novaland and its phases are conceptualised by us, the actual course remained open until the end and depended on participants' decisions. Participants received monthly income in Novas (explained to be at parity around 1 Euro), had living expenses (e.g., housing and food), participated in an election, experienced a natural catastrophe (i.e., forest fire), and could donate from their savings in Novaland. In total, there were five subsequent Novaland phases (the Appendix includes exemplary screenshots of Novaland). Each phase represented one month in Novaland. Table 1 presents the five Novaland phases and their scope.

*Table 1: Phases of the Novaland Experience*

| Phase | Period | Scope |
|---|---|---|
| 1 to 3 | 7.00 to 7.29 pm | Phase 1: Introduction and story telling |
| | | Phase 2: Personal characteristics and life |
| | | Phase 3: Election |
| 4 | 7.30 to 7.39 pm | Election outcome and catastrophe |
| 5 | 7.40 to 8.00 pm | Effect of the catastrophe including post-survey |

Participants' experience in Novaland ended with a post-survey. This survey included questions on demographics (e.g., age, education, and gender), political attitudes (e.g., towards welfare chauvinism and governmental scope), social trust (e.g., trust in fellow citizen and institutions), and participant evaluations (e.g., on interest, difficulty, and sensitivity).

There were four random assignments planned sequentially in the data collection, all with equal group sizes and orthogonal to one another in a 3 X 2 X 2 X 2 design. Table 2 illustrates the treatments.

*Table 2: Overview of treatments in Novaland*

| Order and group numbers | Content | Characteristics |
| --- | --- | --- |
| 1st treatment (3 groups) | Net monthly income | 1400 Novas, 2000 Novas, or 3500 Novas |
| 2nd treatment (2 groups) | Corruption experience during a medical treatment | Option to bribe for 500 Nova, no corruption |
| 3rd treatment (2 groups) | Unemployment | Unemployment due to insolvency of the company (drop in net income), no unemployment |
| 4th treatment (2 groups) | Losing all assets due to a wildfire | Losing all assets due to a wildfire, no loss |

Unfortunately, there was a programming error leading only to a fully randomised assignment of the income group. The 2nd to 4th treatment were assigned, but in a fully correlated way. One group received no corruption – unemployment – wildfire.

In order to gain some insights, we will describe the no corruption – unemployment group against the corruption – unemployment groups and compare them against the May results for the same variables as an artificial control group.

### *Samples*

For this study, we mainly used data from a study that was conducted online on 9 November 2022. The study took place between 7 and 8 pm CET. Overall, 529 persons registered to take part in the study, out of which 347 persons showed up and started (another 15 participants showed up late and were not allowed to start). Of these, 324 completed the study. Thus, we have a dropout rate of about 6%. Participants had a mean age of about 41 years and about 45% of them were female. In terms of education, about 3% graduated from a lower secondary school, 17% from an intermediate secondary school, and 80% from a college preparatory secondary school. Overall, about 28% of participants used a smartphone, 11% a tablet, and 62% a computer device.

As a comparison, we also resort to a small feasibility study from 09 May 2022. The study took place between 7 and 8 pm CET. Overall, 48 persons registered to take part in the study, out of which 31 persons showed up and started. Of these, 29 completed the study. Thus, we have a dropout rate of about 6.5 %. Participants had a mean age of about 27.2 years and about 51.6 % of them were female. In terms of education, 0 % graduated from a lower secondary school, 3.2 % from an intermediate secondary school, and 96.7 % from a college preparatory secondary school. Overall, about 3.4 % of participants used a smartphone, 3.4 % a tablet, and 93.1 % a computer device.

## 6. Empirical Results

We will first give an overall assessment of the quality of data collection and then, second, review our evidence in the realm of internal and external validity before we turn to our substantive results.

### General Assessment of Data Collection

Overall, the data collection worked technically well. The programmed platform was capable of hosting more than three hundred participants online at the same time. The number of incompletes was small (13 out of 323). We experienced some problems with the randomisation as already foreshadowed before. The random assignment of income in three groups worked well, the others did not. This kind of mistake may be typical of future mistakes in the formation of the virtual world. As everything needs to be designed and programmed, the risk of a small discrepancy jeopardising part of the experience and thereby of the data collection is large.

### Internal Validity: Comprehension Checks and Open Comments

In order to infer participants' commitment during their time in Novaland we included five comprehension checks. These checks asked participants questions on the following aspects: 1 citizenship in Novaland, 2 right to vote, 3 governmental financing, 4 income in Novaland, and 5 income in relation to others. The first three questions were asked in the first phase and the last two questions were asked in the second phase. Table 2 presents the results of the comprehension checks.

*Table 3: Outcome of the comprehension checks*

| Comprehension check | Correct (%) | Incorrect (%) | Nonresponse (%) |
|---|---|---|---|
| 1 | 92.6 | 4.3 | 3.1 |
| 2 | 91.4 | 2.8 | 5.9 |
| 3 | 93.8 | 1.5 | 4.6 |
| 4 | 80.6 | 12.0 | 7.4 |
| 5 | 78.4 | 17.6 | 4.0 |

Note. Because of rounding the percentages may not add up 100%.

As shown in Table 2, the majority of participants successfully pass the comprehension checks indicating that they pay close attention to the information provided in Novaland. However, a substantially higher number of participants passes the checks in the first phase (>90%) than the checks in the second phase (> 78%). This suggests that participants either got inattentive over time or that the comprehension checks in the second phase were more difficult. Considering the fact all comprehension checks were placed relatively close to the beginning of Novaland we find the latter explanation more convincing. Non-response to the comprehension check items varied between 3% and 8%. The evidence reflects some variation as to the perception of the Novaland Experience. Some people were not concentrated enough to answer all questions correctly. This hints at the problem of inattentiveness in this purely online study where compliance with rules, such as concentration and commitment, cannot be controlled by us.

At the end of the study, participants were invited to leave open comments. Out of the 323 finishers, 163 (50.5 %) left a comment (see Table 4). This prevalence is quite remarkable as it is unusual after an online survey, and especially surprising given that it took one hour to complete. This high proportion is already indicative of a high level of engagement. When content-coding the 163 comments, we see that they were overwhelmingly positive (generally positive 28.8 %, experience perceived as interesting 28.8 %, as fun 12.3 %, as immersive 9.3 %, exciting 9.2 %, as realistic 6.1 %, as thought-provoking 5.5 %). Only 8.6 % gave a generally negative comment. This pattern of open comments is good news as it shows that significant

proportions of the participants connected to the Novaland Experience despite its text-heavy and lengthy nature.[2]

*Table 4: Codes with prevalence greater than 5 % among open comments*

| Code | N | % |
|---|---|---|
| General positive comment | 47 | 28.8 |
| Perceived participation as interesting | 47 | 28.8 |
| Perceived participation as fun | 20 | 12.3 |
| Perceived participation as immersive | 16 | 9.8 |
| Perceived participation as exciting | 15 | 9.2 |
| Perceived Novaland as realistic | 10 | 6.1 |
| Perceived Novaland as thought-provoking | 9 | 5.5 |
| General negative comment | 14 | 8.6 |

Note. Comments may be coded with more than one code. % relates to N = 163 open comments. Some codes are not shown (< 5 %).

### *External Validity: Correlations with Post-Experience Survey Answers*

As we will lay out below, we find some correlations in expected directions between post-experience survey answers and in-experience behaviour. This means that there is a carry-over effect of what participants think in the real world and how they behave in Novaland. For example, participants who are more in favour of strong welfare state activities and the reduction of inequality are also, with many other factors being held constant, more likely to donate a higher proportion of their in-experience assets in Nova and less likely to vote for the right parties of Novaland.

### *Descriptive Evidence on Outcome Variables*

Table 5 lists the means or proportions of 7 outcome variables that are collected in the first four phases of Novaland. There are three columns: column 1 shows the figures from the feasibility study in May 2022, column 2 shows the figures for the people who experienced corruption in Novaland but no unemployment in the November 2022 study. Column 3 reveals the numbers for those who experienced unemployment but no corruption in the November study. This three-group comparison allows us a bit of an overview of patterns.

*Table 5: Means or proportions with Y=1 of dependent variables*

| Dependent variables | Baseline study (May 2022) | Treatment Corruption (1 = yes) | Treatment Unemployment (1 = yes) | Unit |
|---|---|---|---|---|
| Donation | 580.0 | 509.8 | 201.1 | Nova |
| Donation as proportion of all assets in Novas | 14 | 13 | 10 | % |
| Voting for economic right parties (ERP) | 10 | 17 | 23 | % |
| Satisfaction with election outcome | 4.3 | 3.9 | 3.9 | 1 low to 5 high |

---

[2] We also find that losers of the virtual election are in tendency less satisfied with the election outcome, another puzzle piece indicating that participants take the experience sincerely.

| | | | | |
|---|---|---|---|---|
| Satisfaction with government programme | 3.3 | 3.6 | 3.6 | 1 low to 5 high |
| Satisfaction with own situation | 3.5 | 3.3 | 2.7 | 1 low to 5 high |
| N | 31 | 161 | 163 | |

As to donations when being told that a wildfire is destroying people's homes, participants who are in the unemployment condition give substantially less, both absolutely in Novas and relatively as to their assets, when compared against the May sample (where unemployment or corruption were not treated). Participants in the unemployment condition gave only 201.1 Novas and 10 % of their assets on average. As to voting, the corruption-treatment and the unemployment-treatment groups revealed much higher likelihoods to vote for one of the right Novaland parties (17 % and 23 %) compared to just 10 % in the May group. Given that a left majority materialised after the election, it is thus not surprising to have lower levels of satisfaction with the electoral outcome in these treatment groups, too (3.9 versus 4.3). Asked about the government programme that the parties promised, the two groups show very similar levels of mean satisfaction in comparison to the May group. Finally, when asked about their level of satisfaction with the personal situation, the unemployment group stands out again as having the lowest level of satisfaction (2.7) compared with the corruption (3.3) and May group (3.5).

Even though this is just descriptive evidence that compares two experimental treatment groups with one artificial control group from another study, we can detect that the experience of unemployment in Novaland with a loss in income and the reception of unemployment benefits generates stronger effects than the corruption experience.

### *Multiple Regression Estimates on In-Experience and Out-of-Experience Predictor Variables*

Table 6 lists the estimates of 6 multiple regressions on the same outcome variables that we have just explored in the comparison across three groups. The first three rows of estimate show the experimental treatments of three income groups with the lowest as a baseline and the binary indicator of being treated in corruption or unemployment.

The assigned income group captures significant variations in four out of the six outcome variables reported here (and barely insignificant for one). This shows that the assigned income causally drives people's behaviour and thinking across the Novaland experience. High-income treated participants give more money to others, vote more for right parties and are less satisfied with a left government wanting to increase taxes. Real income as reported in the survey, which is not a random sample of the population, does not matter except for a small captured effect for the last dependent variable (Satisfaction with own situation).

The unemployment-treated individuals show significantly different predicted values on the dependent variables compared to corruption-treated individuals. Unemployment causally decreases the absolute amount of donations (mainly via the associated loss in income), the relative proportion of assets that are given in donations and decreases satisfaction with one's personal situation.

*Table 6: Multiple Regressions on 6 outcome variables*

| Independent variables | 1 OLS | 2 Fractionalised GLM | 3 Fractionalised GLM | 4 OLS | 5 OLS | 6 OLS |
|---|---|---|---|---|---|---|
| | Donation | Donation (in relation to assets) | Voting for economic right parties | Satisfaction with election outcome | Satisfaction with government programme | Satisfaction with own situation |
| **In-experience variables** | | | | | | |
| Novaland income (1400 Novas low as ref.) | | | | | | |
| 2000 Novas | 111.80 | 0.02 | 2.16*** | -0.16 | -0.17 | 0.83*** |
| | (78.87) | (0.18) | (0.57) | (0.14) | (0.15) | (0.12) |
| 3500 Novas | 586.80*** | 0.21 | 2.47*** | -0.23 | -0.45** | 1.29*** |
| | (77.55) | (0.19) | (0.57) | (0.13) | (0.14) | (0.12) |
| Treatment (corruption as ref.) | | | | | | |
| Unemployment (1 = yes) | -328.31*** | -0.33* | 0.26 | -0.01 | 0.11 | -0.63*** |
| | (63.35) | (0.15) | (0.36) | (0.11) | (0.12) | (0.10) |
| **Out-of-experience variables** | | | | | | |
| Age (years) | 0.50 | -0.00 | -0.04* | 0.00 | 0.01* | -0.00 |
| | (2.44) | (0.01) | (0.02) | (0.00) | (0.00) | (0.00) |
| Female (1 = yes) | -49.41 | -0.08 | -0.31 | -0.03 | -0.07 | -0.23* |
| | (64.39) | (0.15) | (0.36) | (0.11) | (0.12) | (0.10) |
| Education (low as ref.) | | | | | | |
| high (1 = yes) | 39.70 | -0.05 | -1.28* | 0.39** | -0.18 | -0.28* |
| | (86.44) | (0.19) | (0.53) | (0.15) | (0.16) | (0.13) |
| Real income (thousand Euros) | -34.49 | 0.04 | 0.01 | 0.02 | 0.09 | -0.10* |
| | (29.09) | (0.08) | (0.17) | (0.05) | (0.05) | (0.00) |
| Support for welfare expenses (1 low to 5 high) | 53.82 | 0.17* | -0.51** | 0.23*** | -0.08 | 0.06 |
| | (31.24) | (0.08) | (0.17) | (0.05) | (0.06) | (0.05) |
| Support for inequality reduction (1 low to 5 high) | 48.94 | 0.19* | -0.91*** | 0.28*** | 0.15* | -0.10 |
| | (34.85) | (0.08) | (0.19) | (0.06) | (0.07) | (0.05) |
| Intercept | -45.84 | -3.16*** | 4.13*** | 1.65*** | 3.05*** | 3.45*** |
| | (192.72) | (0.42) | (1.19) | (0.33) | (0.36) | (0.29) |
| $R^2$/ Mc Fadden-$R^2$ | 0.26 | 0.02 | 0.32 | 0.24 | 0.09 | 0.37 |

N=310, p-values of a two-sided test beta=0. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. All dependent variables were measured before the treatment "forest fire". Models 1, 4, 5 and 6 are based on OLS regressions, whereas models 2, 3 and 7 are based on fractionalised General Linear Models (Williams 2019).

The unemployment-treated individuals show significantly different predicted values on the dependent variables compared to corruption-treated individuals. Unemployment causally decreases the absolute amount of donations (mainly via the associated loss in income), the

relative proportion of assets that are given in donations and decreases satisfaction with one's personal situation.

We have thus replicated the causal effects of two known experiences in the real world in Novaland: namely the income that can be spent and unemployment. This is important evidence as to the external validity of Novaland as we can produce a context in which known factors from the outside world can be reproduced in the Novaland Experience. This is an important first step on the feasibility assessment of Novaland before we will turn to other forms of causal factors about which less is known.

Let us now turn lastly to the out-of-experience variables. Please recall that this is a convenience sample of people recruited through a social media campaign. The sample is younger and more educated compared with the resident adult population in Germany. Still, we have substantial variation on many out-of-experience background variables whose influence we can explore in the regression analysis, too. In a nutshell, we find very little significant coefficients for age, gender, education, and real income. But the political attitudes measuring support for welfare state spending and for inequality reduction by the state, two of the most common indicators of welfare state attitudes, do capture significant variation. People who report to being more pro-redistribution outside of Novaland, give more donations in Novaland, are less likely to vote for right parties, are more satisfied with a left majority and the ensuing governmental programme. These pieces of evidence give us some first indication that the attitudes that are linked to deep, underlying ideology are so hard-wired that they come to play even in the fully virtual experience.


## 7. Conclusions

Our main contributions are: (1) The newly programmed Novaland Experience worked technically well in that over 300 participants could simultaneously interact online at the same time. (2) Participants behaved in an internally valid manner even though there are some signs of inattentiveness. (3) The political ideology of volunteers was reflected in how they behaved in Novaland, whereas socio-demographic variables showed no systematic effects. (4) The assigned income and unemployment treatments worked in that they causally changed the way in which the participants behaved towards others and towards politics and policies in Novaland with unclear evidence as to the corruption experience. Overall, this data collection exercise is encouraging as it allows us to explore the potential for relevant social science research in virtual worlds in the 2020s.

Our next steps are: (1) To preregister hypotheses from our current model and then to replicate the same version of Novaland while fully controlling multiple randomised assignments. (2) To extend the data collection to volunteers from the German-speaking parts of Switzerland, Germany and Austria. (3) To empirically test whether linking the financial pay-out to Novaland behaviour changes the results. (4) To empirically test whether the physical presence of participants changes the results and to translate the findings into observable implications in an international survey and field experiments.

There are a number of ways in which the Novaland platform can be used in other applications. First, it can be employed in **teaching**. Students can be taught to think about institutional design rules, implement these institutions as text and then apply them in their version of Novaland to see how participants, such as their class peers, react to their design features. Second, the platform can be used as a **virtual policy lab**. Before policy-makers are implementing a certain public policy, they can see how participants react to that policy in the artificial world of Novaland. The contents of the experience are fully exchangeable. This is all the more of relevance in a political system like Germany in which piloting of public policies is constrained by strict legal rules that prohibit the unequal treatment of citizens in Germany that

a policy pilot will always entail. Third, **empirical democratic theorists** can implement utopian or dystopian versions of democratic systems in the Novaland experience. For instance, decision by delegation to mini-publics or other features of modern direct democracy can be tried out on participants.

The virtual world approach of Novaland has its limitations. First of all, it takes us as social scientists to the edge of our training. Even though we can learn how to code on a basic level, we will always be dependent on programmers to do the heavy lifting and can only quality-check their work up to a certain degree. Secondly, there are a number of powerful studies in the real world that measure causal effects for political solidarities and related concepts. For instance, natural disasters create a short-term increase in prosocial behaviour and social trust (Li, Li, and Ruan 2021). The as-if-random placement of refugees in Sweden in the 1980s and 1990s reduced support for public redistribution (Dahlberg, Edmark, and Lundqvist 2012). Field experiments are another way of manipulating the treatment and have some control over compliance in the real world. Future iterations of Novaland must be able to derive expectations for these real-world assessments as the ultimate tests of its helpfulness.

## 8. References

Atlas, Stephen A. 2008. "Inductive Metanomics: Economic Experiments in Virtual Worlds." *Journal Virtual Worlds Research* 1 (1): 1–15.

Awad, Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. 2018. "The Moral Machine experiment." *Nature* 563 (7729): 59–64. https://doi.org/10.1038/s41586-018-0637-6.

Bainbridge, William S. 2007. "The Scientific Research Potential of Virtual Worlds." *Science* 317 (5837): 472–76.

Bauer, Michal, Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts. 2016. "Can War Foster Cooperation?" *Journal of Economic Perspectives* 30 (3): 249–74. https://www.aeaweb.org/articles?id=10.1257/jep.30.3.249.

Busemeyer, Marius R., and Achim Goerres. 2020. "Policy feedback in the local context: analysing fairness perceptions of public childcare fees in a German town." *Journal of Public Policy* 40 (3): 513–33. https://www.cambridge.org/core/article/policy-feedback-in-the-local-context-analysing-fairness-perceptions-of-public-childcare-fees-in-a-german-town/F6CF70723BA0BBAC209DCE41ABA002F6.

Cavaillé, Charlotte, and Kris-Stella Trump. 2015. "The Two Facets of Social Policy Preferences." *The Journal of Politics* 77 (1): 146–60.

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree—An open-source platform for laboratory, online, and field experiments." *Journal of Behavioral and Experimental Finance* 9: 88–97. https://www.sciencedirect.com/science/article/pii/S2214635016000101.

Chesney, T., S. H. Chuah, and R. Hoffmann. 2009. "Virtual world experimentation: An exploratory study." *Journal of Economic Behavior & Organization* 72 (1): 618–35.

Christensen, Tamlin C., Lisa Feldman Barrett, Eliza Bliss-Moreau, Kirsten Lebo, and Cynthia Kaschub. 2003. "A practical guide to experience-sampling procedures." *Journal of Happiness Studies* 4 (1): 53–78.

Dahlberg, Matz, Karin Edmark, and Heléne Lundqvist. 2012. "Ethnic Diversity and Preferences for Redistribution." *Journal of Political Economy* 120 (1): 41–76.

Del Ponte, Alessandro, and Peter DeScioli. 2019. "Welcome to the Saving Game." http://pdescioli.com/orangegame/demos.html.

Duffy, J. 2011. "Trust in Second Life." *SOUTHERN ECONOMIC JOURNAL* 78 (1): 53–62.

Eifler, Stefanie, and Knut Petzold. 2019. "Validity Aspects of Vignette Experiments: Expected "What-If" Differences Between Reports of Behavioral Intentions and Actual Behavior." In *Experimental Methods in Survey Research*, eds. P. Lavrakas, M. Traugott, C. Kennedy, A. Holbrook, E. de Leeuw and B. West, 393–416.

Faralla, Valeria, Alessandro Innocenti, Stefano Baraldi, Luca Lusuardi, Maurizio Masini, Vincenzo Santalucia, Diletta Scaruffi, and Matteo Sirizzotti. 2021. "The Effect of Future Self-Continuity on Delaying Immediate Gratification. A Virtual Reality Experiment." In *Proceedings of the 32nd European Conference on Cognitive Ergonomics. ECCE '21*. New York, NY, USA: Association for Computing Machinery.

Gabriel, Allison S., Nathan P. Podsakoff, Daniel J. Beal, Brent A. Scott, Sabine Sonnentag, John P. Trougakos, and Marcus M. Butts. 2019. "Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement." *Organizational Research Methods* 22 (4): 969–1006.

Goerres, Achim. 2021. *The Research Agenda of POLITSOLID: The Ties that Bind - Experimental Analyses of Political Solidarities in Modern European Democracies*. Social Science Research Network. http://dx.doi.org/10.2139/ssrn.3792243.

Goerres, Achim, and Jan K. Höhne. forthcoming [2023]. "Evaluating the Response Effort and Data Quality of Established Political Solidarity Measures: A Pre-Registered Experimental Test in an Online Survey of the German Adult Resident Population in 2021." *Quality & quantity*.

Groenland, E.A.G., A.A.A. Kuylen, and J. G. Bloem. 1996. "The dynamic process tracing approach: Towards the development of a new methodology for the analysis of complex consumer behaviours." *Journal of Economic Psychology* 17 (6): 809–25. https://www.sciencedirect.com/science/article/pii/S0167487096000402.

Höhne, Jan K. 2020. "Experience sampling." In *SAGE research methods foundations*, eds. Paul Atkinson, Sara Delamont, Alexandru Cernat, Joseph W. Sakshaug and Richard A. Williams. London: Sage.

Innocenti, Alessandro. 2017. "Virtual reality experiments in economics." *Journal of Behavioral and Experimental Economics* 69: 71–77. https://www.sciencedirect.com/science/article/pii/S221480431730068X.

Jetten, Jolanda, Frank Mols, Nikita Healy, and Russell Spears. 2017. ""Fear of Falling": Economic Instability Enhances Collective Angst among Societies' Wealthy Class." *Journal of Social Issues* 73 (1): 61–79.

Kozlov, Michail D., and Mark K. Johansen. 2010. "Real Behavior in Virtual Environments: Psychology Experiments in a Simple Virtual-Reality Paradigm Using Video Games." *Cyberpsychology, Behavior and Social Networking* 13 (6).

Levy, Jacob T. 2017. "Against Fraternity: Democracy without Solidarity." In *The strains of commitment: the political sources of solidarity in diverse societies*. still image, eds. Keith G. Banting and Will Kymlicka. Oxford/New York, NY: Oxford University Press, 107–24.

Li, Yao, Haoyang Li, and Jianqing Ruan. 2021. "Do Long-Term Natural Disasters Influence Social Trust? Empirical Evidence from China." *International Journal of Environmental Research and Public Health* 18 (14).

Mummolo, Jonathan, and Erik Peterson. 2019. "Demand Effects in Survey Experiments: An Empirical Assessment." *American Political Science Review* 113 (2): 517–29. https://www.cambridge.org/core/article/demand-effects-in-survey-experiments-an-empirical-assessment/043386DC63A69098E859414EF9932EBC.

Naumann, Elias, Christopher Buss, and Johannes Bähr. 2016. "How Unemployment Experience Affects Support for the Welfare State: A Real Panel Approach." *European Sociological Review* 32 (1): 81–92.

Navarrete, C. D., Melissa M. McDonald, Michael L. Mott, and Benjamin Asher. 2012. "Virtual morality: Emotion and action in a simulated three-dimensional "trolley problem"." *Emotion* 12: 364–70.

Orne, Martin T. 2017 [1970]. "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications." In *Sociological Methods*, ed. Norman K. Denzin. New York [E-book]: Routledge, Chap. 18.

Quidt, Jonathan de, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." *American Economic Review* 108 (11): 3266–3302. https://www.aeaweb.org/articles?id=10.1257/aer.20171330.

Rothstein, Bo, and Daniel Eek. 2009. "Political Corruption and Social Trust." [English]. *Rationality and Society* 21 (1): 81–112.

Shellman, Stephen M. 2001. "Active learning in comparative politics: A mock German election and coalition-formation simulation." *PS: Political Science and Politics* 34 (4): 827–34.

Simonovits, Gábor, Gábor Kézdi, and Péter Kardos. 2018. "Seeing the World Through the Other's Eye: An Online Intervention Reducing Ethnic Prejudice." *American Political Science Review* 112 (1): 186–93.

Slater, Mel, and Maria V. Sanchez-Vives. 2016. "Enhancing Our Lives with Immersive Virtual Reality." *Frontiers in Robotics and AI* 3. https://www.frontiersin.org/articles/10.3389/frobt.2016.00074.

Stover, William J. 2005. "Teaching and Learning Empathy: An Interactive, Online Diplomatic Simulation of Middle East Conflict: An Interactive, Online Diplomatic Simulation of Middle East Conflict." *Journal of Political Science Education* 1 (2): 207–19.

Tanjitpiyanond, Porntida, Jolanda Jetten, and Kim Peters. 2022. "How economic inequality shapes social class stereotyping." *Journal of Experimental Social Psychology* 98: 104248. https://www.sciencedirect.com/science/article/pii/S0022103121001517.

Titmuss, Richard M. 2019 [1958]. *Essays on "The Welfare State"*. Bristol: Policy Press.

van Berkel, Niels, Denzil Ferreira, and Vassilis Kostakos. 2018. "The experience sampling method on mobile devices." *ACM Computing Surveys* 50 (6): Article 93.

Williams, Richard. 2019. "Analyzing Proportions: Fractional Response and Zero One Inflated Beta Models." https://www3.nd.edu/~rwilliam/stats3/fractionalresponsemodels.pdf.