# Extended Supplementary Materials for "The Transmission Game: Testing behavioral interventions in a pandemic-like simulation"

Jan K. Woike
Max Planck Institute for Human Development, Berlin, Germany; University of Plymouth, United Kingdom

Sebastian Hafenbrädl
IESE Business School, Barcelona, Spain

Patricia Kanngiesser
Freie Universität Berlin, Germany; University of Plymouth, United Kingdom

Ralph Hertwig
Max Planck Institute for Human Development, Berlin, Germany

Abstract

This is the extended version of the Supplementary Materials for "The Transmission Game: Testing behavioral interventions in a pandemic-like simulation".

## Contents

**E1 Organization of the Supporting Material** **9**
E1.1 Contents of different documents . . . . . 9
E1.2 File location, reproducibility and code . . . . . 10
E1.3 Selection for this document . . . . . 10
E1.4 Links to preregistrations . . . . . 11
E1.5 Intervention names in main text and Supplementary Materials . . . . . 11

**I The Transmission Game Paradigm** **12**

**E2 Background and Study Goals** **12**
E2.1 Approaches to investigating an ongoing pandemic . . . . . 12

Jan K. Woike https://orcid.org/0000-0002-6816-121X, Sebastian Hafenbrädl https://orcid.org/0000-0002-5148-766X, Patricia Kanngiesser https://orcid.org/0000-0003-1068-3725, Ralph Hertwig https://orcid.org/0000-0002-9908-9556
Correspondence e-mail: woike@mpib-berlin.mpg.de

E2.2 Contributions of the behavioral sciences . . . 13
E2.2.1 The need for the behavioral sciences . . . 13
E2.2.2 Surveys, Experiments and Games . . . 13
E2.3 Brief summary of game paradigm and chosen interventions . . . 14
E2.3.1 Rationale for the construction of the transmission game paradigm 14
E2.3.2 Choice of interventions for Study 2 . . . 14
E2.4 Study goals . . . 14
E2.4.1 Goals of Study 1 . . . 14
E2.4.2 Goals of Study 2 . . . 15

**E3 Transmission Game Framework 15**
E3.1 Rules of the standard transmission game . . . 15
E3.2 Illustrations of game rules and implications . . . 16
E3.3 Development of the framework . . . 19
E3.3.1 Neutral framing . . . 20
E3.3.2 Relevant features of the health crisis . . . 21
E3.3.3 Mapping these features . . . 22
E3.3.4 Asynchronous implementation in an online setting . . . 26
E3.3.5 No deception . . . 29
E3.3.6 Balance between complexity and comprehension . . . 29

**II Formal analysis of the Transmission Game 32**

**E4 Summary of Theoretical Simulation Results 32**
E4.1 Degree of cooperation and color switching rates . . . 32
E4.2 The social dilemma at the heart of the transmission game . . . 35
E4.3 Optimal individual strategy . . . 37
E4.4 Flattening the curve: A collusive optimum . . . 37

**E5 Formal analysis of the transmission game 40**
E5.1 Simulations and formal analysis . . . 40
E5.2 Modelling the transmission game: Uniform behavior . . . 40
E5.2.1 Population states and game results . . . 40
E5.2.2 Strategies . . . 41
E5.2.3 Roundwise transmissions . . . 41
E5.2.4 Putting the parts together . . . 46
E5.3 Modelling the transmission game: Non-uniform behavior with unilateral deviation . . . 49
E5.3.1 Population states and game results . . . 49
E5.3.2 Strategies and transmission risk . . . 50
E5.3.3 Roundwise transmissions . . . 51
E5.3.4 Putting the parts together . . . 62
E5.4 Extensions of the analysis tool . . . 66
E5.4.1 Partial games . . . 66
E5.4.2 Modelling player groups: Some progress and obstacles . . . 67

**E6 Analytical results for the transmission game** **71**
E6.1 Confirming the results of previous simulations . . . 71
E6.1.1 Revisiting the flattening of the curve . . . 71
E6.1.2 Roundwise results . . . 73
E6.2 Collusive equilibrium . . . 76
E6.2.1 Analysis results . . . 76
E6.2.2 Implications of results . . . 77
E6.2.3 Testing the implications . . . 77
E6.3 Search for the individual equilibrium . . . 79
E6.3.1 Individual HG-patterns . . . 80
E6.3.2 Varying group behavior . . . 82
E6.3.3 Best possible response analysis . . . 83
E6.3.4 Considering a mixed-strategy equilibrium . . . 87
E6.4 Some caveats in interpreting the results . . . 89
E6.4.1 Introducing a utility function . . . 90
E6.4.2 Repeating the analysis for expected utility . . . 90
E6.4.3 Other utility-considerations . . . 90

**III Supplementary Materials for Study 1** **93**

**E7 Study 1: Selected Materials** **93**
E7.1 Transmission game instructions . . . 93
E7.1.1 Page 1/5 . . . 93
E7.1.2 Page 2/5 . . . 94
E7.1.3 Page 3/5 . . . 97
E7.1.4 Page 4/5 . . . 99
E7.1.5 Page 5/5 . . . 101
E7.2 Transmission Game . . . 101
E7.2.1 Introduction . . . 101
E7.2.2 Rounds . . . 105
E7.2.3 Final results . . . 106
E7.3 Measures and additional variables used in Study 1 . . . 107
E7.3.1 Demographics . . . 107
E7.3.2 Cognitive measures . . . 107
E7.3.3 Political preferences and attitudes . . . 107
E7.3.4 Personality scales . . . 108
E7.3.5 Personality single-item measures . . . 108
E7.3.6 Economic games and social preferences . . . 109
E7.3.7 Brief health measures . . . 109
E7.3.8 COVID-19 related measures . . . 109
E7.3.9 Postquestionnaire . . . 109

**E8 Study 1: Selected Results** **110**
E8.1 Behavior in the transmission game . . . 110

E8.2 Postquestionnaire results . . . 111
E8.2.1 Game motivation . . . 111
E8.2.2 Strategy descriptions . . . 112
E8.2.3 Further qualitative data . . . 113
E8.2.4 Estimates of other participants' behavior . . . 113
E8.3 Simulation results based on Study 1 data . . . 115
E8.4 Predictors of game results . . . 116
E8.5 Political measures and COVID-19 . . . 118

**IV Supplementary Materials for Study 2 120**

**E9 Study 2: Interventions 120**
E9.1 Vicarious Learning . . . 120
E9.2 Norm-based interventions . . . 120
E9.2.1 Descriptive norms . . . 121
E9.2.2 Injunctive norms . . . 121
E9.3 Boosting interventions . . . 122
E9.3.1 Chain reaction . . . 122
E9.3.2 Simulator . . . 123

**E10 Study 2: Selected Materials 123**
E10.1 Transmission Game: Condition-specific instructions . . . 124
E10.1.1 Instructions: Injunctive norms condition . . . 124
E10.1.2 Screenshots: Injunctive-norms . . . 124
E10.1.3 Instructions: Chain reaction condition . . . 126
E10.1.4 Screenshots: Chain reaction . . . 126
E10.1.5 Instructions: Simulator condition . . . 136
E10.1.6 Screenshots: Simulator . . . 137
E10.2 Condition-specific parts in the game . . . 142
E10.2.1 Introduction: Addition in the vicarious learning condition . . . 142
E10.2.2 Introduction: Addition in the descriptive norms condition . . . 144
E10.2.3 Rounds (vicarious learning condition only) . . . 146
E10.2.4 Rounds (descriptive norms condition only) . . . 150
E10.2.5 Rounds (injunctive norms condition only) . . . 153
E10.3 Measures and additional variables used in Study 2 . . . 155
E10.3.1 Demographics . . . 155
E10.3.2 Cognitive measure . . . 155
E10.3.3 Political preferences and attitudes . . . 155
E10.3.4 Personality scales . . . 155
E10.3.5 Personality single-item measures . . . 155
E10.3.6 Social preferences . . . 156
E10.3.7 COVID-19 related measures . . . 156
E10.3.8 Postquestionnaires . . . 156

**E11 Study 2: Selected Results 156**

E11.1 Behavior in the transmission game . . . 156
E11.2 Risky choices by round . . . 156
E11.3 Distribution of risky choices across participants . . . 158
E11.4 Behavioral categories and predictor variables . . . 160
E11.4.1 Behavioral categories across conditions . . . 160
E11.4.2 Behavioral categories and social value orientation . . . 163
E11.4.3 Behavioral categories and risk preference . . . 165
E11.4.4 Behavioral categories and cognitive reflection . . . 166
E11.5 Postquestionnaire-results . . . 168
E11.5.1 Expectations . . . 168
E11.5.2 Expectation for self versus others . . . 170
E11.5.3 Motivations (general postquestionnaire) . . . 171
E11.5.4 Motivations (specific postquestionnaires) . . . 176
E11.6 Predictor variables and game results . . . 178
E11.7 Game results, political attitudes and COVID-19 related scales . . . 180
E11.8 Qualitative Data . . . 182
E11.8.1 Transmission game associations . . . 182
E11.8.2 Messages to other players . . . 183

**E12 Study 2: Selected Simulation Results 187**

**V Discussion and Extensions 192**

**E13 Discussion of Selected Results 192**
E13.1 Results for predictor variables . . . 192
E13.1.1 Social value orientation . . . 192
E13.1.2 Risk-taking . . . 192
E13.1.3 Cognitive reflection . . . 193
E13.1.4 Personality . . . 193
E13.2 Postquestionnaire results . . . 194
E13.2.1 Linear transmission curves . . . 194
E13.2.2 Optimism . . . 194
E13.2.3 Survey responses vs. behavior . . . 195
E13.3 Attitudes towards COVID-19 . . . 195
E13.3.1 Political ideology . . . 195
E13.3.2 Conspiratorial thinking . . . 196
E13.4 Intervention results . . . 197
E13.4.1 Vicarious learning . . . 197
E13.4.2 Descriptive norms intervention . . . 198
E13.4.3 Injunctive norms intervention . . . 199
E13.4.4 Boosting interventions . . . 200

**E14 Variants, extensions, and research ideas 201**
E14.1 Paper and pencil version of the Transmission Game . . . 201
E14.1.1 Game sheets . . . 201

E14.1.2 Determining results . . . 201
E14.1.3 Possible applications of the paper-and-pencil variant . . . 204
E14.2 Varying parameters in the simulation . . . 205
E14.2.1 Reducing the number of players . . . 207
E14.2.2 Extending the number of rounds to 40 . . . 208
E14.2.3 Lowering the transmission probabilities further . . . 208
E14.2.4 Further variations . . . 212
E14.2.5 Notes on the choice of parameters . . . 212
E14.3 Adding elements to the simulation: Infectious periods and immunity . . . 214
E14.3.1 Simulation code . . . 214
E14.3.2 Simulation results . . . 218
E14.3.3 Discussion . . . 224
E14.4 A look forward . . . 224

# List of Figures

E1 Representation of players: two possible states. . . . 16
E2 Population at the start of the game and before the first round. . . . 17
E3 Pairing of players and color changes. . . . 18
E4 Probability of switching from blue to purple conditional on player actions. 18
E5 Progression of the game to the end. . . . 19
E6 Cumulative number of positive cases in the United States since March 15, 2020. . . . 22
E7 Distribution of observed proportions of purple participants by round (50% ALLG-players and 50% ALLH-players). . . . 33
E8 Average proportion of purple participants after each round for each simulated number of H-players. . . . 34
E9 Average rate of staying blue as a function of the number of AllH-players. 35
E10 Average expected payoff as a function of the number of AllH-players. . . . 36
E11 Average proportion of purple participants after each round for each possible collective switching point from H-actions to G-actions. . . . 38
E12 Scatterplot of probability of staying blue and the average expected outcome for all possible switching points. . . . 39
E13 Average proportion of purple participants after each round for each possible collective switching point from H-actions to G-actions . . . 74
E14 Derived roundwise distributions of computed population states . . . 75
E15 Transformation of a strategy into an HG-strategy by repeatedly swapping G and H. . . . 78
E16 Difference in survival rates for players starting with H or ending with it. . 79
E17 Difference in survival rates for focus individuals starting with H or ending with it. . . . 82
E18 Difference in survival rates for focus individuals with GH- and HG-strategies. 83
E19 Simulation results for the analysis of HG-strategy combinations. . . . 86
E20 Analytical results for HG-strategy combinations. . . . 87

E21 Simulation results for the analysis of different action proportions in round 12. . . . 89
E22 Utility function example. . . . 91
E23 Analytical results for expected utility when combining HG-strategies. . . . 92
E24 Game Instructions Page 1 . . . . 94
E25 Game Instructions Page 2 . . . . 96
E26 Game Instructions Page 3 . . . . 98
E27 Game Instructions Page 4 . . . . 100
E28 Game Instructions Page 5 . . . . 101
E29 Game Introduction . . . . 103
E30 Game matrix (presented below game introduction) . . . . 104
E31 Game round 1 (example) . . . . 105
E32 Game round 2 (example) . . . . 106
E33 End of game (example) . . . . 107
E34 Histogram of scores after round 25. . . . 110
E35 Proportion and 95% CI of proportion of H-choices for each round of the game. . . . 111
E36 Histograms of postquestionnaire-responses and game score . . . . 112
E37 Boxplots of estimates of the number of purple players. . . . 114
E38 Boxplots of estimates of the number of H-actions. . . . 115
E39 Simulation results for games based on participants' choices in Study 1. . . . 116
E40 Predictors of game behavior: Histograms, correlations and scatterplots. . . . 117
E41 Political attitudes and COVID-19 scales: Histograms, correlations and scatterplots. . . . 119
E42 Instructions (injunctive norms) . . . . 125
E43 Instructions (chain reaction, page 1/9) . . . . 127
E44 Instructions (chain reaction, page 2/9) . . . . 128
E45 Instructions (chain reaction, page 3/9) . . . . 129
E46 Instructions (chain reaction, page 4/9) . . . . 130
E47 Instructions (chain reaction, page 5/9) . . . . 131
E48 Instructions (chain reaction, page 6/9) . . . . 132
E49 Instructions (chain reaction, page 7/9) . . . . 133
E50 Instructions (chain reaction, page 8/9) . . . . 134
E51 Instructions (chain reaction, page 9/9) . . . . 135
E52 Instructions (simulator, page 1/2) . . . . 138
E53 Instructions (simulator, page 2/2, upper part) . . . . 139
E54 Instructions (simulator, page 2/2, lower part) . . . . 140
E55 Instructions (simulator, example of 99% H results) . . . . 141
E56 Instructions (simulator, example of 25% H results) . . . . 141
E57 Instructions (vicarious learning condition). . . . 143
E58 Instructions (descriptive-norms condition). . . . 145
E59 Game round 1 (example in vicarious condition) . . . . 147
E60 Graph in game round 5 (example in vicarious condition with tooltip shown) 148
E61 Graph in game round 25 (example in vicarious condition) . . . . 149
E62 Interface in game round 1 (example in descriptive-norms condition) . . . 151

E63 Interface in game round 6 (example in descriptive-norms condition) . . . 152
E64 Game round 5 (example in injunctive-norms condition) . . . . . 154
E65 Transmission Game round-wise results in Study 2 . . . . . 157
E66 Distribution of percentage of risky choices for all experimental conditions. 159
E67 Game results across conditions in Study 2. . . . . 160
E68 Mosaic plot of the proportion of behavioral patterns observed in the six experimental conditions. . . . . 162
E69 Behavioral pattern and SVO. . . . . 164
E70 Behavioral pattern and risk preference. . . . . 165
E71 Behavioral pattern and CRT score: Mosaic plot. . . . . 167
E72 Postquestionnaire expectations. . . . . 169
E73 Relative optimism: expected color switching for self and others. . . . . 171
E74 Game results and postquestionnaire motivations (part 1). . . . . 173
E75 Game results and postquestionnaire motivations (part 2). . . . . 175
E76 Game results and condition-specific postquestionnaire responses. . . . . 177
E77 Predictor variables and game results. . . . . 179
E78 Game results, political attitudes and COVID-19 related scales. . . . . 181
E79 Categorization of messages by game behavior. . . . . 185
E80 Categorization of messages by game condition. . . . . 186
E81 Categorization of messages split by game condition. . . . . 187
E82 Results for simulated games based on the observed decisions in Study 2. . 189
E83 Comparison of original simulation results and simulation results after resampling. . . . . 190
E84 Average outcomes across simulated games by strategy and condition. . . . 191
E85 Game slips for the paper-and-pencil version. . . . . 202
E86 Filled in game slips (paper-and-pencil version). . . . . 203
E87 Twenty-sided die for the paper-and-pencil version of the game. . . . . 203
E88 Game round tracker and die result table. . . . . 204
E89 Final state of the game slips. . . . . 205
E90 Payoff calculation sheet. . . . . 206
E91 Simulation results for a parameter variation with 24 players (variant 1). . 208
E92 Simulation results for two parameter variations with 40 rounds (variants 2 and 3). . . . . 209
E93 Simulation results for two parameter variations with two purple participants at the start of the game (variants 4 and 5). . . . . 211
E94 Simulation results for two parameter variations (variants 6 and 7). . . . . 213
E95 Simulation results for a parameter variation (variant 8). . . . . 214
E96 Results for simulations with infectious periods and 0% H-players . . . . . 219
E97 Results for simulations with infectious periods and 25% H-players . . . . 220
E98 Results for simulations with infectious periods and 50% H-players . . . . 221
E99 Results for simulations with infectious periods and 75% H-players . . . . 222
E100 Results for simulations with infectious periods and 100% H-players . . . . 223

# List of Tables

E1 Mapping of names and abbreviations of interventions in the manuscript to names and abbreviations in the Supplementary Materials. . . . . 11
E2 List of parameter values . . . . . 30
E3 Progression of population states across three rounds in the toy example when all players choose the strategy $S = (1, 1, 0)$. . . . . 49
E4 Progression of population states across three rounds in the toy example when all players choose the strategy $S = (1, 1, 0)$. . . . . 64
E5 Progression of population states across three rounds in the toy example when group players choose the strategy $S_g = (0, 0, 0)$ and the focus individual chooses $S_f = (1, 1, 1)$. . . . . 65
E6 Categories for associations with the transmission game in open-format asnwers. . . . . 182

## E1 Organization of the Supporting Material

To achieve the dual goals of comprehensiveness and accessibility of important information, we have organized the supporting material on two levels. This document is on the top level and gives an overview over all results. On the second level, there are seven documents containing detailed descriptions of results and materials including the generating computer code. Additionally, we have compiled materials relating to the development and administration of the transmission game in this document, together with a selection of simulation results and empirical results to highlight our main findings.

### E1.1 Contents of different documents

The full supporting material on the Dataverse consists of eight separate documents with this documents also serving as a summary. The other seven documents are the following:

1. Study Materials Study 1 (**SM1**)
2. Study Materials Study 2 (including JavaScript code; **SM2**)
3. Variables and Measures Study 1 (including R code; **VM1**)
4. Variables and Measures Study 2 (including R code; **VM2**)
5. Supporting Analyses Study 1 (including R code; **SA1**)
6. Supporting Analyses Study 2 (including R code; **SA2**)
7. Simulation Code and Results (including R and Matlab code; **SCR**)

There are two documents each (one for each Study) for study materials, scales, and analyses. A seventh document contains the code and results for all simulated games. Study materials include all item texts and screenshots of many parts of the study as they were shown to participants. The documents focusing on variables and measures present

inter-item correlations and Cronbach's alphas for scales used in the study and document sources for variables. These documents were compiled using knitR within Overleaf, and contain the full R code used to create results using the raw data files. The documents with supporting analyses were created with knitR as well and likewise include all written code used to generate results and figures. Finally, the simulation documents provide the Matlab code used for simulations, and contain the R code for analysing the simulations' output files. This document was compiled directly from RStudio.

### E1.2 File location, reproducibility and code

These documents are hosted on the Harvard Dataverse in two Dataverses together with all data referenced and used within these documents. For Study 1 the data set and documents (356 pages) are published as:

Woike, J. K., Hafenbrädl, S., Kanngiesser, P., & Hertwig, R. (2021). Replication Data for: Study 2 (The Transmission Game: Testing behavioral interventions in a pandemic-like simulation), *Harvard Dataverse*, V1, https://doi.org/10.7910/DVN/RHR2A5.

1. **SM1**: 113 pages
2. **VM1**: 181 pages
3. **SA1**: 62 pages

For Study 2 the data set and documents (754 pages) are published as:

Woike, J. K., Hafenbrädl, S., Kanngiesser, P., & Hertwig, R. (2021). Replication Data for: Study 2 (The Transmission Game: Testing behavioral interventions in a pandemic-like simulation), *Harvard Dataverse*, V2, https://doi.org/10.7910/DVN/ILDEM0.

1. **SM2**: 128 pages
2. **VM2**: 124 pages
3. **SA2**: 96 pages
4. **SCR**: 162 pages
5. **E**: Extended version of the Supplemnentaty Materials for Study 1 and Study 2, 244 pages (this document)

### E1.3 Selection for this document

The information in this document is mostly a selection of results taken from the seven other documents. We added a short section detailing the game rules and discuss some possible game variants in the final section. We present the material related to the game framework (Study 1) and the five interventions (added in Study 2), omitting comprehension checks and postquestionnaires (see **SM1** and **SM2**). We added a brief description of selected explanatory variables.

We present results supporting the selection of scales from Study 1 that were subsequently implemented in Study 2 (for more details about all measures see **VM1** and **VM2**).

We briefly discuss the relationship between these predictors, game results, political preferences and reactions to COVID-19 (for more results see **SA1** and **SA2**).

As the game rules contain stochastic components that will potentially change outcomes in repeated play-throughs of a game with given choices, we use the simulation method to illustrate likely outcomes and their distribution across 1,000,000 simulations per condition, both for the actual data and for theoretical data. The latter helps to illustrate maximizing and dominated strategies and the social dilemma at the core of the game. The theoretical results also provide context for evaluating the empirical results (for more simulations and results, see **SCR**).

We connect our approach to many recent studies focused on the current pandemic. Due to the relatively short amount of time since the beginning of the pandemic, many of these studies have not been peer-reviewed at this stage. We do not understand our referencing as an unconditional endorsement of these studies, but rather as an attempt of helping to connect independent strands of research and documenting parallel findings and inter-relationships.

### E1.4 Links to preregistrations

We preregistered the materials for both studies on the Open Science Foundation platform. Please find links to the two preregistrations below:

- Study 1: https://osf.io/kc2an/,
- Study 2: https://osf.io/uacfy/.

### E1.5 Intervention names in main text and Supplementary Materials

After completing the Supplementary Materials, we decided to change the names of our intervention conditions to render them more intuitive for a multidisciplinary audience. We did not change the names in the Supplementary Materials, though, as they appear in variable names and throughout the pre-registration, in figures and results files. Table E1 presents the mapping from names and abbreviations used in the manuscript to names and abbreviations used throughout the Supplementary Materials (including this document) and code files.

**Table E1**

*Mapping of names and abbreviations of interventions in the manuscript to names and abbreviations in the Supplementary Materials.*

| Manuscript | | | Supplement | |
|---|---|---|---|---|
| Intervention | Short name | | Intervention | Short name |
| Learning from others' outcomes | LFO | $\longrightarrow$ | Vicarious learning | VIC |
| Observing behavior | OBS | $\longrightarrow$ | Descriptive norms | DES |
| Messaging | MES | $\longrightarrow$ | Injunctive norms | INJ |
| Simulator | SIM | $\longrightarrow$ | Simulator | SIM |
| Chain of Infection | COI | $\longrightarrow$ | Chain | CHN |

# Part I
# The Transmission Game Paradigm

## E2 Background and Study Goals

At the time of writing these lines, the world experiences a rare pandemic impacting and restricting many areas of life and society. The SARS-CoV-2 virus originated in Wuhan, China (Liang et al., 2020), but has since spread to the entire globe. The virus has transformed formerly commonplace social behaviors into threats to health and livelihoods. Based on the data on the "Worldometers" page[1], more than 29.7 million coronavirus cases have been confirmed, with over 939,000 resulting in deaths. While some countries have experienced the crisis in waves, world-wide there are currently over 250,000 cases a day. The high degree of infectiousness and the broad distribution render a fast return to pre-crisis habits unlikely in the short- and mid-term. Rapidly increasing case numbers put a strain on hospitals and medical systems with far-reaching consequences. The pandemic has been considered a "humanitarian crisis" (Nacoti et al., 2020).

### E2.1 Approaches to investigating an ongoing pandemic

The novel coronavirus has become a singular target of research. This project coincides with a virtual deluge of proposals, working papers, and peer-reviewed studies in many disciplines, investigating various aspects of the pandemic and its impact on society and individuals. Much of the immediate response centered on projections of the course of the pandemic into the future (Metcalf et al., 2020; Walker et al., 2020). When data from multiple countries started to accumulate, modellers explored retrospectively the effectiveness of counter-measures, inspired by earlier and (still) hypothetical modelling (e.g., N. M. Ferguson et al., 2005; N. M. Ferguson et al., 2006; Litvinova et al., 2019). Epidemiological, mechanistic models (Holmdahl & Buckee, 2020) simulate the dynamics of transitions between susceptible, infectious and recovered population members (SIR models; Dehning et al., 2020; Leung et al., 2020), with extensions addressing exposure (SEIR models; Block et al., 2020; Prem et al., 2020), quarantine (Maier & Brockmann, 2020), or different degrees of symptom severity (Giordano et al., 2020). Alternatively, econometric models capture the effects of interventions using regression or general forecasting approaches (Hsiang et al., 2020) using country-level or regional data across points in time. Results of both types of models may jointly support the public health response (Shea et al., 2020).

In the absence of an effective vaccine against the novel coronavirus, and given the lack of effective pharmaceutical treatment (López & Rodó, 2020; Tian et al., 2020), non-pharmaceutical interventions are the primary counter-measures. Thus, modellers have focused their efforts on testing the effects of travel bans and restrictions (Chinazzi et al., 2020; Gatto et al., 2020; Kraemer et al., 2020), lockdowns (Davies et al., 2020; Dehning et al., 2020; Flaxman et al., 2020; Giordano et al., 2020; Lavezzo et al., 2020), social distancing (Baumgartner et al., 2020; N. Ferguson et al., 2020; Siedner et al., 2020; Tian et al., 2020; J. Zhang et al., 2020), contact-tracing (Aleta et al., 2020; Ferretti et al., 2020), quarantines (Aleta et al., 2020; Pan et al., 2020), social bubbles (Block et al., 2020), mask-wearing

[1] https://www.worldometers.info/coronavirus/; 16.09.2020

(Brauner et al., 2020; Goldberg et al., 2020), business closures (Brauner et al., 2020), and school closures (Davies et al., 2020; Prem et al., 2020; J. Zhang et al., 2020).

Some studies have made use of the fact that systematic differences in policies and local conditions have resulted in natural experiments that allow for investigations into causal mechanisms. In most cases, the ecological validity of natural experiments (Haushofer & Metcalf, 2020) involves a tradeoff with internal validity, as naturally occurring "control conditions" often feature confounds and unobserved variables. A different strand of research focuses on the development of technological tools to support or construct counter-measures, such as the COVID Symptom Tracker (Drew et al., 2020) or digital contact tracing (Ferretti et al., 2020).

## E2.2 Contributions of the behavioral sciences

### *E2.2.1 The need for the behavioral sciences*

Given the nature of the threat in the current COVID-19 crisis, initial responses naturally deferred to medical and epidemiological advice. Yet, studies on population-wide interventions stop short of demonstrating how these interventions can be successfully implemented when they depend on the co-operation of millions of citizens. Van Bavel et al. (2020) highlighted the role of behavioral science in providing tools for large-scale changes in behavior and in assisting decision-makers in understanding and managing the crisis. Insights from behavioral science inform government messaging and communication (Dorison et al., 2020; Lunn et al., 2020; Michie et al., 2020), offer a critical reflection of fundamental response strategies (Chater, 2020), and help to create and support interventions (West et al., 2020). Moreover, social scientists have collaborated to compile large and international survey datasets (e.g., Perrotta et al., 2020; Yamada et al., 2020). The need to react to a rapidly developing situation has resulted in calls to reform behavioral research and suggestions for how the research cycle can be aligned with the necessity for timely interventions (Hahn et al., 2020).

### *E2.2.2 Surveys, Experiments and Games*

Behavioral scientists responded to the crises quickly and primarily conducted surveys, reporting on reactions to and perceptions of the crisis, adherence to counter-measures, and statistical predictors of these reactions. While such studies can contribute to identifying relationships and possible mechanisms, they may not be sufficient for laying a foundation for large-scale interventions. Randomized control trials in the lab (Haushofer & Metcalf, 2020) together with considerations for how to scale them up to the desired level of implementation (Al-Ubaydli et al., 2020) seem to be ethically and functionally advisable as a preceding step.

Behavioral game paradigms have been successfully used to study human behavior and motivation, but there have been fewer applications to epidemiology. Ibuka et al. (2014), Böhm et al. (2016), and Lim and Zhang (2020) used the "vaccination game", a social dilemma to study the willingness to symbolically vaccinate (lose part of a monetary endowment) in order to achieve "herd immunity" with benefits for the group (a monetary bonus). They investigated predictors for and the likelihood of free-riding. Another recent study used a variant of the dictator game, the "risk dictator game", to study the link between the

willingness to incur risks for other participants and the willingness to take counter-measures against Covid-transmission (Campos-Mercade et al., 2020). Earlier, epidemiologically relevant observations were made—unexpectedly—in a very different game environment. When a programmed disease was set free in the open world of War of Warcraft, multiple epidemiological mitigation strategies failed until the server had to be reset due to the high number of player characters succumbing to the virus. Based on these observations, Lofgren and Fefferman (2007) highlighted under-utilised potential of game environments for gaining insight into the dynamics of outbreaks.

### E2.3 Brief summary of game paradigm and chosen interventions

#### *E2.3.1 Rationale for the construction of the transmission game paradigm*

Our study proposes a novel paradigm to experimentally study decision-making in pandemic-like scenarios and the effectiveness of intervention measures. In developing the transmission game paradigm, we were guided by a list of desirable properties that we wanted to achieve by design: We wanted to capture central elements of the health crisis and introduce them in a neutrally-framed game scenario. Participants in large groups were to encounter the equivalent of a spreading infection among their group and face a social dilemma. Individually profitable actions as opposed to safer and less profitable actions should dynamically increase the risk for self and others to incur severe losses at the end of the game. We intended the game paradigm to be communicable to broad parts of the population and to be flexible enough for groups of various sizes (including very large groups), both offline and online. In addition, the paradigm should be deception-free (Hertwig & Ortmann, 2008). We wanted participants to make choices independent of each other to increase the statistical power of group comparisons. Finally, we wanted to create a paradigm that would allow us to test a range of interventions. We explain and justify the details of the game and interventions in separate sections (section E3 and section E9).

#### *E2.3.2 Choice of interventions for Study 2*

We implemented five different interventions in Study 2. They are explained and justified in detail in section E9. Specifically, we included a vicarious learning intervention showing the percentage of "infected" participants in a previous game (Study 1) up to the round for which a decision had to be made. Further, we tested two norm-based interventions focusing on how many participants chose the low risk option in the previous game and a brief message explaining the negative consequences of choosing the higher risk option for oneself and other players. Finally, we included two different boosting interventions (1) highlighting how high-risk choices can initiate cascading consequences and result in transmission chains of infection and (2) illustrating the game dynamics through a simulator of game choices and outcomes.

### E2.4 Study goals

#### *E2.4.1 Goals of Study 1*

In Study 1, we implemented the paradigm without interventions in a single group of 100 participants on Prolific, all located in the United States. In addition to the game, we

included a postquestionnaire, a broad range of possible predictors, and several survey scales measuring COVID-19-related attitudes. The preregistered study had three main goals: (1) demonstrate that participant behavior shows substantial variance, (2) collect data for realizing some of the interventions planned for Study 2, and (3) perform an exploratory analysis of predictors and select a promising subset for Study 2.

#### *E2.4.2 Goals of Study 2*

In Study 2, we compared game behavior in a control condition (between-subjects) with behavior in five intervention groups (one group per intervention). A total of 600 players participated on Prolific, all located in the United States. There were two main goals in Study 2: (1) Estimate the relative effectiveness of the five interventions in reducing the number of risky decisions by participants, and (2) attempt to replicate the findings for predictor variables in Study 1.

## E3 Transmission Game Framework

In this section, we will present the game rules with some illustrations and discuss design considerations for the framework. We will also discuss why some of these guiding principles were chosen and how they were implemented. We will also refer to the framework as an experimental paradigm, especially when discussing its implementation in the laboratory.

### E3.1 Rules of the standard transmission game

For easy reference, we offer a brief summary of the rules at this point. We also reproduce the instructions given to participants in the "Selected Materials" section in this document.

We will start with a brief list of all rules and then explain their implications in more detail.

1. There are 100 players in the game.
2. Each player is in one of two states, blue or purple.
3. No player can see their own state or the state of other players.
4. A player can change from blue to purple but a purple player stays purple until the end of the game.
5. At the start of the game, all 100 players are blue.
6. The game lasts 25 rounds.
7. Before the first round only, eight randomly selected players change into purple players.
8. At the beginning of each round, players are paired at random.
9. Each player chooses an action, G or H for the round.
10. A player who chooses G receives 8 points, a player who chooses H receives 40 points.

11. A blue player that is paired with a purple player may change into a purple player.
12. If none of the paired players chooses H, a blue player that is paired with a purple player changes into a purple player with $p = 0.05$.
13. If any one but not the other of the paired players chooses H, a blue player that is paired with a purple player changes into a purple player with $p = 0.15$.
14. If both players choose H, a blue player that is paired with a purple player changes into a purple player with $p = 0.25$.
15. The game proceeds until the end of round 25.
16. If a player is blue at the end of the game, the their points scored across the 25 rounds are converted into a bonus payment at a rate of 2 points: 1 penny [0.01 GBP].
17. A purple player does not receive a bonus payment; scored points are thus valueless for purple players.

It should be noted that a further rule is implicit due to the asynchronous nature of play: No player observes the actions of other players. This conveys the two benefits discussed above (i.e., the paradigm is scalable to groups of different sizes, and independent choices allow for more statistical power).

### E3.2 Illustrations of game rules and implications

To illustrate the game rules and their implications, we will use several illustrations of players and populations. We will represent players by blue and purple stones (see Fig. E1).

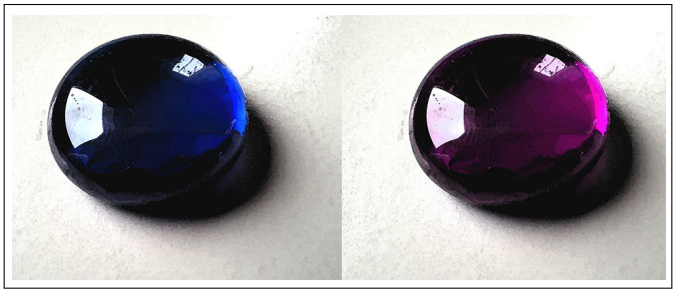

**Figure E1**
*Representation of players: two possible states.*
*Photo Credit: Jan K. Woike, University of Plymouth.*

Every player starts the game in the blue state, and the initial population is therefore represented by 100 blue stones (see Fig. E2, left side). Before the first round, eight randomly chosen players change their color to purple (see Fig. E2, right side).

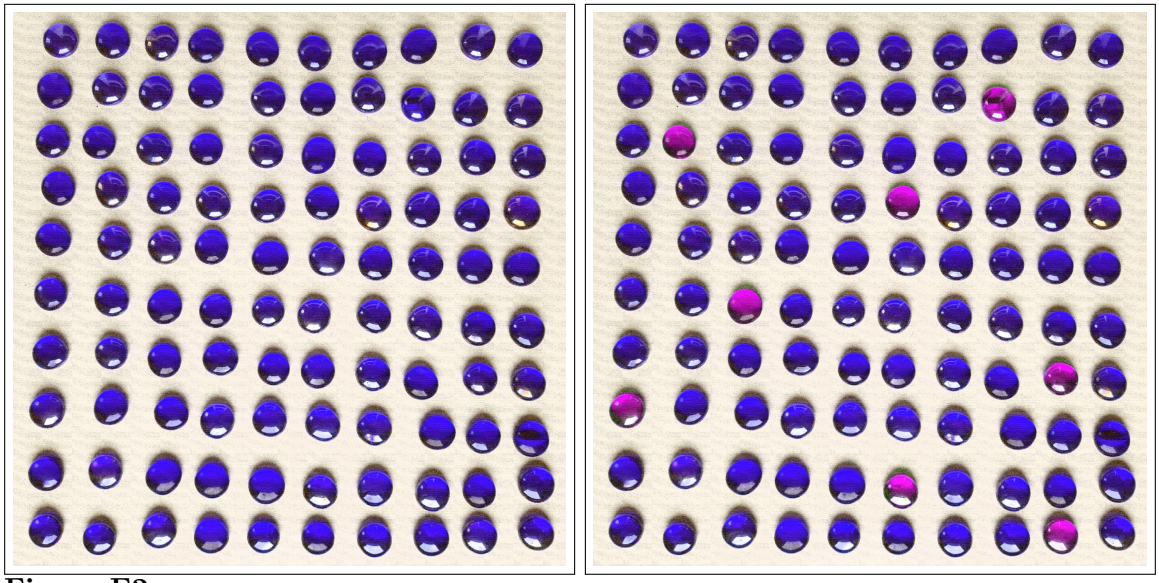

**Figure E2**

*Population at the start of the game (left side) and before the first round (right side).*
*Photo Credit: Jan K. Woike, University of Plymouth.*

The pairing of players in each round corresponds to randomly rearranging the stones in rows of two (see Fig. E3, left side). Each pairing results in one of four combinations of stone colors: blue/blue, blue/purple, purple/blue, or purple/purple. For two of these combinations, the actions chosen by the players have no impact on their color at the end of the round (see Fig. E3, right side): Two paired blue players will stay blue and two paired purple players will stay purple. Their action choice will still have consequences for the points they receive in the round: Players choosing action G receive 8 points, players choosing action H receive 40 points[2].

[2]Receiving points does not equate to receiving bonus money. If both players are purple, their points will ultimately not be converted into bonus payments.

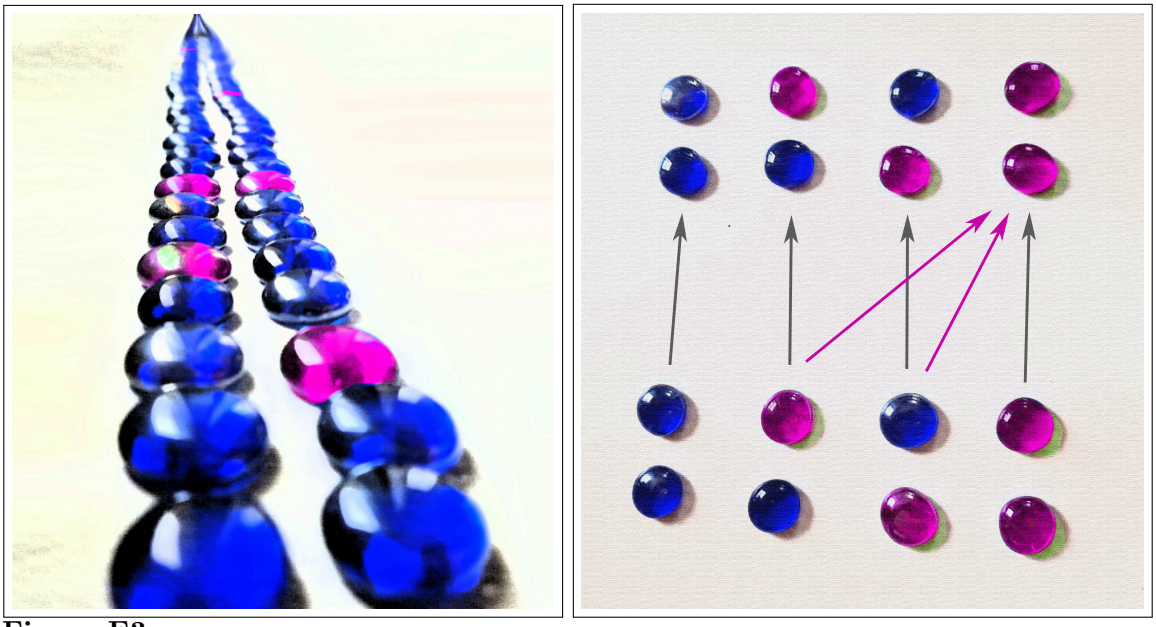

**Figure E3**
*Pairing of players (left side) and possible color change (right side).*
*Photo Credit: Jan K. Woike, University of Plymouth.*

A color change can only occur, if a blue player is paired with a purple player. In this case, the blue player may change into a purple player (not the other way round). The minimum probability is $p = 0.05$ (if both choose action G), but it increases by 0.1 for each H-action chosen by any of the two players up to a maximum of $p = 0.25$ (see Fig. E4)[3].

| | $\mathbf{G}_j$ | $\mathbf{H}_j$ |
|---|---|---|
| $\mathbf{G}_i$ | 5% | 15% |
| $\mathbf{H}_i$ | 15% | 25% |

**Figure E4**
*Probability of switching from blue to purple conditional on player actions.*

These are all of the game mechanics. Figure E5 illustrates how a game could progress to the final round. Note that the stones representing specific players remain in fixed positions and purple players stay purple until the end of the game.

[3]To stay within the stone version and to implement this mechanism mechanically, the blue player's new state could be randomly drawn out of 20 stones, of which one, two, or three were purple, depending on the chosen actions.

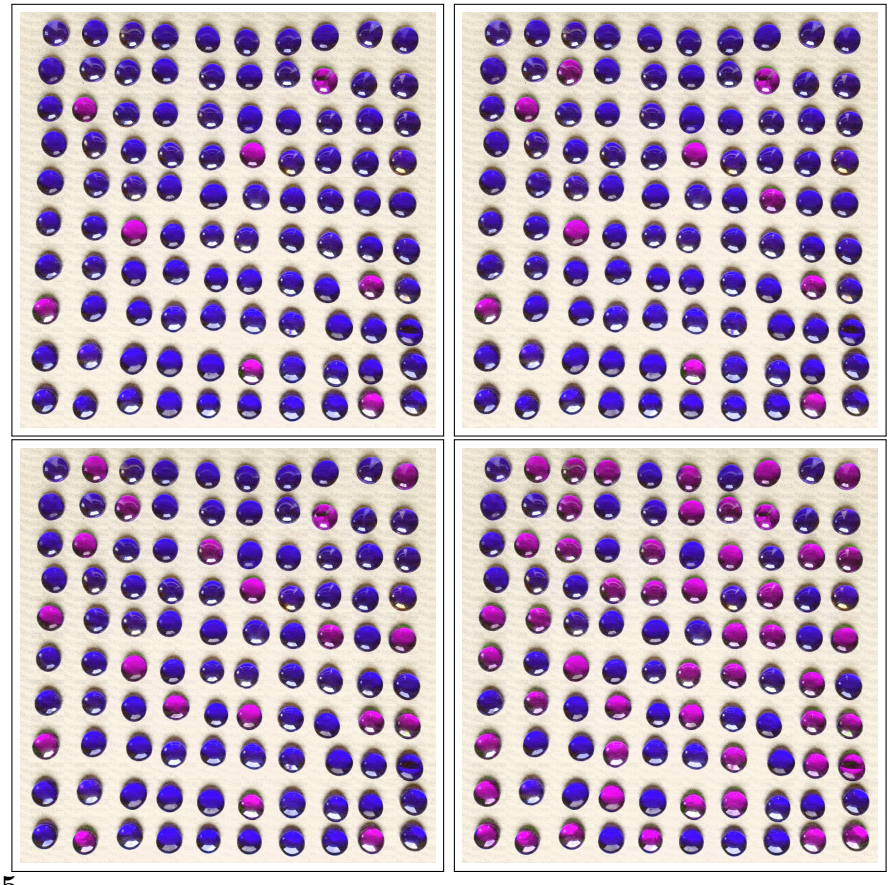

**Figure E5**
*Progression of the game in four steps to the end (from top left to bottom right): Stone colors are shown after rounds 1, 5, 15, and 25.*
*Photo Credit: Jan K. Woike, University of Plymouth.*

### E3.3 Development of the framework

In this section, we provide details of how we implemented the desirable properties mentioned above and justify some of our decisions. Our goals were the following:

1. Map relevant elements of the COVID-19 health crisis onto game elements.
2. Establish a neutral and non-politically charged context.
3. Create a paradigm that is communicable to the general public, comprehensible and simple enough.
4. Design the paradigm to be deception-free.
5. Be able to apply this paradigm to groups of various sizes including very large groups.
6. Be able to collect data online and offline.
7. Collect data from individuals independently, preserving power for the analysis.

8. Design the paradigm to be flexible enough to test multiple forms of interventions.

#### *E3.3.1 Neutral framing*

We decided to implement our framework using neutral terms (e.g., color change) to allow for some distance to the ongoing COVID-19 pandemic. Research practices differ between disciplines and this decision might need some explanation. Ariely and Norton (2007) discussed the methodological differences between economists and psychologists as a "gap in abstraction".

Our goal was to mirror the incentive structure inherent in the problem situation and include the relevant elements that form the input for psychological processes that result in decisions. In doing so, we do not want to imply that the semantic context does not matter, but that a rich context adds elements that would pose substantial problems for the analysis of interventions (as explained below) that lies at the center of our interests. Note that a similarly neutral framing is employed in the Reckless or Responsible game (Erev et al., 2020; Plonsky et al., 2021; Roth et al., 2020), and the authors likewise chose a neutral framing to "avoid framing and impression management effects" (Roth et al., 2020).

Bardsley et al. (2020) argued that the isolation of factors in the laboratory can be considered as one of its biggest advantages, "since without it the interplay of causes in the outside world cannot be disentangled" (p.215). In line with our reasoning, they regarded mimicry of contexts as a reasonable basis for generalization, but neither as a requirement nor as the best choice in all situations. Further, an empirical comparison of variations of context across several experiments in political science by (Brutger et al., 2020) produced the finding that choices between levels of abstraction produced very limited changes in behavioral observations and they found no impact on central findings in conceptually replicated studies. Below, we explain why we would consider a rich context a threat to reaching the aims of our study, why we consider it a threat to replicability and, in addition, ethically problematic given the interventions in our experiment.

**Exposure to real-world interventions.** There are good reasons to assume that a thick framing would create particular problems in this study. Several months into the pandemic, all participants are likely to have been confronted with a wide variety of messages through multiple channels. Participants on crowdsourcing platforms at the time of the study were also exposed to a multitude of pandemic-focused surveys and experiments. Using a strong pandemic framing might render working interventions seemingly ineffective, because they copy interventions participants have been already exposed to (see, e.g., Gelfand et al., 2021). This might create a misleading picture of relative effectiveness, and this point is of particular concerns as some of our interventions were modeled after interventions observed in the early phases of the crisis.

Participants would be able to rely on beliefs or attitudes formed in response to the current pandemic over extended periods of time. As a consequence, their beliefs and attitudes towards the topic would likely be independent of the game scenario but might be used as shortcuts to decision-making. Thus, participants may become unresponsive to manipulations and crucial scenario features. Of course, these beliefs are certainly not irrelevant for their response to the health crisis in the real-world, but our interests centers on interventions that help in response to a novel situation, we are less interested in interactions between uncontrolled experienced interventions and our manipulation. Any result

using a thick context would become fundamentally ambiguous and interwoven with individual learning experiences outside the lab. Abstraction should increase the replicability of findings, even when attitudes towards a specific health crisis shift.

**Country-specific conditions: Politicization of the pandemic.** A rich context would also impact the local specificity of observations. The situation in the United States is characterized by a strong politicization of both the pandemic and interventions that seems specific to both place and time (Gelfand et al., 2021; Rothgerber et al., 2020; Xu & Cheng, 2021). Political conservatism was found to be a better (negative) predictor of compliance with countermeasures than case numbers in the US. This has been linked to targeted messaging and misinformation by conservative media personalities (Graham et al., 2020; Rosenfeld et al., 2021). At the same time, the messaging about masks underwent repeated shifts (Manjoo, 2020), so that results obtained at any point in time during the pandemic have to be seen in light of the most recent statements. While these processes are certainly interesting and could be studied through variation in framing, they are again obstacles for the question at the center of the present study. Different health crises in the US have led to different reactions by the parties (Beall et al., 2016; Nyhan, 2014; Van Bavel, 2020). Here, we are interested in how participants' response to a novel situation can be influenced by interventions, not in the specific state of the political discourse. Thus, we hope that our results can be replicated across time and—to a reasonable degree—across locations.

**Protecting participants.** In addition to the potential interaction with existing interventions, there is an ethical argument to be made for a more neutral framing: Neutral terminology that does not directly relate to the pandemic reduces the risk of boomerang effects in interventions. For example, if one observes an increase in undesired behaviors in a neutral scenario these behaviors remain confined to the lab setting. In contrast, a reduction in willingness to engage in countermeasures against the virus may lead to serious consequences for the health of participants and those they encounter outside the lab. In discussing our results we will point out incidences in which this negative impact from study setting to the real world indeed occurred.

#### *E3.3.2 Relevant features of the health crisis*

From the beginning our focus was on the complex dilemma involved in mitigating the effects of the novel coronavirus. Social distancing, mask-wearing, and taking individual precautions are costly, both for the individual and organizations. At the same time, the consequences of not engaging in these and other counter-measures are costly as well. Violations of preventive rules can provide benefits to the rule violators, but incur costs on others who are negatively affected by it. For example, younger people contracting the virus might expect a milder course of the disease, but they can infect others in vulnerable groups whose experience might be much worse. Similarly, current evidence suggests that masks have a protective effect both for the wearer and those in proximity to the wearer (Howard et al., 2021; Peeples, 2021). Importantly, the protective effect for those close by might be stronger than the self-protective effect for the wearer.

In addition, the implementation of counter-measures poses a collective action problem in the sense that they are personally costly but will remain ineffective if an insufficient number of people adheres to them. At the same time, widespread adherence might confer

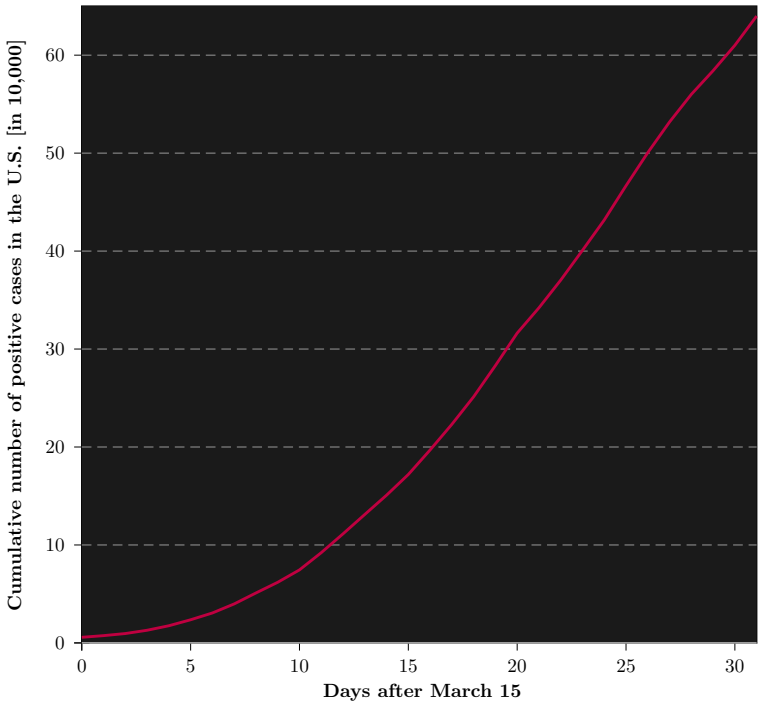


**Figure E6**
*Cumulative number of positive cases in the United States since March 15, 2020, based on the data provided by the COVID Tracking project.*

benefits even to isolated violators. As a consequence, groups may suffer from the prevention paradox: Due to the absence of negative consequences, people may question the legitimacy of the measures that prevented those negative consequences in the first place. This could easily lead to a situation in which free-riders—who only remained unharmed due to widespread adherence—might become seemingly successful models for those who restricted themselves.

Another relevant aspect of the current crisis is the exponential nature of virus transmissions. At the beginning of the pandemic in the United States, the number of observed cases since the first case followed a clearly non-linear pattern. Figure E6 shows the development of total case numbers for the days following the 15th of March 2020.

Linear extrapolation of infection fails to appreciate the risks associated with rising numbers and accelerated growth. Likewise, daily interactions cannot be simply classified into high-risk and safe actions. Many interactions involve a form of micro-risk that accumulates over time with pernicious effects for entire groups. While a single interaction might be relatively safe, a large-group event compounds these risks into foreseeable case clusters.

### *E3.3.3 Mapping these features*

Mapping the features of the current crisis onto elements in the game involved both a positive selection of features to include and a negative selection of features not to map. We decided to exclude a number of elements that play a role in the current crisis to keep

the game rules at a level accessible by a typically motivated survey participant. While we made this selection for the standard paradigm, there is no reason not to include (some of) these features in future variants of the game. We would advise not to include everything at the same time, though. We believe that the current level of complexity should not be exceeded by more than a few additions.

**Health status.** To move from a health domain that—given the current political discourse—intersects with the political domain into a more neutral territory, we chose to simplify the scenario by replacing health status by a simple pair of colors: blue and purple. Blue represents a healthy individual, purple represents an infected individual. We did not choose green and red (Lim & Zhang, 2020), as these colors have multiple connotations due to their use in warning signs and standard icons. Attitudes towards the subject matter can override the objective incentive-structure in a game scenario, as shown in Böhm et al. (2016). While these attitudes are relevant for interventions in the health domain, the history of belief formation lies completely outside of the game paradigm and would add condition-independent variance. Further, results would be very dependent on time and location of data collection.

**Binary state.** We simplified the complex medical relationships by reducing health status to these two states. In our paradigm, we ignored the differentiation between mild and severe cases, and between infectious and non-infectious states. We also eliminated the delay between the moment of infection and becoming infectious and spreading the virus. We did not introduce symptoms that would allow interaction partners to adjust their risk expectations. This corresponds to the estimated impact of asymptomatic spread of COVID-19. In addition, we reduced the number of state transitions to a single possibility: blue participants can change to purple but not the other way round. A purple participant stays purple until the end of the game. Thus, we excluded one feature present in the current crisis, namely the possibility of recovering from the infection (Gatto et al., 2020).

**Consequences of color change.** The outcome of contracting the novel coronavirus is stochastic, influenced by age, ethnicity, occupation, pre-existing conditions, medical insurance, and access to healthcare (Fielding-Miller et al., 2020; Lan et al., 2020; Mukherji, 2020; Rentsch et al., 2020; Simione & Gnagnarella, 2020). Consequences range from mild discomfort to death, but we simplified all of this in the game to a single outcome: Changing one's color to purple implies losing all potential bonus money for the game. Obviously, these consequences are less severe than the consequences of contracting the actual virus, but they correspond to the maximum level of negative consequences that can be ethically justified for participants in our experimental setting. Base payments guarantee that participants are not exploited, while maintaining the disincentive associated with switching colors (losing all bonus payments).

**Time period and initial state.** To study transmission we created a game with multiple rounds that starts with a low proportion of purple participants. This corresponds to the state of nations at the beginning of the epidemic after small numbers of positive cases were reported. We chose a length of 25 rounds to allow for the potential of considerable spread. Rounds have no literal correspondence in the real-world, but might realistically map onto time periods of about a week. At the beginning of the game, 8 players are randomly flipped to purple. We chose this number as it is small enough for people to consider themselves blue with high confidence and large enough to allow for the desired

growth dynamics (see below). While the current crisis appears to occur in waves with case numbers rising and ebbing (based on seasonal variation, and levels of regulations, interventions, and risk-taking in a society; Leung et al., 2020), our game focuses on a single wave.

**Interactions and transmission.** We simplified virus spread through a population by only considering dyadic interactions. In each round, players are grouped in pairs and have one interaction in their pair, resulting in a total of 50 interactions per round (in a population of 100 players). We did not introduce tie strengths or social networks (Christakis & Fowler, 2013) into the game, even when similarity (Centola, 2011) could increase the likelihood of interaction and people might attend differently to others depending on their distance in social networks (Bond et al., 2012). Furthermore, we only allow color changes when a blue player is paired with a purple player, and only from blue to purple. This corresponds to the following transmission characteristics present in the current situation: (a) infected individuals (purple) can transmit the disease to healthy individuals (blue), but not vice versa, (b) healthy individuals cannot infect each other, and (c) already infected individuals cannot be infected again (while being infected).

**Protective measures and points.** The risks of any real-world interaction can be mitigated by protective measures. While several types of these measures (e.g., mask-wearing, social distancing, hand-washing) exist, we simplified the choice to engage in these mitigating behaviors to a binary choice. To reduce moral and political connotations, we offer two abstract actions, namely "G" and "H". Action "G" represents a choice in favor of mask-wearing and social distancing, and action "H" represents a choice against these measures. Protective measures incur costs in terms of materials, time, and restricted opportunities to engage with others both for personal and commercial purposes. To mirror these differential costs, we assigned a score of 8 to the restricted behavior (action G) and a score of 40 to the unrestricted behavior (action H)[4]. All players choose one of these two actions in each round of the game.

We simplified participants' choices to two actions; yet, people can show more nuanced behaviors during a pandemic. We also simplified consequences to two fixed values. Outcomes in a pandemic depend on circumstances and timing: For example, there is a difference between postponing non-essential medical procedures out of consideration for increased risks and postponing the treatment of acute emergencies (Kazi et al., 2020; Westgard et al., 2020). Moreover, the adverse impact of self-isolation on mental health (Ammerman et al., 2020; Azim et al., 2020; Filgueiras & Stults-Kolehmainen, 2020; Hyland et al., 2020; Okruszek et al., 2020; Wissmath et al., 2020) as well as physical health can compound over time, for example due to decreases in physical activity and changes to eating habits (Ammar et al., 2020). Finally, one may argue that individual outcomes depend on group outcomes: If high infection rates impact national economies, most actors within these economies might suffer as a consequence.

[4]Multiplying both values by a constant would not change the game dynamics (and could result in the same payoffs if the score to payment ratio is changed accordingly. Adding a constant to both values *can* change the dynamics, a positive value would make the G-action relatively more attractive. The stakes of the game—in terms of bonus money participants can gain or potentially lose—depend on the maximum and minimum possible score and the score-to-payment ratio.

**Protective measures and transmission risk.** To map the relation between risk-mitigating behavior and transmission, we implemented transmission risk from a purple to a blue player as a function of both players' action choices. If both players choose action G, the blue player has a risk of 5% to turn purple. If one of the two players chooses action H and the other player chooses action G, the risk of turning purple increases by 10% to 15% (it does not matter which player chooses H). If both players choose action H the risk of turning purple increases to the maximum of 25%. We chose this value as it reflects the fact that risky interactions do not result in an infection as the most likely outcome, even for an infectious disease such as COVID-19. At the same time, these numbers are close to risk assessments measuring, for example, the protective value of mutual or non-mutual mask-wearing, for example. We simplify, of course, by reducing the risk of one time period to a single interaction (and the game to 25 rounds). We also stated probabilities as objective probabilities that are common knowledge. Public discourses about masks indicate that there is variation both in government guidelines regarding the use and efficacy of masks, and in people's perceptions of their benefits (Leder et al., 2020). Moreover, the protective function of masks for wearers and others around them are likely asymmetric (Pfattheicher et al., 2020), which further complicates incentives and preferences for wearing them (Betsch et al., 2020; Campos-Mercade et al., 2020).

Even after our data collection was completed, data on the effectiveness of protective measures only allowed for estimates with a high degree of uncertainty. For example, Howard et al. (2021) argue in their review of the evidence for face mask wearing, that "understanding of the relative merits of different cloth mask designs and materials is still limited", while the authors still unambiguously recommend mask mandates in combination with other measures, such as physical distancing. Rader et al. (2021) likewise find strong effects of mask-wearing, but their modeling approach stresses the simultaneous consideration of physical distancing with widely different consequences of mask-wearing depending on the level of physical distancing.

Transmission risks also fundamentally depend on the virus variant. One example with a similar relationship between the adoption of preventive measures and outcomes as in our game is given by Y. Zhang et al. (2021). In their model, total infection rates for elementary schools after 90 days can vary between slightly above 22% and close to 90%, depending on the use of masks and regular testing (see also Cha, 2021). This is comparable to the minimum and maximum final infection rates that can be reached in our games based on participants level of risk-taking.

Our framework does not commit to a precise definition what one interaction with another player represents in terms of time scale or type and duration of contact. In this sense, one round could represent a week of interactions, single interactions or anything in between. Transmission risks can be reduced by physical distancing, and many models of transmission allow for continuity in the reduction of risk. For example, De-Leon and Pederiva (2020) include the physical distance between individuals in their model equation, Contreras et al. (2021) model the effective transmission rate as a fraction of $R_0$ proportional to the number of contacts relative to the number of pre-pandemic contacts. Our parameter settings constitute special (but not irregular) cases of either conception of reduction of transmission risk.

**Stochastic elements.** We reduced the number of stochastic influences to three elements: (1) the initial selection of purple participants, (2) the round-wise matching of participants, and (3) outcomes of pair-wise interactions in mixed-color pairs. In contrast, point scores for actions and the consequences of color change are deterministic. Also, after the game has started, blue players cannot turn purple if they do not interact with purple players. Overall, we aimed at creating a game that simultaneously conveys uncertainty about one's own state (blue or purple), the state of one's interaction partners, and the outcomes of interactions and the entire game. At the same time, we did not want to create the impression that outcomes are entirely random and ultimately unrelated to one's choices.

**Non-linear development.** To create the potential of non-linear growth of the number of purple participants in the population, several requirements have to be met. The population has to be large enough to allow for a gradual acceleration of the infection rate until the majority of the population is infected (and a slowing convergence against total infection afterwards). Simultaneously, we needed a game that lasts long enough for these phases to be reached. To meet these requirements, we chose a group size of 100 participants to balance the benefits of large populations with the costs of collecting data for a single condition[5]. We first measured the impact of several design parameters on hypothetical game outcomes, namely, the initial proportion of purple participants, the probabilities of color change during the game, and the number of rounds. Lower initial rates of purple participants necessitated either more game rounds or higher probabilities of color change during the game. Increasing the initial rate had a strong impact on the number of rounds necessary before differences in behavior could result in clear differences in final color proportions. We finally chose an initial rate of 8% to achieve a compromise: Switching to purple before the first round stays relatively unlikely, and we were able to keep the maximum color change risk at 25% and the number of rounds at 25. The simulation results for these parameters (see **SCR** and the simulation section below) confirm that this creates the desired progression in purple rates.

#### *E3.3.4 Asynchronous implementation in an online setting*

We designed the paradigm to be implementable in an asynchronous online setting. We considered asynchronous gameplay as the preferable option for two main reasons: selective sample attrition and non-independence of observations.

**Selective attrition.** The first reason concerns attrition. Large synchronous groups necessitate waiting rooms and many rounds in a game require participants to wait for the slowest participant (or the count-down of a timer). Almost inevitably, some participants will miss round actions due to connection issues or sample attrition (Arechar et al., 2018). Allowing each participant to start and stop on their own does not filter participants who are impatient or who are indifferent to leaving groups behind with potentially negative consequences for bonus payments. This type of attrition cannot be assumed to be independent of study questions. As a point of comparison: Molnar (2019) limited the maximum number of synchronous players in SMARTRIQS to eight out of concern that higher numbers of players are unlikely to generate complete group responses.

[5]Obviously, the game rules can be easily adapted to larger and smaller populations.

There is evidence from many studies that this type of attrition happens to a substantial degree on online platforms, even with small groups. Synchronous experiments on crowdsourced platforms have been conducted in the past decade (Mao et al., 2011). Roth et al. (2020) excluded participants with more than 20% of choices not given within a time limit and lost 43% of the original participants in two studies (152 of 352). Arechar et al. (2018) observed an attrition rate of 18% (84 participants) despite taking steps to keep this rate as low as possible. Suri and Watts (2011) accepted up to 10% missing responses that were replaced by random computer choices. They had to resort to this procedure even after going through a multi-step selection procedure for participants. They report that studies with 24 simultaneous participants could not be conducted ad hoc due to lack of workers in the waiting rooms. They solved this problem by teaching game rules in smaller group-games and invited participants from a selected panel of about 150 participants through direct messaging via the platform to conduct many experiments. Mason and Suri (2012) admit that even when following this procedure the data of 21 out of 94 experiments with 20–24 players had to be discarded (22% of experiments), and suggest an upper limit of 4–8 participants for successful synchronous experiments—the same limit implemented by Molnar (2019) in SMARTRIQS.

The attrition observed in the studies is most likely selective at multiple steps in the process. Waiting rooms are built on the first-come first-serve principle with late participants excluded from the study (Mao et al., 2011). This is likely to select for attributes such as conscientiousness and against participants who are more likely to be distracted temporarily in their working environment. Attrition during the study is often caused by impatience with waiting times caused by participants taking different amounts of time for entering roundwise decisions.

Even among impatient people, not everyone would be willing to drop out of a running experiment as this creates costs for everyone else involved in the game (including the experimenter). It is not far-fetched to assume that there might be systematic differences in social preferences between those dropping out and those staying in the game. Finally, technological problems might cause participants to drop out involuntarily. These problems are sometimes randomly distributed, but participants with better equipment are less likely to suffer from technological problems than others, potentially linked to socio-economic status,

Why does this type of selective attrition constitute a problem for our study? If attrition affected each of the intervention conditions equally, the main harm would be distortions in the studied samples drawn from the population of platform participants. Unfortunately, impatience and lack of self-control have been identified as key variables in predicting resistance against interventions (Graham et al., 2020). Moreover, our intervention conditions offer different amounts of information to participants. Offering more information will increase the variance in response times when some participants try to take this information into account when making their decisions and other participants do not. Differential familiarity with graphs will increase this variance. Thus, it would be likely that conditions with larger variance in response times lose more participants than others. As these participants tend to be impatient, this might lead to seemingly paradoxical findings as reported in Zhou and Fishbach (2016).

**Non-independence of observations.** The second reason concerns response independence. If group members' decisions are not influenced by other members, individual

choice sequences are independent from each other. At the same time, responses are not influenced by the stochastic outcomes (i.e., potential color changes) in each round, as color information is not disclosed to participants during the game. This allows for a separation of data analysis and payoff calculations: If there were six games with the same rules and with 100 participants per game (600 total), data analysis could treat the 600 participants as independent, and expected game results could be computed by repeated sampling of 100 participants out of the full 600. Only the calculation of payoffs would require six actual games to be played out by assigning the participants to six separate games and simulating each of them exactly once. This would fulfill the requirement for a deception-free implementation of the paradigm.

A non-independent realization would require a hierarchical design with participants clustered in six games (with fixed assignment) and severely reduce statistical power: the effective sample size would be equivalent to the number of groups, requiring 1,000 participants for an effective sample of 10. Our goal in this study is to compare the effectiveness of interventions, and increases in effective sample size allow us to detect differences between interventions.

**Synchronous versus asynchronous games.** Given the problems described above, it is understandable that asynchronous designs are very popular when working with online samples even in recent years (Evans et al., 2021). The literature on comparing behavior between synchronous and asynchronous designs draws a mixed and somewhat inconsistent picture regarding differences. Part of this literature focuses on differences between employing the strategy method (making decisions for all possible actions of asynchronous others) versus responding to actions by synchronously connected players. Some studies, such as Brandts and Charness (2000), did not find any behavioral differences between these two variants. Surveying a range of 29 studies, Brandts and Charness (2011) saw differences reported in a minority of these studies (nine showed mixed evidence, and only four showed clear differences). These differences did not rise to the level of observing novel effects by using the strategy method. More recently, Columbus and Böhm (2021) found more selfish behavior when using the strategy method compared to a simultaneous decision protocol.

More relevant for our design are the findings by Evans et al. (2021) who similarly studied a Prolific sample and compared a simultaneous matching protocol for a prisoner's dilemma with an asynchronous variant. The authors likewise used comprehension questions and used a postquestionnaire to inquire about participants' experience of the game with a sample of over 1,200 participants. They found no significant differences between conditions and also observed no differences in perceived realism. Evans et al. (2021) concluded that "synchronous designs and asynchronous designs can produce similar results in studies of online cooperation" (p.838).

In line with the lack of differences in perceived realism, we would argue that there is no guarantee that participants trust studies with synchronous interaction more than studies with asynchronous interaction. Unfortunately, some researchers on crowdsourced platforms advocate for the use of deception, even when their studies could have been perfectly realized without it. Summerville and Chartier (2013) advocated for using pseudo-dyadic interaction by making participant believe that they are matched with partners that in reality do not exist[6]. The simultaneous presence of deceptive and non-deceptive variants of synchronous

[6]Even these authors advise against using groups larger than ten. They argue that participants would not

games destroys any assumed advantage for synchronicity.

**Lowering the barrier for replications.** We should mention one further important advantage of the asynchronous format. We believe that the comparatively reduced demands regarding technology lower the barriers for those who are interested in replicating the work. This can include those working inter-culturally and in environments with restricted access to technology. For this purpose, we also discuss a paper-and-pencil variant of the game that can be played and evaluated without the need for computers (see section E14.1). We actively invite other researchers to test this game in a variety of settings to learn more about the temporal and cultural stability of game behavior and its behavioral determinants.

#### *E3.3.5 No deception*

It is crucial for us that implementations of the paradigm are deception-free and we consider it unacceptable to use fake participants, fake choices by other players, etc. We reject the use of deception both for ethical and pragmatic reasons. Deceptive paradigms should not be used when non-deceptive alternatives are available, according to ethical guidelines for experimentation (American Psychological Association, 2017; Kanngiesser et al., 2020). At closer inspection, there is rarely a case that unambiguously warrants the choice of deceptive methods based on this principle. On a pragmatic level, participants' trust in the veracity of the game rules is necessary for the game to tap into social preferences at all. Reducing the game to an individualistic optimization exercise makes it inappropriate for exploring interventions. Experiencing deceptive variants will shatter trust in the existence of other players and change future game behavior. Thus, decisions by some researchers who do not want to spend funds for real participants or to invest effort into implementing deception-free paradigms have the potential to hurt future research (see also Hertwig & Ortmann, 2008; Ortmann & Hertwig, 2002).

#### *E3.3.6 Balance between complexity and comprehension*

As we have described throughout this section, there were several components of pandemics that could be added to a game scenario, but that we decided not to include. The reason for this decision is that we needed to strike a balance between realism and comprehensiveness on the one side and feasibility and understanding on the other side. We decided to develop a paradigm that can be used with diverse samples, not a selected group of trained university students in technical disciplines.

**Parameters in the basic game.** Our final design can be fully characterized by nine parameters (see Table E2).

The number of rounds (1), the transmission probabilities (46), the values of the action (78) were shown at each round in the game. The conversion rate is applied to each score, indicating the possible payoff for the points scored up to the current round. Thus, there is no need to memorize numbers with the exception of the population size and the initial color switch rate (which is asked for in comprehension questions and clearly stated in the instructions).

---

believe to be matched in larger groups without long waiting times.

**Table E2**
*List of parameter values*

| Number | Parameter | Value |
|---|---|---|
| 1 | Number of rounds | 25 |
| 2 | Number of participants | 100 |
| 3 | Number of purple participants at the start | 8 |
| 4 | Transmission probability: GG | 0.05 |
| 5 | Transmission probability: GH/HG | 0.15 |
| 6 | Transmission probability: HH | 0.25 |
| 7 | Value of action G | 8 |
| 8 | Value of action H | 40 |
| 9 | Points per GBP | 200 |

Of course, one could add more parameters, more complex outcome structures, additional rules and more. One should not forget though, that our aim in this study was to compare different interventions that themselves add complexity to the task. We demonstrate the consequences of a seemingly simple addition below, but the general observation will hold for many possible additions. On their own they seem to add little burden to the game rules, but in addition to the basic game and interventions, these additions can easily increase the amount of confused responses, increase variance in responses based on participants disconnecting from the game and reduce the power of studies with little gain in exchange.

**Example: Making the payoff function stochastic.** Our payoff function is quite simple, it is all or nothing depending on the color state. If we had added randomness to the final calculation in the case of infection/purple color, then at any point in the game, participants would have had to calculate expected values for their current point score for both colors and build the difference to have an idea of the loss potential in the case of color switching. For example, we could have differentiated between mild, moderate, and severe consequences of color change, resulting in different fixed or relative reductions in bonus money. For example we could have told participants that these consequences occur with the probabilities $p_{mild} = 0.5$, $p_{moderate} = 0.35$, and $p_{severe} = 0.15$. The consequences could have been reductions of 25%, 50%, and 90% of the points. This would have required us to reduce the number of points paid out for H actions accordingly. But how does a participant interpret a score of 150 in this scenario and how many participants would be able to compute the expected value in the case of infection? The correct solution is:

$$\begin{aligned} E(score) =& 0.5 \cdot (100\% - 25\%) \cdot 150 + 0.35 \cdot (100\% - 50\%) \cdot 150 + 0.15 \cdot (100\% - 90\%) \cdot 150 \\ =& 56.25 + 26.25 + 15 = 97.5 \end{aligned} \tag{1}$$

The expected value could be calculated for participants in each round (similar to the conversion into pounds we use). But the expected value is not synonymous with the outcome distribution. Purple participants with a score of 150 would earn a payoff described by the gamble $(0.5, 112.5; 0.35, 75; 0.1, 15)$. Risk-averse participants would certainly choose

the expected value over the gamble. To make matters more complex, this calculation is necessary on top of the necessity to estimate one's chances to change from blue to purple. Based on our teaching experience, we would not assume that all university students would be able to make the necessary calculations, and we believe that the added realism of differential outcomes is outweighed by the costs of many participants losing their ability to interpret the possible consequences of game outcomes.

**Attention and Comprehension checks.** The attention checks and technical checks in the beginning addressed the problem of participants trying to fake their physical location, and checked that participants were willing and able to read questions. This should not be taken as self-evident, as multiple failed attempts to answer simple questions (such as "which of these are names of animals?") demonstrate.

One should also keep platform dynamics in mind. During the Covid crisis, crowdsourcing platforms saw an influx of new participants to crowdsourced platforms (Peyton et al., 2021). Our studies were not affected by the TikTok problem observed on Prolific in July 2021[7].

We took care in our game to include every relevant element in the game in comprehension check questions. While we allowed for multiple attempts, we also recorded the number of false entries to measure the degree of comprehension and prevent players who did not read the instructions properly to take part in the game. We included 16 comprehension checks. A small number of participants did not pass the checks, even after passing attention checks. The analysis of the distribution of errors (see **VM2**) shows that while the majority of participants makes a small number of errors (the median is 1 error, the third quartile is 3), there are some players with ten or more errors. Allowing for repeated attempts gives an opportunity to learn and participate and avoids the selection of specific subgroups of participants. The number of errors also shows that the chosen level of complexity is challenging for some participants and that further increases might easily lead to substantial selection effects.

We believe that the number of comprehension checks necessary to cover all game parameters is about the maximum number that can be used without losing frustrated participants or adding too many details that are unlikely to be simultaneously memorized.

**Post-questionnaire.** The qualitative data in the post-questionnaire (see **VM2**) gives a good indication that most participants made sense of the game situation in the intended way. Most participants give plausible estimates of the number of purple players across rounds, their messages to other players make sense and their comments are expected comments given the game. Still, some participants also indicate that they felt challenged by the scenario, calling it a "complicated math problem". Some of the interventions led some players to consider them "confusing", but testing their appropriateness was a central question in this study. Taken as a whole, the responses indicate that the complexity was appropriate.

[7] see https://blog.prolific.co/we-recently-went-viral-on-tiktok-heres-what-we-learned/

# Part II
# Formal analysis of the Transmission Game

## E4 Summary of Theoretical Simulation Results

### E4.1 Degree of cooperation and color switching rates

Choosing action H can be considered a competitive action and choosing action G a cooperative action. Choosing H increases the risk of color change (blue to purple) and thereby the risk of ending the game without bonus payment. At the same time, choosing action H increases a player's point score substantially, without producing any benefit for other players. Choosing action G lowers the risk of color switching both for the player themself and other players.

To explore how the game plays out if different numbers of players choose cooperative and competitive actions, we conducted several sets of simulations (see **SCR** for code and results for all of these simulations). In a first set, we determined a fixed proportion of players who would consistently choose H-actions (AllH) with the remaining players choosing G-actions for every round in the game (AllG). By systematically varying the proportion of AllH-players, we were able to observe how the simulation would play out depending on the amount of cooperation the group.

Figure E7 summarizes the results of 1,000,000 games with 50% AllH and 50% AllG players. Many elements in the game are stochastic in nature, but results across games are centered around a trajectory of small increases of the proportion of purple players in the beginning, accelerated growth in the middle of the game and some degree of slowing down in the final rounds. This dynamic is due to a combination of exponential infection dynamics with a finite population. That is, the number of purple participants starts at 8% and each color switch leads to an acceleration of further color switching until half of the population is purple. Beyond this point, there will be fewer pairings of blue with purple players [8]

[8] If $\pi_p$ is the proportion of purple players, there are $n_d = \pi_p \cdot (1 - \pi_p) \cdot 100^2$ possible pairings with players of different colors, which has a maximum at $\pi = 0.5$, when 2,500 of the full set of potential 4,950 pairings are between participants of different color.

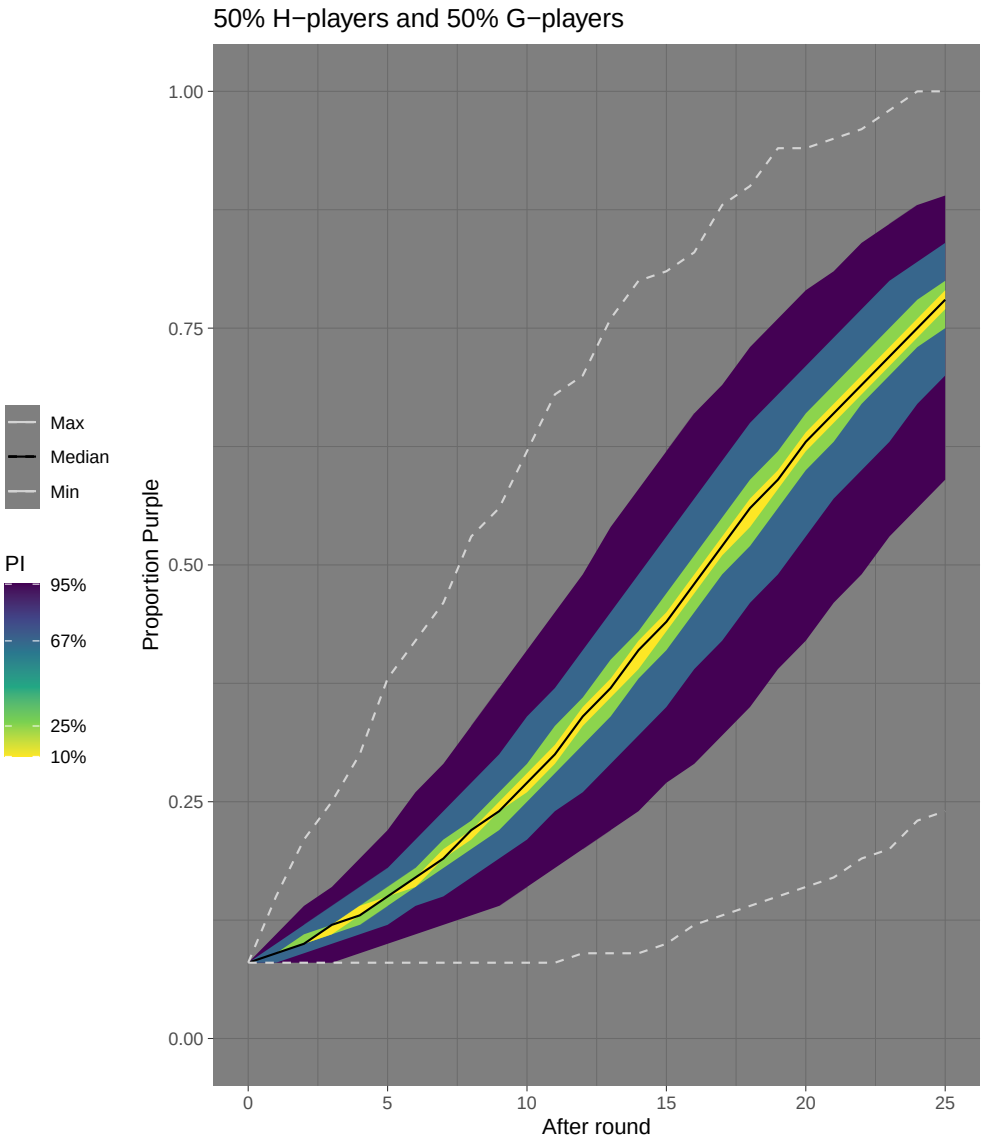


**Figure E7**

*Distribution of observed proportions of purple participants by round across 1,000,000 games with 50% ALLG-players and 50% ALLH-players: lines mark the median, minimum, and maximum results, and areas show the inner 10%, 67% and 95% of results.*

Varying the percentage of AllH players systematically from 0 to 100% in 20 steps and simulating 1,000,000 games for each percentage, reveals a clear pattern in the mean proportion of purple participants across rounds (see Fig. E8). The expected final proportion of purple participants is very sensitive to the number of risk-taking players in the group, starting from well below 0.25 and going up to over 0.95 (less than five players stay blue on average, in the latter case). Color switching can at no point be completely avoided, but it can certainly be reduced by collective cooperation. For lower percentages of AllH players final proportions of purple players change in larger increments as these groups do not face the growth limits incurred when half of the population has already switched colors. In contrast, in the upper range of proportions of AllH players a large majority of the population has already switched to purple in the final rounds of the game.

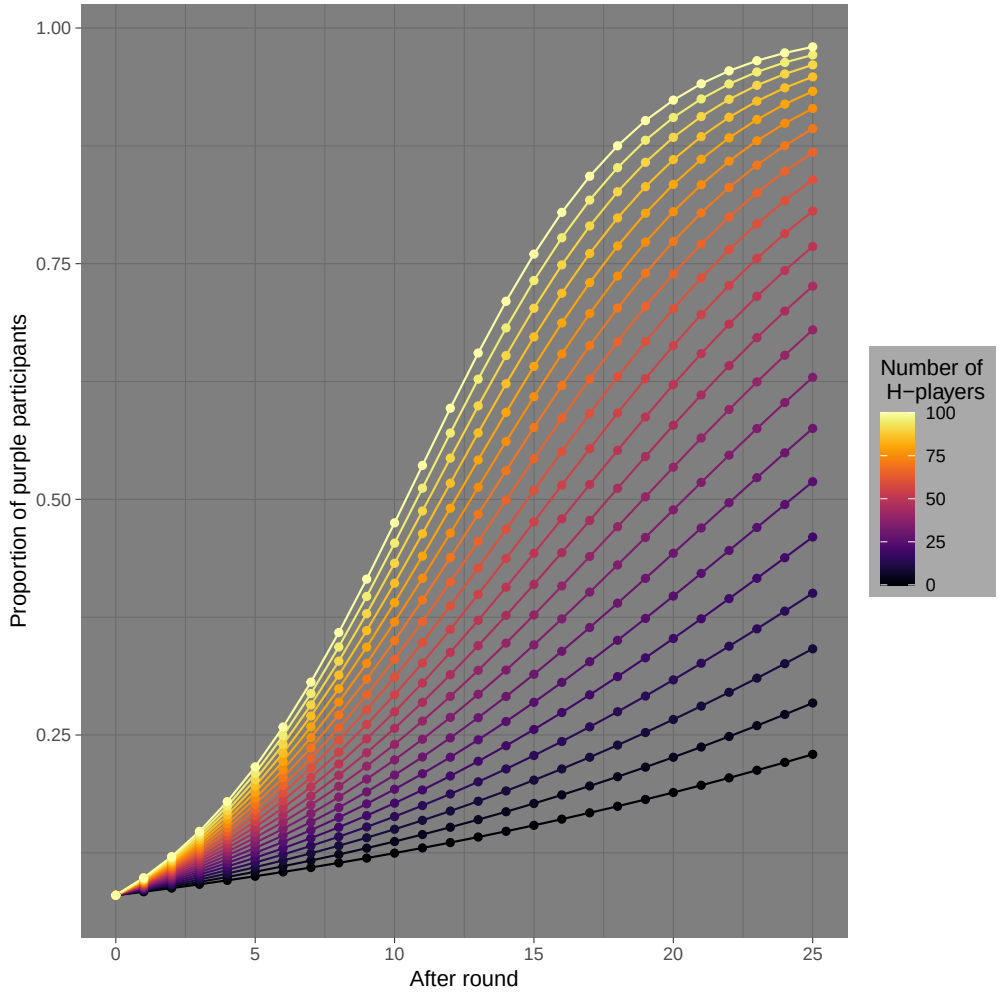


**Figure E8**

*Average proportion of purple participants after each round for each simulated number of H-players: points mark the mean across 1,000,000 games.*

### E4.2 The social dilemma at the heart of the transmission game

Central to the transmission game is a social dilemma. At first glance, there are two consequences of choosing H-actions over G-actions: there is a gain in scored points and an increase in risk of switching to purple. But it is less clear how these two outcomes trade off against each other. AllG-players will always score $8 \cdot 25 = 200$ points if they stay blue and 0 otherwise. AllH-players will always score $40 \cdot 25 = 1,000$ points if they stay blue and 0 otherwise. Obviously, staying blue will depend both on the player's own actions and other players' actions.

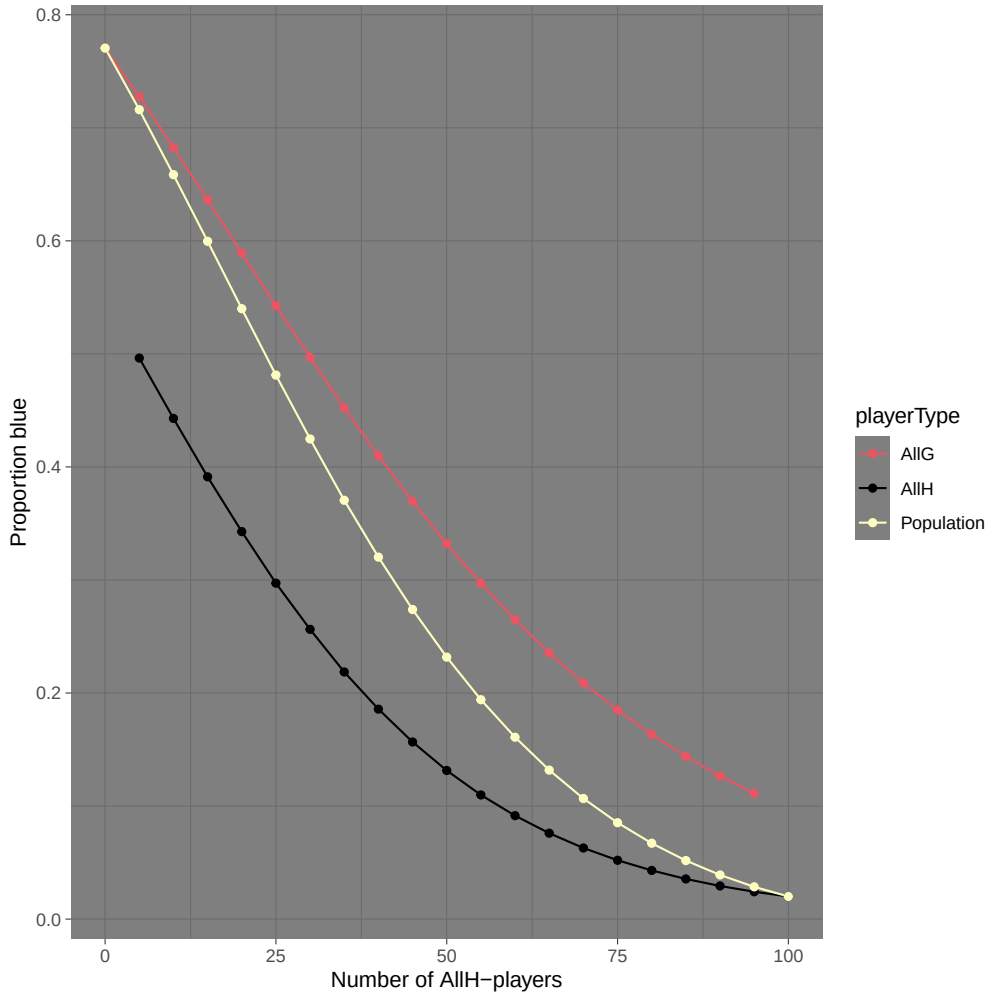


**Figure E9**
*Average rate of staying blue as a function of the number of AllH-players for the full group (population) and the two subgroups of players.*

In a first step, we show that the average survival rates depend on the proportion of AllH-players. We summarize average rates and strategy-specific rates in Figure E9. Irrespective of the proportion of AllH-players, AllG-players stay blue at a higher rate than AllH-players. At the same time, with increasing numbers of AllH-players, the rate of blue

players drops for both AllH- and AllG-players. With only a third of all players being AllH-players, the chance to stay blue even for AllG-players is even lower than the chance of AllH-players to stay blue if there are only a few AllH-players. Unsurprisingly, the average rate of staying blue in the group lies between the two subgroups's rate (with the position depending on the relative group size). In spite of this advantage in terms of survival rates, AllG-players score fewer points than AllH-players at a point rate of 5:1. Figure E10 plots average expected payoffs for the player groups.

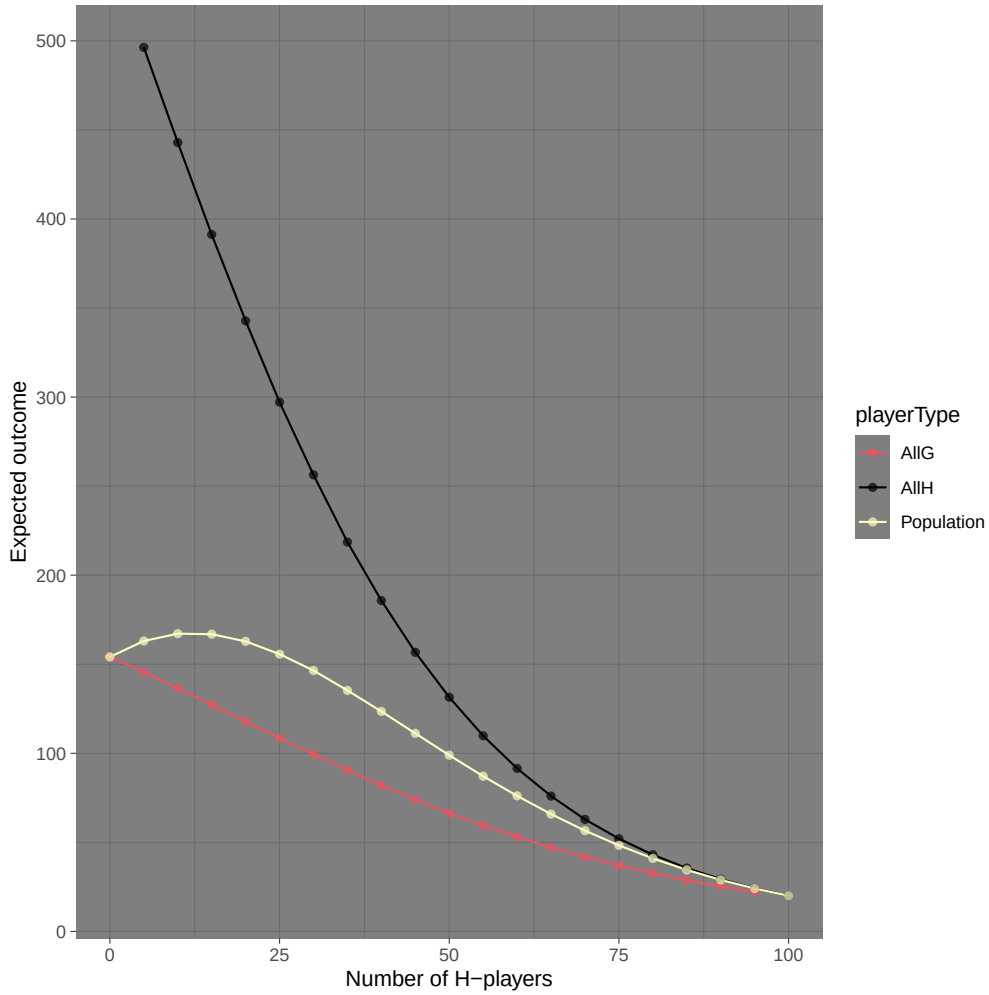


**Figure E10**
*Average expected payoff as a function of the number of AllH-players for the full group (population) and the two subgroups of players.*

Irrespective of the proportion of AllH-players, AllH-players have a higher expected payoff than AllG-players. In other words, the AllH-strategy dominates the AllG-strategy in terms of expected payoffs. Thus, there is always an individual incentive not to cooperate. At the same time, if all players choose the AllH-strategy, even the expected outcomes for AllH-players will be lower than the expected outcomes for AllG-players with perfect cooperation.

### E4.3 Optimal individual strategy

Based on the observations in the previous section, the transmission game can be compared to an $n$-person prisoner's dilemma (Hamburger, 1973), although our game has two specific features: (1) the average payoff does not decrease monotonously with competition, and (2) the relatively higher expected payoff for competition is linked to a higher risk of leaving the game with nothing. An optimal individual strategy thus depends at least on two types of individual preferences: social preferences and risk preference. A player who places a high enough value on average group outcomes or who is sufficiently averse to inequality will not choose the strategy with the highest expected personal outcome. At the same time, a player who is sufficiently averse to risk will not choose the riskiest strategy, even if this player is indifferent to the outcomes of other players. AllG and AllH are not the only available strategies, players can choose a number of H-choices between 0 and 25. Without assumptions about individual preferences any number of H-choices might be potentially optimal for a rational player trying to maximize their expected utility.

### E4.4 Flattening the curve: A collusive optimum

The analysis of average group outcomes further shows that a group playing a universal AllG-strategy will reach the highest rate of staying blue, but it does not ensure the maximum average expected outcome: a limited number of H-actions increase the expected average value at the cost of a small drop in payout rates. The benefits of these choices would disproportionately favor AllH-players. In **SCR** we explore collective strategies that aim to maximize the average expected value, thus achieving a collusive optimum. We demonstrate that both the distribution of H-choices across players and the spacing of H-actions across rounds impact the final outcomes. In particular, an even distribution of H-actions across players ensures that the accumulated points are not lost due to increased risk of switching colors. Earlier H-actions lead to better outcomes than later H-actions. In early rounds, fewer players are purple and the risky interactions lead to fewer color switches than in later rounds. Thus, the best strategy we identified for a group of players that intends to maximize everyone's expected outcome is to start with a number of H-actions and then switch to G-actions. This strategy is defined by choosing a switching point between H and G. We simulated all possible switching points (including AllG and AllH) and played out 1,000,000 games for each of them. The resulting proportions of purple participants by round are shown in Figure E11.

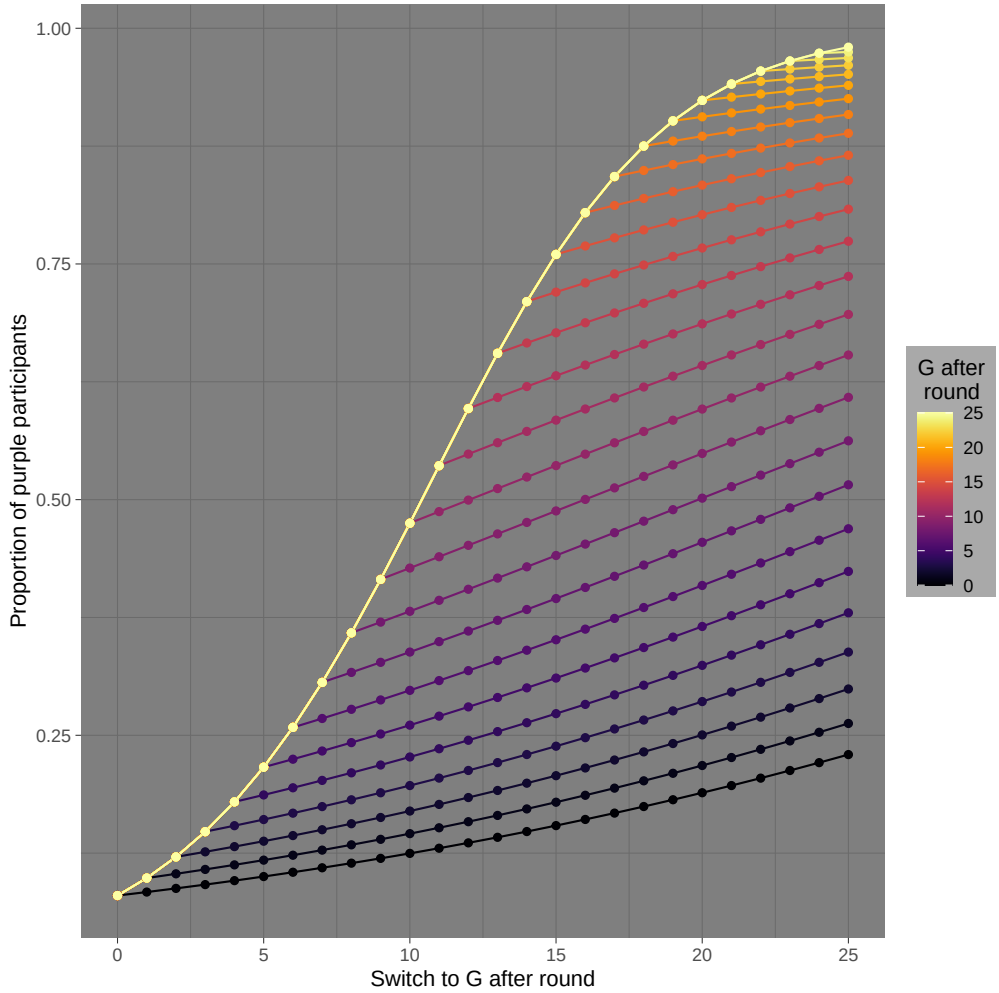


**Figure E11**

*Average proportion of purple participants after each round for each possible collective switching point from H-actions to G-actions: points mark the mean across 1,000,000 games.*

It is clear from the setup that all games have the same expected values before the switching point. The upper line corresponds to the AllH-condition with 100% AllH-players, the lowest line to the AllG-condition with 0% AllH-players. One can call the effect of switching "flattening the curve", as the change of strategy leads to a strong reduction in the slope of the curve. As the extreme strategies correspond to the extremes in the previous simulation, the range of possible final outcomes is the same. To determine the quality of switching points, we computed both the final proportion of blue players (the complement of the proportion of purple players after round 25) and the expected value for each player (all players use the same strategy here). This allows us to identify the best options (see Fig. E12).

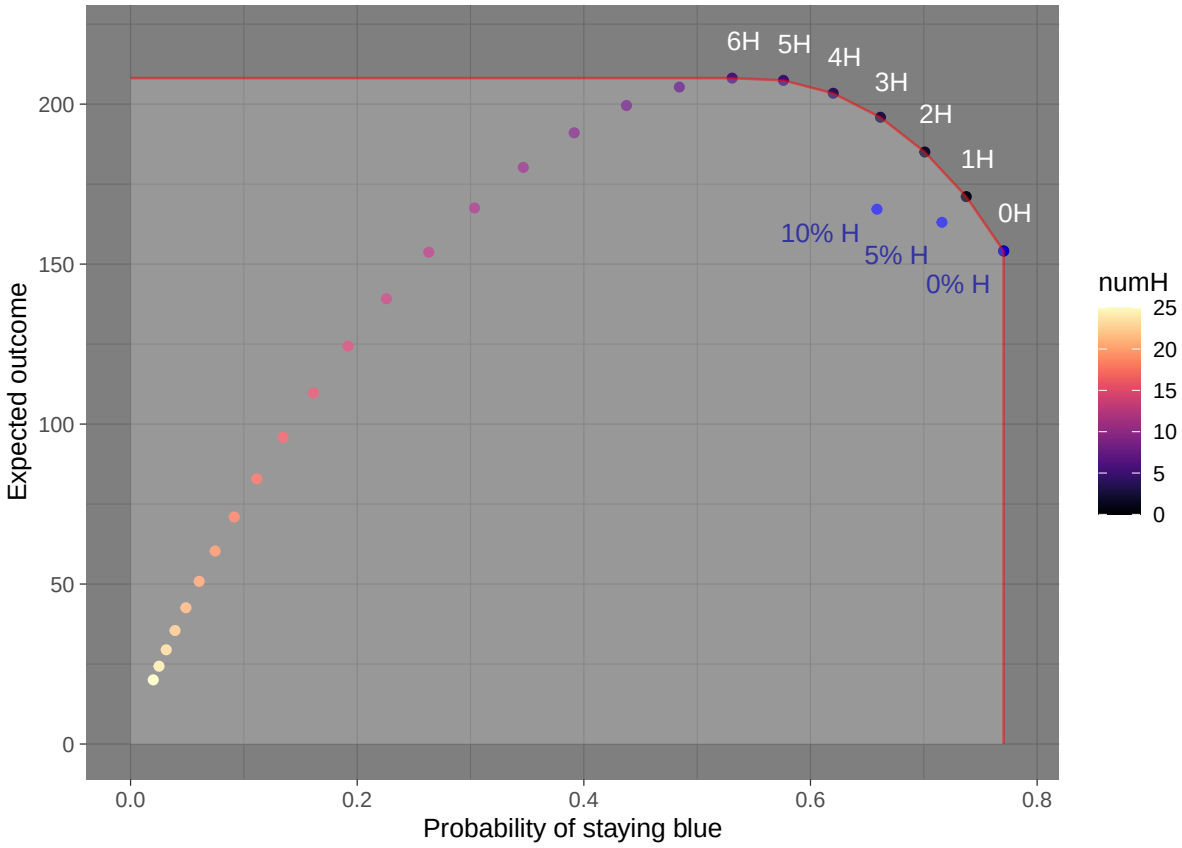


**Figure E12**

*Scatterplot of probability of staying blue (x-axis) and the average expected outcome (y-axis) for all possible switching points: The red line marks the pareto-frontier, pareto-efficient strategies are labelled in white. For comparison, the pareto-efficient strategies for fixed proportions of AllH-players are marked in blue.*

Again, we observed the tradeoff between the probability of receiving a bonus payment and the size of the bonus payment. Depending on the degree of risk-aversion, switching points between 0 and 6 are pareto-efficient: there is no other solution that outperforms these solutions on both dimensions. Only a risk-seeking group of players should even consider switching points above 6. The collusive perspective assumes that the group's aim is to achieve the best result for each individual. The highest expected outcome is achieved with a 6H-19G strategy, but the highest probability of being paid is achieved with a 25G-strategy (also corresponding to a fixed percentage of 0% H-players). The best solutions for fixed proportions above 0 achieve lower expected outcomes and lower probabilities of staying blue than these H-G-strategies. An optimal collusive strategy thus requires each individual player to understand that the temporal order of actions matters and to switch from H to G after a small number of rounds. One can assume that this realization requires some cognitive sophistication.

## E5 Formal analysis of the transmission game

During the review process, reviewers asked to supplement the simulations with a formal analysis of the properties of the transmission game. In particular, there was interest in a more precise analysis of game equilibria. We will discuss the relative merits of simulations and formal analysis and introduce some tools that we developed to derive analytical solutions to replicate the search for a collusive optimum without resorting to stochastic simulation methods. We will then report on a search for individual-level equilibrium for individual behavior and discuss the relevance of this solution for explaining actual game behavior.

### E5.1 Simulations and formal analysis

Monte Carlo simulation methods are established tools in many disciplines and used frequently in biology (DiStefano III, 2015), artificial intelligence (Doukidis & Angelides, 1994), artificial life (Adami, 1998), statistics (Robert et al., 2004), and cognitive psychology (Gigerenzer et al., 1999; Hertwig et al., 2019; Ratcliff & Rouder, 1998). At the same time, some regard simulation results in the absence of formal proof tentative and incomplete. In some areas of economics and theoretical biology, formal methods are the strongly preferred approach.

We believe that both approaches complement each other. Some processes are too complex for analytical solution and would require simplifications for these approaches to work. Simulations offer the advantage of modelling processes in a step-by-step fashion. Simulations with large numbers of repetitions are likely to approach correct results with little error. In the analyses so far, we used 1,000,000 trials to estimate outcomes.

On the negative side, simulation approaches are time-consuming and outright wasteful if fast analytical approaches are available. As stated before though, analytical approaches do not always readily present themselves, and may sometimes even be slower to compute than simulation estimates. Both approaches are prone to errors: programming errors can lead simulations astray, and mathematical errors or formalization mismatches can lead the analytical approach astray. We decided to develop a formal approach to the game to complement and extend our earlier simulations. Seeing convergence between simulations and analysis can also further increase confidence in the findings.

### E5.2 Modelling the transmission game: Uniform behavior

We will describe the components of the analytic solution and use a toy example to demonstrate the type and form of results at each step. In the toy example, we use a population of eight individuals over three rounds with 2 infected individuals before the first round.

#### *E5.2.1 Population states and game results*

The most important decision in approaching this problem is to find a good representation of relevant states in the game. Certain details of interactions during one round (e.g., which specific infected individual was paired with which other individual) are irrelevant for subsequent rounds, while other details, such as infection counts, need to be known. We will demonstrate our choice of representation with the toy example mentioned above.

We consider populations of $n$ individuals, with population infection states in each round $i$ defined as a probability distribution

$$I_i = (p_{i,0}, p_{i,1}, \ldots, p_{i,n}), \tag{2}$$

with $\sum_{j=0}^{n} p_{i,j} = 1$.

The only certain point in the game is the situation before the first round, with a defined number of $k$ intial infections ($p_{0,k} = 1$ and $p_{0,\bar{k}} = 0$. For the toy example (2 infections in a population of 8 players), this translates into

$$I_0 = (p_{0,0} = 0, p_{0,1} = 0, p_{0,2} = 1, p_{0,3} = 0, p_{0,4} = 0, p_{0,5} = 0, p_{0,6} = 0, p_{0,7} = 0, p_{0,8} = 0). \tag{3}$$

.

#### *E5.2.2 Strategies*

Strategies can be described by an action vector across 25 rounds:

$$S = (s_1, s_2, s_3, \ldots, s_{25}). \tag{4}$$

Each $s_i$ can take the values 0 (action G) or 1 (action H). There is a wide range of possible combinations of decisions in the original game. Each player can choose between $2^{25} = 33,554,432$ strategies that reduce to $2^3 = 8$ in the simplified game. Each player can choose their strategy independently, with $2^{2,500}$ unique playing patterns. To reduce the complexity for analysis, we focus first on situations with a single strategy played by all population members. Only in this case is the population state at the beginning of a round sufficient information to determine the state at the end of the round. In all other cases, the probability of distinct strategy types of being infected will likely differ and need to be jointly considered (see the more complex analysis below). In the toy example, there are eight possible strategies: $S_1 = (0,0,0)$, $S_2 = (0,0,1)$, $S_3 = (0,1,0)$, $S_4 = (0,1,1)$, $S_5 = (1,0,0)$, $S_6 = (1,0,1)$, $S_7 = (1,1,0)$, $S_8 = (1,1,1)$.

#### *E5.2.3 Roundwise transmissions*

For a transmission to occur, an infected individual needs to be paired with an uninfected individual. Thus, the number of possible infections depends on the population state at the beginning of the round. For the situation with shared strategies, there is a uniform probability of transmission for all mixed pairs. Therefore, the infection state of the population after the round can be derived from the infection state before the round, the distribution of the number of mixed pairs, and the distribution of infections for a given number of mixed pairs.

**Number of pairings with a specific number of mixed pairs.** The number is a function of population size and the number of infected (and healthy) individuals within this population. The distribution can be derived using combinatoric principles. Given a population of size $n$ ($n \; mod \; 2 = 0$, $n > 0$) with $h$ healthy and $100 - h = i$ infected members, a random pairing of all members can result in 0 to $min(h, i)$ pairings, in which an infected individual is paired with a healthy individual, henceforth called mixed pairings. The number

of (equiprobable) unique pairings of population members can be calculated as the number of possible partitions into groups of 2:

$$T(h,i) = \frac{(h+i)!}{2^{\frac{h+i}{2}} \cdot (\frac{h+i}{2})!} = \frac{n!}{2^{\frac{n}{2}} \cdot (\frac{n}{2})!} \tag{5}$$

It is also possible to determine the number of pairings that produce exactly $k$ mixed pairs. This number is calculated using three parts: the number of possible set of $k$ mixed pairs ($P$), the number of ways to arrange the remaining healthy members into non-mixed pairs ($H_{hty}$) and the number of ways to arrange the remaining infected members into non-mixed pairs ($H_{inf}$). These counts can be derived as:

$$P(h,i,k) = \frac{\frac{h!}{(h-k)!} \cdot \frac{i!}{(i-k)!}}{k!}, \tag{6}$$

$$H_{hty}(h,k) = \frac{(h-k)!}{2^{\frac{h-k}{2}} \cdot (\frac{h-k}{2})!}, \tag{7}$$

and

$$H_{inf}(i,k) = \frac{(i-k)!}{2^{\frac{i-k}{2}} \cdot (\frac{i-k}{2})!}. \tag{8}$$

$P(h,i,k)$ corresponds to the number of unique selections of $k$ elements out of both $i$ and $h$. The denominator corresponds to the number of possible orderings of $k$ pairs, as the order of pairs is irrelevant. There are $h-k$ healthy population members in non-mixed pairs. The enumerator corresponds to the number of orderings of these elements. Both the position within pairs, and the order of pairs are irrelevant, and the denominator corresponds to the product of the number of possibilities for each. $H_{inf}$ is determined in parallel to $H_{hty}$.

The number of unique pairings with $k$ mixed pairs is therefore

$$M(h,i,k) = P(h,i,k) \cdot H_{hty}(h,k) \cdot H_{inf}(i,k) \tag{9}$$

$$= \frac{\frac{h!}{(h-k)!} \cdot \frac{i!}{(i-k)!}}{k!} \cdot \frac{(h-k)!}{2^{\frac{h-k}{2}} \cdot (\frac{h-k}{2})!} \cdot \frac{(i-k)!}{2^{\frac{i-k}{2}} \cdot (\frac{i-k}{2})!} \tag{10}$$

This number is calculated in the Matlab function below (see Listing 1).

Listing 1: Calculating the number of mixed pairs in Matlab

```
1
2  function [ numPairings,percTotal ] =
       findNumCombinationsofMixedPairings( healthy,infected,numMixed)
3  if mod( (healthy+infected),2)>0
4      %odd number of individuals: pairing impossible
5      numPairings=0;
6      percTotal=0;
7    elseif numMixed>min(healthy,infected)
8      % there cannot be more mixed pairings than the minimum number
           of
9      % healthy and infected
10     numPairings=0;
11     percTotal=0;
12   elseif mod( min(healthy,infected)-numMixed,2)>0
13       % the remaining single individ would have to be in a mixed
             pair
14       numPairings=0;
15       percTotal=0;
16   else
17       %There is at least one combination possible
18       numTotalSequences=factorial(healthy+infected)/(2^( (healthy
             +infected)/2)*...
19           factorial((healthy+infected)/2));
20       %number of unique mixed pairs
21       %for each pair: choose one of remaining infected
22       %         choose one of remaining healthy (independently)
23       %   num of healthy: healthy!/(healthy-numPairs)!
24       % num of infected: infected!((infected-numPairs)!
25       numMixedSequences=(factorial(healthy)/(factorial(healthy-
             numMixed))*...
26           (factorial(infected) / factorial(infected-numMixed)));
27       %order of pairs is irrelevant
28       numMixedSequences=numMixedSequences/factorial(numMixed);
29       %number of remaining healthy-healthy pairs (given each
             unique mix):
30       numRemainingHealthy=factorial(healthy-numMixed)/(factorial(
             (healthy-numMixed)/2)*...
31           2^((healthy-numMixed)/2));
32       numRemainingInfected=factorial(infected-numMixed)/(
             factorial( (infected-numMixed)/2)*...
33           2^((infected-numMixed)/2));
34       numMixedSequences=numMixedSequences*numRemainingHealthy*
             numRemainingInfected;
35       numPairings=numMixedSequences;
```

```
        percTotal=numMixedSequences/numTotalSequences;
    end
end
```

**Example for the toy problem.** As an example, consider the question of how many pairings with 0 mixed pairs can be created for the population in the toy problem in the first round (2 infected, 6 healthy). If the infected individuals are named $a$ and $b$, respectively, and the healthy individuals $u$,$v$,$w$,$x$,$y$,$z$, then all pairings can be enumerated:

$$
\begin{array}{lll}
((a,b),(u,v),(w,x),(y,z)) & ((a,b),(u,w),(v,x),(y,z)) & ((a,b),(u,v),(w,x),(y,z)) \\
((a,b),(u,x),(v,w),(y,z)) & ((a,b),(u,z),(v,w),(x,y)) & ((a,b),(u,v),(w,y),(x,z)) \\
((a,b),(u,w),(v,y),(x,z)) & ((a,b),(u,v),(w,y),(x,z)) & ((a,b),(u,x),(v,y),(w,z)) \\
((a,b),(u,z),(v,x),(w,y)) & ((a,b),(u,v),(w,z),(x,y)) & ((a,b),(u,w),(v,z),(x,y)) \\
((a,b),(u,v),(w,z),(x,y)) & ((a,b),(u,x),(v,z),(w,y)) & ((a,b),(u,z),(v,y),(w,x))
\end{array}
$$

All of the pairings must include the pair $(a,b)$, as there are only two infected individuals. Splitting the two individuals up will result in not one but two mixed pairs. In fact, there is no way to find a complete pairing with exactly one mixed pair. The orderings of pairs within the pairing are irrelevant: $((a,b),(u,v),(w,x),(y,z))$ is the same pairing as $((u,v),(a,b),(w,x),(y,z))$, each of the $4! = 24$ orders of pair are equivalent. Similarly, the order of individuals within each pair does not matter and $(a,b)$ is equivalent to $(b,a)$. There are $2^4 = 16$ ways to arrange individuals within pairs and this is independent from the order of pairs. The number can also be computed by using the equation for $M(h,i,k)$:

$$
\begin{aligned}
M(6,2,0) =& \frac{\frac{6!}{(6-0)!}\cdot\frac{2!}{(2-0)!}}{0!}\cdot\frac{(6-0)!}{2^{\frac{6-0}{2}}\cdot(\frac{6-0}{2})!}\cdot\frac{(2-0)!}{2^{\frac{2-0}{2}}\cdot(\frac{2-0}{2})!} \\
=& \frac{1\cdot 1}{1}\cdot\frac{(6)!}{2^3\cdot 3!}\cdot\frac{2}{2\cdot 1} = \frac{6\cdot 5\cdot 4\cdot 3!}{8\cdot 3!} = \frac{120}{8} \\
=& 15
\end{aligned}
$$

There are more orderings that result in two mixed pairs:

$$
\begin{aligned}
M(6,2,2) =& \frac{\frac{6!}{(6-2)!}\cdot\frac{2!}{(2-2)!}}{2!}\cdot\frac{(6-2)!}{2^{\frac{6-2}{2}}\cdot(\frac{6-2}{2})!}\cdot\frac{(2-2)!}{2^{\frac{2-2}{2}}\cdot(\frac{2-2}{2})!} \\
=& \frac{\frac{6!}{4!}\cdot\frac{2!}{0!}}{2!}\cdot\frac{4!}{2^2\cdot 2!}\cdot\frac{0!}{2^0\cdot 0!} = \frac{30\cdot 2}{2}\cdot\frac{24}{4\cdot 2}\cdot\frac{1}{1\cdot 1} \\
=& 30\cdot 3\cdot 1 = 90
\end{aligned}
$$

It is evident that these are all possible orderings, as

$$
\begin{aligned}
T(6,2) =& \frac{(6+2)!}{2^{\frac{6+2}{2}}\cdot(\frac{6+2}{2})!} = \frac{8!}{2^4\cdot 4!} \\
=& \frac{8\cdot 7\cdot 6\cdot 5}{16} = 105
\end{aligned}
$$

**The probability distribution of the number of mixed pairs.** The probability that a randomly generated pairing creates exactly $k$ mixed pairs is simply the ratio of the number of unique orderings fulfilling the criterion and the total number of unique orderings:

$$p(m = k; h, i) = \frac{M(h, i, k)}{T(h, i)}, \tag{11}$$

or

$$p(m = k; h, i) = \frac{\frac{\frac{h!}{(h-k)!} \cdot \frac{i!}{(i-k)!}}{k!} \cdot \frac{(h-k)!}{2^{\frac{h-k}{2}} \cdot (\frac{h-k}{2})!} \cdot \frac{(i-k)!}{2^{\frac{i-k}{2}} \cdot (\frac{i-k}{2})!}}{\frac{(h+i)!}{2^{\frac{h+i}{2}} \cdot (\frac{h+i}{2})!}} \tag{12}$$

The number of mixed pairs resulting from a random orderings is thus defined as:

$$D(h, i) = (p(m = 0; h, i), p(m = 1; h, i), \ldots, p(m = min(h, i); h, i)). \tag{13}$$

In Matlab, this can be done by repeatedly applying the function in Listing 1, as implemented in the function below (see Listing 2).

Listing 2: Calculating the distribution of the number of mixed pairs in Matlab

```
function [ probDistMixedPairs ] = findProbDistMixedPairs(
    healthy,infected,maxPairs )
 probDistMixedPairs=zeros(maxPairs+1,1);
 for numPairs=0:maxPairs
     [~,probDistMixedPairs(numPairs+1,1)]=...
         findNumCombinationsofMixedPairings(healthy,infected,
             numPairs);
 end
end
```

**Distribution for the toy problem.** All components were already computed above. The distribution for the toy problem is:

$$D(6, 2) = (p(m = 0; 6, 2), p(m = 1; 6, 2), p(m = 2; 6, 2)) = (\frac{15}{105}, 0, \frac{90}{105}) = (\frac{1}{7}, 0, \frac{6}{7}) \tag{14}$$

For this simple case, an alternative way of deriving this distribution would be to consider the chance that individual a is paired with individual b: As there are seven other individuals, the chance is $1/7$. Given that there can only be 0 or two mixed pairs, the probability of seeing two pairs must be $1 - 1/7 = 6/7$.

**Distribution of infections given the number of pairs.** Each mixed pair is associated with an opportunity for the infected person to infect the healthy person. These opportunities are independent of each other. We restricted the analysis to games in which every player chooses the same strategy. Therefore the event occurs with the same probability across pairs, namely: $p = 0.05 + 0.2 \cdot s_i$ in round $i$ (0.05 if players choose action G, and 0.25 if players choose action H).

In this case, the number of infections in the round is binomially distributed. The probability to observe $k$ infections in round $i$ is:

$$p(n_{inf} = k) = f(k, n_{mixed}, 0.05+0.2\cdot s_i) = \binom{n_{mixed}}{k} \cdot (0.05+0.2\cdot s_i)^k \cdot (0.05+0.2\cdot s_i)^{n_{mixed}-k} \tag{15}$$

The implementation in Matlab of calculating this probability given the number of mixed pairs and the infection risk is straightforward (see Listing 3).

Listing 3: Calculating the probability of a given number of infections in a round in Matlab

```
function [ infectionDistribution ] = calcInfectionDistribution(
    probDistMixedPairs,infectionRisk,maxPairs )
infectionDistribution=zeros(maxPairs+1,1);
for ci=1:maxPairs+1
   checkNumber=ci-1;
   if probDistMixedPairs(ci)>0
      infectCount=0:checkNumber;
      outProb=binopdf(infectCount,checkNumber,infectionRisk)';
      infectionDistribution(1:checkNumber+1,1)=...
                  infectionDistribution(1:checkNumber+1,1)+...
                        probDistMixedPairs(ci,1)*outProb;
   end
  end
end
```

**Example.** If there are two mixed pairs in the toy example and all players choose action H in a round, then the number of infections follows the $B(k, 2, 0.25)$ distribution. Specifically:

$$f(0, 2, 0.25) = \binom{2}{0} \cdot (0.25)^0 \cdot (0.75)^{2-0} = \frac{9}{16}, \tag{16}$$

$$f(1, 2, 0.25) = \binom{2}{1} \cdot (0.25)^1 \cdot (0.75)^{2-1} = \frac{3}{8}, \tag{17}$$

and

$$f(2, 2, 0.25) = \binom{2}{2} \cdot (0.25)^2 \cdot (0.75)^{2-2} = \frac{1}{16}. \tag{18}$$

#### *E5.2.4 Putting the parts together*

At this point, all components are in place to be put together to model the full game. We can make repeated use of the chain rule: $P(A \cap B) = P(A|B) \cdot P(B)$ and the law of total probability: $P(A) = \sum_i P(A \cap B_i)$. We know:

1. the probability distribution of population states $I$ before the round,
2. the probability distribution of the number of mixed pairs $M$ conditional on the population state, and

3. the probability distribution of the number of infections $p(n_{inf})$ conditional on the number of mixed pairs.

For round $r$ we can determine the distribution of population states after the round ($I_{r+1}$) using:

$$p_{r+1,i^*} = \sum_{k \leq i^*}^{n} \sum_{m=i^*-k}^{min(n-k,k)} p_{r,k} \cdot p(m, h = n-k, i = k) \cdot f(n_{inf} = i^* - k, m, 0.05 + s_{r+1} \cdot 0.2) \tag{19}$$

There are a few shortcuts in this definition, as the number of infections cannot be lower after a round than before. Thus only population states are relevant to consider that do not have a higher number of infections than the target number. The number of mixed pairs is limited by the minimum of infected and healthy individuals. The Matlab implementation (see Listing 4) uses this and similar shortcuts, but essentially implements this equation for each round. As the population state has sufficient information to calculate results, each round can be implemented by just inheriting the population state from the previous round. The algorithm creates the distribution of infection numbers for each round of the game.

Listing 4: Calculating the development of population states across rounds in Matlab

```
tic

maxPairs=50;
numIndivs=100;
numRounds=25;

probDist=zeros(100+1,numRounds+1);
% probdist 1: 0
% probsdit 25: 24 infected

%Round0: 8 infected
probDist(8+1,1)=1;

%Just an example for a tested strategy:
roundH=[1;1;1;1;1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0];
%probDist: Devlopment per round:

for ri=1:numRounds
  newProbDist=zeros(numIndivs+1,1);

  for distCond = 1:numIndivs+1

      if probDist(distCond,ri)>0
          %determine number of mixed pairs for pop with distCond
             -1 infected

```

```
            numInfected=distCond-1;
            numHealthy=100-numInfected;

            distMixedPairs=findProbDistMixedPairs(numHealthy,
                numInfected,...
                 maxPairs);

            infectionRisk=0.05+roundH(ri,1)*0.2;
            infectDist=calcInfectionDistribution(distMixedPairs,
                infectionRisk,...
                  maxPairs);
            maxIncrease=min(distCond-1+maxPairs,numIndivs)-(
                distCond-1);

            newProbDist(distCond:distCond+maxIncrease)=...
                 newProbDist(distCond:distCond+maxIncrease)+...
                     probDist(distCond,ri)*infectDist(1:maxIncrease
                         +1,1);
        end
    end
         probDist  (1:numIndivs+1,ri+1)= newProbDist;
end

 toc
 meanDist=[0:100]*probDist
 finalDist=probDist(1:numIndivs,numRounds+1);
```

The final part of the code allows to calculate the mean infection rate:

$$\pi_{infection} = \frac{1}{100} \cdot \sum_{i=0}^{100} p_{25,i} \cdot i \tag{20}$$

Likewise, the expected score for each population member is calculated by counting the number of H-choices and multiplying final scores with the survival rate (the probability of not becoming infected, i.e., $1 - \pi_{infection}$):

$$E(O) = (1 - \pi_{infection}) \cdot (200 + 32 \cdot \sum_{i=1}^{25} s_i) \tag{21}$$

**Example.** Applying the (slightly modified) routines to the toy example, one can observe the distribution of population states across the three rounds. Here, we assume a strategy of choosing H in the first two rounds and G in the last round. The resulting development of game states across rounds is shown in Table E3.

The expected average score in this game would be

$$E(O) = (1 - \pi_{infection}) \cdot (24 + 32 \cdot \sum_{i=1}^{3} s_i) = 0.625 \cdot 88 = 55.009 \tag{22}$$

**Table E3**
*Progression of population states across three rounds in the toy example when all players choose the strategy* $S = (1, 1, 0)$.

| | Round | | | |
|---|---|---|---|---|
| Pop. state | 0 | 1 | 2 | 3 |
| 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 |
| 2 | 1 | 0.625 | 0.391 | 0.358 |
| 3 | 0 | 0.321 | 0.382 | 0.374 |
| 4 | 0 | 0.054 | 0.175 | 0.194 |
| 5 | 0 | 0 | 0.045 | 0.060 |
| 6 | 0 | 0 | 0.008 | 0.012 |
| 7 | 0 | 0 | 0.001 | 0.001 |
| 8 | 0 | 0 | <0.001 | 0.001 |
| $\pi_{infection}$ | 0.250 | 0.304 | 0.362 | 0.375 |

## E5.3 Modelling the transmission game: Non-uniform behavior with unilateral deviation

One limitation of the analytical approach presented so far is the assumption of uniform behavior. While it would be ideal to model 100 independently varied strategies, we presently do not believe that is easily done. We will show some steps towards this goal in the following section. Here, we again focus on a special case that is useful for a range of analyses: We model a group of 99 individuals with a uniform strategy in contrast to a single individual that can make different choices. This framework allows to search for strategies that make unilateral deviation unattractive (or, more specifically, Nash equilibria). In this section we will present the required changes to the previous code for this approach to work. We will call the deviating individual 'focus individual' and the remaining individuals 'group'.

### *E5.3.1 Population states and game results*

One essential change from the previous approach is the need to represent game states differently. It is not enough to add an infection probability for the focus individual, as the distribution of group states is not independent of this probability. A solution that allows for this dependency but keeps the necessary book-keeping to a minimum is to extend population states to the following form.

We consider populations of $n$ individuals consisting of one focus individual and 99 group individuals. Population infection states in each round $i$ are defined as

$$I_i = ((p_{i,0,+}, p_{i,0,-}), (p_{i,1,+}, p_{i,1,-}), \ldots, (p_{i,n-1,+}, p_{i,n-1,-})), \tag{23}$$

with $\sum_{j=0}^{n-1} p_{i,j,+} + \sum_{j=0}^{n-1} p_{i,j,-} = 1$.

In this notation $p_{i,k,+}$ represents the probability that in round $i$, $k$ members of the group are infected and the focus individual is infected, as well. In contrast, $p_{i,k,-}$ represents the probability that in round $i$, $k$ members of the group are infected and the focus individual

is not infected. One should note that the probability of the focus individual being infected in a round $p*_{i,+}$ is only implicitly coded, as

$$p*_{i,+} = \sum_{j=0}^{n-1} p_{i,j,+}. \tag{24}$$

This notation also implies a change to the initial population state with $k_{inf}$ individuals, with

$$p_{0,i,+} = \begin{cases} 0, i \neq k_{inf} - 1 \\ \frac{k_{inf}}{n}, i = k_{inf} - 1, \end{cases} \tag{25}$$

and

$$p_{0,i,-} = \begin{cases} 0, i \neq k_{inf} \\ \frac{n-k_{inf}}{n}, i = k_{inf}. \end{cases} \tag{26}$$

Both states with positive probabilities correspond to the same number of infected individuals ($k_{inf}$), but in the first case, the focus individual is one of them, while in the second case the focus individual is not infected. As the group of $k_{inf}$ individuals is selected randomly out of the $n$ individuals in the population, there is probability of $p = \frac{k_{inf}}{n}$ to include the focus individual.

**Example.** We will continue to use the toy example to illustrate the use of the notation for special cases. For the toy example (2 infections in a population of 8 players) with a focus individual, the initial population state is:

$$\begin{aligned} I_0 = ((p_{0,0,+} = 0, p_{0,0,-} = 0), (p_{0,1,+} = \frac{1}{4}, p_{0,1,-} = 0), (p_{0,2,+} = 0, p_{0,2,-} = \frac{3}{4}), \\ (p_{0,3,+} = 0, p_{0,3,-} = 0), (p_{0,4,+} = 0, p_{0,4,-} = 0), (p_{0,5,+} = 0, p_{0,5,-} = 0), \\ (p_{0,6,+} = 0, p_{0,6,-} = 0), (p_{0,7,+} = 0, p_{0,7,-} = 0), (p_{0,8,+} = 0, p_{0,8,-} = 0)). \end{aligned} \tag{27}$$

.

### *E5.3.2 Strategies and transmission risk*

Strategies can still be described by an action vector across 25 rounds, but now two strategies need to be considered, one for the focus individual and one for the group:

$$S_f = (s_{f,1}, s_{f,2}, s_{f,3}, \ldots, s_{f,25}). \tag{28}$$

and

$$S_g = (s_{g,1}, s_{g,2}, s_{g,3}, \ldots, s_{g,25}). \tag{29}$$

Each $s_{f,i}$ and $s_{g,i}$ can take the values 0 (action G) or 1 (action H). Each strategy vector has $2^{25} = 33,554,432$ possible action combinations that can be chosen independently ($2^{50}$ unique combinations). The probability of infection in pairs is no longer the same across all mixed pairs, as it depends on whether the focus individual is part of that pair or not. It does not matter for the probability whether the focus individual is healthy or infected (it does matter for the potential to become infected, of course):

$$p_{inf,i} = \begin{cases} 0, & \text{both healthy or both infected} \\ 0.05 + 0.2 \cdot s_{g,1}, & \text{mixed pair without focus individual} \\ 0.05 + 0.1 \cdot s_{g,1} + 0.1 \cdot s_{f,1}, & \text{mixed pair with focus individual} \end{cases} \tag{30}$$

#### *E5.3.3 Roundwise transmissions*

The analysis of roundwise transmissions follows the same principles as before with the added complication that two different strategies and two different parts of the population need to be considered.

**Number of pairings with a specific number of mixed pairs.** The number of mixed pairings stays unchanged. At the same time, it becomes necessary to separate cases in which the focus individual is involved from those in which the individual is not involved. We therefore differentiate between the condition in which the focus individual is infected and the condition in which the individual is not infected and calculated the probability of involvement in both cases. If there are $h$ healthy and $i$ infected individuals in total, then the probability of involvement is given by the ratio of mixed pairs and the respective group size:

$$p_{inv} = \begin{cases} \frac{n_{mixed}}{h}, \text{focus healthy} \\ \frac{n_{mixed}}{i}, \text{focus infected} \end{cases} \tag{31}$$

The changes made to the Matlab function are shown below (see Listing 5).

Listing 5: Changes made to the function for calculating the number of mixed pairings

```
function [ numPairings,percTotal,probFocusInMixedPair ] =
    findNumCombinationsofMixedPairingsEQ( healthy,infected,
    numMixed, focusInfected)

[... THE OMITTED PART STAYS THE SAME ...]

        if (focusInfected==1)
            %an already infected focus cannot change again, but any
                pair the
            %focus is in is in danger of flipping a healthy group
                member
            %the chance of infection for the group still changes
            probFocusInMixedPair=numMixed/infected;
        else
            %what if the focus is healthy?
            %probability of healthy focus in one of the pairs
            % draw x out of healthy: chance focus is in x: numPairs
                /numhealthy
            probFocusInMixedPair=numMixed/healthy;
        end
```

```
16
17     end
18 end
```

**The probability distribution of the number of mixed pairs.** The modified population representation also leads to change in the routine collecting these distributions. The distribution of mixed pairs is calculated twice for a given number of healthy and infected individuals: once under the assumption that the focus individual is healthy and added to the number of healthy players, and another time under the assumption that the focus individual is infected and added to the number of infected players. Both calls to the function shown before generate distributions and probabilities for the focus individual to be in a mixed pair that are stored in two-dimensional arrays (see Listing 6).

Listing 6: Calculating the distribution of the number of mixed pairs in Matlab in the split-group analysis

```
1  function [ probDistMixedPairs,probFocusInPair ] =
       findProbDistMixedPairsEQ( healthy,...
2                                    infected,maxPairs )
3
4    probDistMixedPairs=zeros(maxPairs+1,2);
5    probFocusInPair=zeros(maxPairs+1,2);
6
7    for numPairs=0:maxPairs
8
9      [~,percTotalInfected, probFocusInMixedPairInfected]=...
10               findNumCombinationsofMixedPairingsEQ(healthy,
                     infected+1,...
11                    numPairs,1);
12
13     [~,percTotalHealthy, probFocusInMixedPairHealthy]=...
14               findNumCombinationsofMixedPairingsEQ(healthy+1,
                     infected,...
15               numPairs,0);
16
17     probDistMixedPairs(numPairs+1,1)= percTotalHealthy;
18     probDistMixedPairs(numPairs+1,2)= percTotalInfected;
19     probFocusInPair(numPairs+1,1)= probFocusInMixedPairHealthy;
20     probFocusInPair(numPairs+1,2)= probFocusInMixedPairInfected
           ;
21
22   end
23 end
```

**Distribution for the toy problem.** We computed the distribution for the toy problem above as:

$$D(6,2) = (p(m=0;6,2), p(m=1;6,2), p(m=2;6,2)) = (\frac{15}{105}, 0, \frac{90}{105}) = (\frac{1}{7}, 0, \frac{6}{7}) \quad (32)$$

This distribution is independent of actions chosen. What does this distribution look like when introducing the group-focus split? It does not exist, as the focus individual is not part of the group-split, but the numbers reoccur in the elements $p(m = k; 6, 2, +)$ when considering $D'(6, 1)$:

$$\begin{aligned} D'(6,1) =&((p(m=0;6,2,+),p(m=0;7,1,-)),(p(m=1;6,2,+),p(m=1;7,1,-)),\\ &(p(m=2;6,+),p(m=0;7,1,-))) \\ =&((\frac{1}{7},0),(0,1),(\frac{6}{7},0)) \end{aligned} \tag{33}$$

Note that there will always be exactly one mixed pair if there is exactly one infected player, therefore $p(m = 1; 7, 1, -) = 1$. The probabilities for the assumption that the focus individual is infected add up to one as do the probabilities for the assumption that the focus individual is healthy, therefore:

$$\sum_{j=0}^{min(h+1,i+1)} p(m = j; h, i + 1, +) = 1 \tag{34}$$

and

$$\sum_{j=0}^{min(h+1,i+1)} p(m = j; h + 1, i, -) = 1. \tag{35}$$

The probability for the focus individual to be included in the mixed pair(s) varies with health status and the number of mixed pairs:

$$\begin{aligned} P'_{inc}(6,1) =&((p_{inc}(m=0;6,2,+),p_{inc}(m=0;7,1,-)),\\ &(p_{inc}(m=1;6,2,+),p_{inc}(m=1;7,1,-)),\\ &(p_{inc}(m=2;6,+),p_{inc}(m=0;7,1,-))) \\ =&((0,0),(0,\frac{1}{7}),(\frac{1}{3},0)) \end{aligned} \tag{36}$$

In contrast to the probabilities above, these probabilities generally do not add up to 1.

**Distribution of infections given the number of pairs.** Calculating the number of infections for each given number of mixed pairs is more complex in this simulation variant, as both the health status of the focus individual and its inclusion in mixed pairs needs to be considered. Infections within the group are still binomially distributed, but this only gives part of the answer. The new routine goes through a case differentiation that might explain the need for the previously calculated inclusion probabilities:

1. The focus individual is healthy and included within the mixed pairs.
2. The focus individual is healthy and not included within the mixed pairs.
3. The focus individual is infected and included within the mixed pairs.

4. The focus individual is infected and not included within the mixed pairs.

In a similar way, the probability distribution of infections needs to differentiate between cases, in which the focus individual is infected and those in which it is not. We chose a similar solution as for population states with a two-dimensional outcome distribution. The routine is also programmed to generate two outcome distributions at once, one for the assumption that the focus individual is healthy (whether included or not), and one for the assumption that the focus individual is infected (in this case, the focus individual cannot leave the round as healthy individual). Each of the two parts consists of values that add up to one.

The case differentiation requires the following changes to the previous routine: If the focus individual is included, then the number of group-group pairings is reduced by 1, and the binomial distribution of group-group infections involves a decremented $n$, or $B(k, n_{mixed} - 1, p_{inf} = 0.05 + 0.2 \cdot s_{g,1})$. If the focus individual is healthy, it cannot infect additional group members, but it can become infected through interaction with an infected group member. If the focus individual is infected, it may or may not increase the number of group infections, and the distribution of group infections has to take this chance into account (the probability of an additional infection is given as $p_{inf} = 0.05 + 0.1 \cdot s_{g,1} + 0.1 \cdot s_{f,1}$). An infected focus individual will stay infected with certainty.

As the previous routine generated the probabilities for the inclusion of the focus individual, all components are available to calculate these two conditional distributions given the conditional distribution of mixed pairs and the conditional probabilities of inclusion of the focus individual.

The added case distinctions make for a more involved algorithm that is reproduced in Listing 7.

Listing 7: Calculation the distribution of infections with group-focus split.

```
function [ infectionDistribution ] =
    calcInfectionDistributionEQv2( probDistMixedPairs,
    probFocusInPair, infectionRiskGroup,infectionRiskMix,maxPairs
    )
    %Structure of variables:
    % probDistMixedPairs: probability distribution of number of
       mixed pairs
    % separate for: focus is healthy (1) and for focus is
       infected (2)
    %probFocusInPair: probability that focus is in one of the
       mixed pairs
    %   separate for: focus is healthy (1)
    %    probDistMixedPairs ( numPairs+1,1)= percTotalHealthy;
    %    probDistMixedPairs ( numPairs+1,2)= percTotalInfected;
    %    probFocusInPair ( numPairs+1,1)=
       probFocusInMixedPairHealthy;
    %    probFocusInPair ( numPairs+1,2)=
       probFocusInMixedPairInfected;

```

```
12      %infectionDistribution: probability that number of new
            infections (IN GROUP)
13      % occurs, from 1 to max number of possible infection events (
            maxPairs)
14      %%second dimension 1: focus healthy (and is not infected)
15      %                  2: focus healthy (and is infected in this
            round)
16      %                  3: focus infected (and is not infected
            after the round; p always 0)
17      %                  4: focus infected (and is also infected
            after the round
18      %                 p=1)
19
20  infectionDistribution=zeros(maxPairs+1,4);
21  for ci=1:maxPairs+1
22    checkNumber=ci-1; %number of events tested (0:maxPairs)
23
24    % IF Focus infected
25    if probDistMixedPairs(ci,2)>0 %skip if impossible
26      infectCount=0:checkNumber;
27      %if focus is not in pairs: distribution of hits in group
28      %this is "withoutFocus" (that is infected)
29      outProbWithoutFocusInfected=binopdf(infectCount,checkNumber,
            infectionRiskGroup)';
30      %the following is for the case that the focus IS included in
            mixed pairs
31      outProbIncludingFocusInfected=zeros(checkNumber+1,1);
32      if checkNumber==0 %there are no mixed pairs
33        outProbIncludingFocusInfected(1,1)=1;   %no new infections
34      else
35        groupOnlyProbsIfIncluded=binopdf(infectCount,checkNumber-1,
              infectionRiskGroup)';
36        %focus individual can add another infection or not
37        outProbIncludingFocusInfected(1,1)=groupOnlyProbsIfIncluded
              (1,1)*...
38            (1-infectionRiskMix);
39        outProbIncludingFocusInfected(2:checkNumber+1,1)=...
40        groupOnlyProbsIfIncluded(2:checkNumber+1,1)*...
41              (1-infectionRiskMix)+...
42              groupOnlyProbsIfIncluded(1:checkNumber,1)*...
43              infectionRiskMix;
44      end
45      % focus infected: focus will always be infected after the
            round
46      % (therefore row 3 stays at 0)
```

```
47      infectionDistribution(1:checkNumber+1,4)=...
48        infectionDistribution(1:checkNumber+1,4)+...
49        probDistMixedPairs(ci,2)*(...
50        probFocusInPair(ci,2)*outProbIncludingFocusInfected(1:
              checkNumber+1,1)+...
51        (1-probFocusInPair(ci,2))*outProbWithoutFocusInfected(1:
              checkNumber+1,1)...
52        );
53    end
54
55    % IF Focus healthy: needs to consider #group infection WITH or
56    % WITHOUT additional focus infection
57    if probDistMixedPairs(ci,1)>0 %skip if impossible
58      infectCount=0:checkNumber;
59      %if focus is not in pairs: distribution of hits in group
60      outProbWithoutFocusHealthy=binopdf(infectCount,checkNumber,
            infectionRiskGroup)';
61      outProbIncludingFocusHealthy=zeros(checkNumber+1,1);
62      if checkNumber==0
63        outProbIncludingFocusHealthy(1,1)=1;
64        %hit numbers can be achieved  via two paths:
65        %aincluding an infection cauxed by focus or not
66        % but with a healthy focus they cannot impact the group
67        %focusInfectionProbIfIncludedHealthy=0;
68      else
69        %if included, one pairing cannot affect the group (only the
70        %focus)
71        outProbIncludingFocusHealthy=binopdf(infectCount,
              checkNumber-1,infectionRiskGroup)';
72      end
73      % focus healthy: Focus is NOT in Pair: all in line 1 (
            infection is
74      % impossible
75      infectionDistribution(1:checkNumber+1,1)=...
76        infectionDistribution(1:checkNumber+1,1)+...
77        probDistMixedPairs(ci,1)*(...
78        (1-probFocusInPair(ci,1))*outProbWithoutFocusHealthy(1:
              checkNumber+1,1)...
79        );
80      % focus healthy: Focus is IN Pair: in line 1 (if not infected
            )
81      % in line 2 if infected; infection probability is (given that
            in pair)
82      % the mix infection prob
83      %infection
```

```
    infectionDistribution(1:checkNumber+1,2)=...
      infectionDistribution(1:checkNumber+1,2)+...
      probDistMixedPairs(ci,1)*(...
      probFocusInPair(ci,1)*infectionRiskMix*...
      outProbIncludingFocusHealthy(1:checkNumber+1,1)...
      );
    %not infection
    infectionDistribution(1:checkNumber+1,1)=...
      infectionDistribution(1:checkNumber+1,1)+...
      probDistMixedPairs(ci,1)*(...
      probFocusInPair(ci,1)*(1-infectionRiskMix)*...
      outProbIncludingFocusHealthy(1:checkNumber+1,1)...
      );
  end
end % loop end  ci=1:maxPairs+1
end  %function end
```

**Example.** The best way to get an idea of what these functions do with what type of input is to look at a concrete example. This example requires the top level routine to give an idea about what values are passed to this function. Listing 8 shows the first part of this routine.

Listing 8: First part of the top level routine for the focus-group split variant.

```
tic
maxPairs=4;
numIndivs=8;
numRounds=3;
probDist=zeros(7+1,2,numRounds+1);
%Round0: 2 infected
probDist(2+1,1,1)=3/4;
probDist(1+1,2,1)=1/4;
%assume G pattern for all
roundHgroup=[0;0;0];
%assume H pattern for focus member
roundHfocus=[1;1;1];

for ri=1:numRounds
  newProbDist=zeros(numIndivs,2);
  for distCond = 1:numIndivs
      if (probDist(distCond,1,ri)+probDist(distCond,2,ri)>0)
          %determine number of mixed pairs for pop with distCond
              -1 infected
          numInfectedGroup=distCond-1;
          numHealthyGroup=numIndivs-1-numInfectedGroup;
          %NEW: addditional member will be added to both possible
               worlds
```

```
22              [probDistMixedPairs,probFocusInPair] =
                    findProbDistMixedPairsEQ(numHealthyGroup,
                    numInfectedGroup,...
23                   maxPairs);
24              infectionRiskGroup=0.05+roundHgroup(ri,1)*0.2;
25              infectionRiskInter=0.05+roundHgroup(ri,1)*0.1+
                    roundHfocus(ri,1)*0.1;
26              infectDist=calcInfectionDistributionEQv3SOM(
                    probDistMixedPairs,probFocusInPair,...
27                   infectionRiskGroup,infectionRiskInter,maxPairs);
28  [...]
```

For the toy example, let us assume that the focus individual chooses action H in the first round, while the group chooses action G. The top level routine will initialize the population state as described above: With $p = 14$ the focus individual is selected as infected and with $p = 34$ they are not. In the first case, the group is always in a state of 1 infections, in the other case in a state with two infections. The variable "probdist" saves the population state for each round. For the first round, its values are thus (note that in Matlab the first dimension always refers to rows the second to columns):

```
1  val(:,:,1) =
2            0            0
3            0       0.2500
4       0.7500            0
5            0            0
6            0            0
7            0            0
8            0            0
9            0            0
```

The top level routine searches for the first row with a sum above 0. The first such row is the second row that corresponds to a population state with 1 infections. For this state, the routine calculates the distribution of mixed pairs and the probability of the focus individual being a member of each number of mixed pairs. This is done by calling

```
1  [probDistMixedPairs,probFocusInPair] =findProbDistMixedPairsEQ(
       numHealthyGroup,numInfectedGroup,...
2                    maxPairs);
```

with the parameters ($h = 6, i = 1$, $max_{mixed} = n/2 = 4$). This results in the values discussed in the context of the function above:

```
1  probDistMixedPairs =
2            0       0.1429
3       1.0000            0
4            0       0.8571
5            0            0
6            0            0
```

and

```
probFocusInPair =
         0         0
    0.1429         0
         0    1.0000
         0         0
         0         0
```

The first distribution shows the distribution of mixed pairs under the two assumptions that the focus individual is healthy (first column) or infected (second column). The first column is irrelevant in this case, as it is impossible in the first round to have seven healthy individuals and one infected (there are two guaranteed infected individuals). In the counterfactual situation that there were seven healthy individuals and one infected individual, one mixed pair would be guaranteed (therefore the 1 in the second column). If the individual is infected, there can be zero or 2 mixed pairs. The second variable gives the probability that the focus individual is in the mixed pairs if either healthy (again, counterfactual) or infected. For two mixed pairs, the probability is obviously zero, and for two mixed pairs, the probability must be one if the focus individual is infected.

The transmission risk for group-group mixed pairs is $p_{inf} = 0.05$, for group-focus mixed pairs it is $p_{inf} = 0.15$. These values (and the maximum possible number of mixed pairs) are the input for the first call to the function that calculates the distribution of infections:

```
infectDist=calcInfectionDistributionEQv3SOM(probDistMixedPairs,
    probFocusInPair,...
                infectionRiskGroup,infectionRiskInter,maxPairs);
```

The algorithm distinguishes several cases. For each possible number of mixed pairs, the probability distribution is calculated once under the assumption that the focus is infected and once under the assumption that the focus is healthy. In practice, the probability of one of these cases is always zero, so only one of these distributions is calculated for each of the number of mixed pairs. For example, it is possible that there are 0 mixed pairs and the focus is infected (in this case, they must be in a pair with the only other infected player).

The counter $ci$ goes from no mixed pairs ($ci = 1$) to the maximum possible number ($ci = maxPairs + 1$). If there are no mixed pairs, the case distinction is simple: The focus must be infected (if there is only one infected individual in the group). The infected individual cannot be in a mixed pair if there are none, and the round is guaranteed to lead to zero new infections. As the probability of seeing 0 mixed pairs for an infected focus individual is 1/7, the probability of this outcome is therefore 1/7 as well ($1/7 \cdot 1$). After this first run through the loop, the variable infectionDistribution has the values:

```
infectionDistribution =
         0         0         0    0.1429
         0         0         0         0
         0         0         0         0
         0         0         0         0
         0         0         0         0
```

The second run through the loop deals with the counter-factual situation of a healthy focus individual and one mixed pair.

```
infectionDistribution =
    0.9357    0.0214         0    0.1429
    0.0429         0         0         0
         0         0         0         0
         0         0         0         0
         0         0         0         0
```

In this situation, it would be most likely that the single pairing does not result in a new infection. If the infected individual meets another group member the chance would be 5% (in 6/7 of all pairings), if meeting the focus individual 15% (in 1/7 of the pairings). The numbers correspond to the multiplication of meeting chance and infection risk with 0.9357 being the complement of the sum to 1.

The next loop deals with the situation that there are two mixed pairs with one infected focus individual included. In this case, there is a single mixed pair with group members, who have a 5% transmission risk with a maximum of 1 infections:

```
groupOnlyProbsIfIncluded =
    0.9500
    0.0500
         0
```

There is a 15% risk of another infection caused by the interaction with the focus individual. This results in the following probabilities of 0,1, or 2 infections in this case:

```
outProbIncludingFocusInfected =
    0.8075
    0.1850
    0.0075
```

The calculation is done in this part of the code:

```
    outProbIncludingFocusInfected(2:checkNumber+1,1) =...
                    groupOnlyProbsIfIncluded(2:checkNumber+1,1) *...
                        (1-infectionRiskMix) +...
                    groupOnlyProbsIfIncluded(1:checkNumber,1) *...
                        infectionRiskMix;
```

The probabilities for this case are multiplied by the probability that there are two mixed pairs including the focus individual if the focus individual is infected ($p = .8571$), and added to the previous probability distribution (in the fourth column).

```
    0.9357    0.0214         0    0.8350
    0.0429         0         0    0.1586
         0         0         0    0.0064
         0         0         0         0
         0         0         0         0
```

All calculations up to this point covered the second row in "probdist", namely the situation in which there is an infected focus individual ($p = 0.25$). The second positive row in this variable is the third row. This covers the case, in which there is a healthy focus individual, and two group individuals are infected ($p = 0.75$).

In this case, only the first two columns of the infection distribution are relevant (as it is impossible that there is an additional infected focus individual):

```
infectDist =
    0.8893    0.0407         0    0.8312
    0.0664    0.0021         0    0.1592
    0.0014         0         0    0.0094
         0         0         0    0.0002
         0         0         0         0
```

Up to two infections can occur. The probability is $p = 0.0014$ that two healthy individuals are infected, and $p = 0.0021$ that one healthy and the focus individual are infected. With a probability of $p = 0.8893$ no individual is infected. Again, these probabilities already incorporate the probability that the focus individual is within the mixed pairs (and the corresponding higher infection risk).

In general, the following four case are thus considered in the determination of the distribution of infections:

1. Focus individual is infected, not included in mixed pairs
2. Focus individual is infected, included in mixed pairs
3. Focus individual is healthy, not included in mixed pairs
4. Focus individual is healthy, included in mixed pairs

In contrast, the columns of the function output correspond to:

1. Assumes that the focus individual is healthy: Probability of all possible numbers of infections in the group, while the focus individual stays healthy.
2. Assumes that the focus individual is healthy: Probability of all possible numbers of infections in the group, while the focus individual is infected in the round (values in column 1 and column 2 add up to 1).
3. Assumes that the focus individual is infected: Probability of all possible numbers of infections in the group, while the focus individual ends the round as healthy. This is impossible, the column will always contain zeros and is kept only for symmetry.
4. Assumes that the focus individual is infected: Probability of all possible numbers of infections in the group, while the focus individual is still infected after the round (values in column 3 and column 4 add up to 1).

#### *E5.3.4 Putting the parts together*

The group-focus split follows the same logic as the previous analysis, but is somewhat complicated by the two-dimensional population representation.

1. The probability distribution of population states $I$ before the round (with the probability of the focus individual being infected implicitly encoded),
2. the probability distribution of the number of mixed pairs $M$ conditional on the population state, and the probability of the focus individual contained in these pairs, and
3. the probability distribution of the number of infections $p(n_{inf})$ conditional on the number of mixed pairs (and the status of the focus individual).

The top level algorithm joins these elements together as before and calculates the population state for each round. Again, the population state contains all the necessary history for the game to be calculated in its entirety. Listing 9 shows the top level code (for the standard game) in its entirety.

Listing 9: Full top level routine for the focus-group split variant.

```
tic
maxPairs=50;
numIndivs=100;
numRounds=25;
probDist=zeros(99+1,2,numRounds+1);
%Round0: 8 infected
probDist(8+1,1,1)=92/100;
probDist(7+1,2,1)=8/100;
%assume G pattern for all
roundHgroup=[0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0];
%assume H pattern for focus member
roundHfocus=[1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1];

for ri=1:numRounds
  newProbDist=zeros(numIndivs,2);
  for distCond = 1:numIndivs
      if (probDist(distCond,1,ri)+probDist(distCond,2,ri)>0)
          %determine number of mixed pairs for pop with distCond
              -1 infected
          numInfectedGroup=distCond-1;
          numHealthyGroup=numIndivs-1-numInfectedGroup;
          %NEW: addditional member will be added to both possible
               worlds
          [probDistMixedPairs,probFocusInPair] =
              findProbDistMixedPairsEQ(numHealthyGroup,
              numInfectedGroup,...
```

```
23                  maxPairs);
24              infectionRiskGroup=0.05+roundHgroup(ri,1)*0.2;
25              infectionRiskInter=0.05+roundHgroup(ri,1)*0.1+
                    roundHfocus(ri,1)*0.1;
26              infectDist=calcInfectionDistributionEQv3SOM(
                    probDistMixedPairs,probFocusInPair,...
27                   infectionRiskGroup,infectionRiskInter,maxPairs);
28              maxIncrease=min(distCond-1+maxPairs,numIndivs-1)-(
                    distCond-1);
29
30              % Focus is healthy after round :
31              %   This means healthy focus that stays healthy (
                    column 1 in infectDist)
32               newProbDist(distCond:distCond+maxIncrease,1)=...
33                  newProbDist(distCond:distCond+maxIncrease,1)+...
34                      probDist(distCond,1,ri)*infectDist(1:maxIncrease
                            +1,1);
35
36               % Focus is infected after round:
37               % infected individuals that stay infected (column 2)
38               % and healthy indivduals that get infected (column 4)
39                newProbDist(distCond:distCond+maxIncrease,2)=...
40                  newProbDist(distCond:distCond+maxIncrease,2)+...
41                      probDist(distCond,1,ri)*infectDist(1:maxIncrease
                            +1,2)+...
42                      probDist(distCond,2,ri)*infectDist(1:maxIncrease
                            +1,4);
43          end
44
45      end
46      probDist (1:numIndivs,1:2,ri+1)= newProbDist;
47  end %rounds loop
48
49  toc
50  finalDist=probDist(1:numIndivs,1:2,numRounds+1);
51  groupSum=sum(finalDist,2);
52  colSum=sum(finalDist);
53
54  avIndiv=colSum(1,2)*100
55  avGroup=100*([0:99]*groupSum)/99
56
57  survIndiv=100-avIndiv
58  survGroup=100-avGroup %note: proability in percent
```

As before, the end state allows to calculate the expected infection probability, in this case separately for group and individual. For the group, this amounts to:

$$\pi_{infection,g} = \frac{1}{99} \cdot \sum_{i=0}^{99} (p_{25,i,+} + p_{25,i,-}) \cdot i \tag{37}$$

For the focus individual, this can be calculated as:

$$\pi_{infection,f} = \sum_{i=0}^{99} p_{25,i,+} \tag{38}$$

The expected score for each population member is again calculated by counting the number of H-choices and multiplying the respective final scores with the respective survival rate (the probability of not being infected):

$$E_g(O) = (1 - \pi_{infection,g}) \cdot (200 + 32 \cdot \sum_{i=1}^{25} s_{g,i}), \tag{39}$$

and

$$E_f(O) = (1 - \pi_{infection,f}) \cdot (200 + 32 \cdot \sum_{i=1}^{25} s_{g,f}). \tag{40}$$

**Example.** To demonstrate the differences in notation while playing the same game as before, we start out by repeating the game played to illustrate the simpler analysis format. We can achieve this by setting both the group's and the focus individual's strategy to be the same ($S_g = s_f = (1, 1, 0)$). The resulting evolution of population states is shown in Table E4.

**Table E4**

*Progression of population states across three rounds in the toy example when all players choose the strategy $S = (1, 1, 0)$.*

| | Round | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0 | | 1 | | 2 | | 3 | |
| Pop. state | - | + | - | + | - | + | - | + |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0.250 | 0 | 0.156 | 0 | 0.098 | 0 | 0.084 |
| 2 | 0.750 | 0 | 0.469 | 0.121 | 0.293 | 0.143 | 0.269 | 0.140 |
| 3 | 0 | 0 | 0.201 | 0.027 | 0.239 | 0.087 | 0.234 | 0.097 |
| 4 | 0 | 0 | 0.027 | 0 | 0.087 | 0.028 | 0.098 | 0.038 |
| 5 | 0 | 0 | 0 | 0 | 0.017 | 0.006 | 0.023 | 0.009 |
| 6 | 0 | 0 | 0 | 0 | 0.0023 | <0.001 | 0.003 | 0.001 |
| 7 | 0 | 0 | 0 | 0 | <0.001 | <0.001 | <0.001 | <0.001 |
| $\pi_{infection,f}$ | 0.250 | | 0.304 | | 0.362 | | 0.375 | |
| $\pi_{infection,g}$ | 0.250 | | 0.304 | | 0.362 | | 0.375 | |

As it should be, the infection probabilities for group members and the focus individual are the same in each round and are also the same as the probabilities observed before (see Table E3). As described above, the infection probability for the individual can be computed by adding the probabilities in the +-column, the probabilities for the group are the average products of row sum and row number (0–7).

A more interesting application of the algorithm is of course a case, in which strategies diverge. This is presented in Table E5 for the situation that the group chooses strategy $S_g = (0, 0, 0)$ and the individual the riskier strategy $S_f = (1, 1, 1)$.

**Table E5**

*Progression of population states across three rounds in the toy example when group players choose the strategy $S_g = (0, 0, 0)$ and the focus individual chooses $S_f = (1, 1, 1)$.*

| | Round | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0 | | 1 | | 2 | | 3 | |
| Pop. state | - | + | - | + | - | + | - | + |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0.250 | 0 | 0.209 | 0 | 0.174 | 0 | 0.146 |
| 2 | 0.750 | 0 | 0.667 | 0.070 | 0.593 | 0.119 | 0.528 | 0.150 |
| 3 | 0 | 0 | 0.050 | 0.003 | 0.087 | 0.020 | 0.114 | 0.043 |
| 4 | 0 | 0 | 0.001 | 0 | 0.006 | 0.002 | 0.012 | 0.006 |
| 5 | 0 | 0 | 0 | 0 | <0.001 | <0.001 | 0.001 | 0.001 |
| 6 | 0 | 0 | 0 | 0 | <0.001 | <0.001 | <0.001 | <0.001 |
| 7 | 0 | 0 | 0 | 0 | <0.001 | <0.001 | <0.001 | <0.001 |
| $\pi_{infection,f}$ | 0.250 | | 0.282 | | 0.314 | | 0.346 | |
| $\pi_{infection,g}$ | 0.250 | | 0.264 | | 0.278 | | 0.293 | |

In each round the infection probability for the focus individual increases by a larger amount than the infection probability for the group. We can also determine the expected outcome for group and individual. The expected average score in this game for the group is

$$E_g(O) = (1 - \pi_{infection,g}) \cdot (24 + 32 \cdot \sum_{i=1}^{3} s_{g,i}) = 0.707 \cdot 24 = 16.97. \quad (41)$$

The expected average score in this game for the individual is

$$E_f(O) = (1 - \pi_{infection,f}) \cdot (24 + 32 \cdot \sum_{i=1}^{3} s_{f,i}) = 0.654 \cdot 120 = 78.48. \quad (42)$$

While the survival probability is lower for the riskier strategy, the expected outcome is clearly higher.

## E5.4 Extensions of the analysis tool

### *E5.4.1 Partial games*

Given that the game state carries all the relevant history, it becomes possible to calculate the results for partial games by entering probability distributions into a routine that simulates a particular range of rounds of the game. Listing 10 shows an implementation of such a function. The function needs a number of input variables: A population state as probability distribution, the range of rounds to compute, and the strategy vector that all population members play.

Listing 10: Function for simulating partial games based on a population state

```
function [ probDistResult ] = distributionOverRoundsUniform(
    probDist,roundStart,...
     roundEnd,roundH)
umIndivs=size(probdist,1)-1;
maxPairs=numIndivs/2;
for ri=roundStart:roundEnd
  newProbDist=zeros(numIndivs+1,1);
  for distCond = 1:numIndivs+1
      if probDist(distCond,1)>0
          %determine number of mixed pairs for pop with distCond
              -1 infected
          numInfected=distCond-1;
          numHealthy=100-numInfected;
          distMixedPairs=findProbDistMixedPairs(numHealthy,
              numInfected,...
               maxPairs);
          infectionRisk=0.05+roundH(ri,1)*0.2;
          infectDist=calcInfectionDistribution(distMixedPairs,
              infectionRisk,...
                maxPairs);
```

```
            maxIncrease=min(distCond-1+maxPairs,numIndivs)-(
                distCond-1);
            newProbDist(distCond:distCond+maxIncrease)=...
                newProbDist(distCond:distCond+maxIncrease)+...
                    probDist(distCond,1)*infectDist(1:maxIncrease
                        +1,1);
        end
    end
    probDist(1:numIndivs+1,1)= newProbDist;
end
probDistResult=probDist;
end
```

One specific use of this function is to simulate all possible pure states (i.e., all possible 101 number of infections with a probability of 1 at a time). Linking each of these distributions with outcome distributions allows to compute the outcome for every possible starting distribution as a weighted distribution across the linked distributions. This can speed up calculations if other parts of the game are to be varied, and it is also useful for analyzing properties of the game (see below).

### *E5.4.2 Modelling player groups: Some progress and obstacles*

Previous analyses have been restricted by the fact that only a single group strategy was considered with a maximum of a single deviating individual. Lifting this restriction comes at considerable costs, though. Here, we consider the case of a single round, in which one group of players chooses action G and the remainder of the group chooses action H. Before and after the round, all players choose the same actions. At first glance, this seems like a simple scenario, but a closer analysis reveals some complications.

Transmission probabilities depend on the combination of strategies within the mixed pairs, in particular the number of HH-combinations, GH-combinations, HG-combinations, and GG-combinations, where the first letter denotes the action chosen by the infected player. To calculate the numbers we introduce a helper functions (see Listing 11). It calculates the number of combinations in which $k$ elements can be drawn from $n$ elements (where the order of draws matters). If four parameters are given, the function calculates the product of both combinations:

$$f(k_1, n_1, k_2, n_2) = \frac{n_1!}{(n_1 - k_1)!} \cdot \frac{n_2!}{(n_2 - k_2)!} \tag{43}$$

Listing 11: Calculating the number of combinations (order matters).

```
function [ numComb ] = numberOfCombinations( numA,totalA,numB,
    totalB )
if nargin <3
    numB=0;
    totalB=0;
end
```

```
if (numA>0) && (numB>0)
    numComb=factorial(totalA)/factorial(totalA-numA)*(factorial(
        totalB)/factorial(totalB-numB));
elseif numA>0
    numComb=factorial(totalA)/factorial(totalA-numA);
elseif numB>0
    numComb=factorial(totalB)/factorial(totalB-numB);
else
  numComb=1;
end
end
```

Listing 12 uses this function to calculate the probability of finding different numbers of combination types for known numbers of G-players, H-players and mixed pairs.

Listing 12: Code for finding the number of combination types.

```
function [ distributionMixedPairtypeProbs ] =
    determineDistributionOfMixedPairsGH...
    (numG,numH,numMixed)

%Distribution: Infected->Healthy: H->H, H->G, G->H, G->G
 typeProb=zeros(numMixed+1,numMixed+1,numMixed+1,numMixed+1);
 popSize=numH+numG;
 totalCombinations=factorial(popSize)/factorial(popSize-2*
    numMixed);%combinations to fill pairs with popmembers

 for numHH=0:numMixed
   for numHG=0:(numMixed-numHH)
       for numGH=0:(numMixed-numHH-numHG)
           numGG=numMixed-numHH-numHG-numGH;
           if (2*numHH+numHG+numGH <= numH) && (2*numGG+numHG+
               numGH <=numG)
               % need to consider number of orders of elements
               % all GG: only one order
            numElements=0;
            numElementTypes=0;
            if numHH>0
                numElements=numElements+numHH;
                numElementTypes=numElementTypes+1;
            end
            if numGH>0
                numElements=numElements+numGH;
                numElementTypes=numElementTypes+1;
            end
            if numHG>0
                numElements=numElements+numHG;
```

```
                numElementTypes=numElementTypes+1;
            end
            if numGG>0
                numElements=numElements+numGG;
                numElementTypes=numElementTypes+1;
            end

            % there are numMixed! orders of the pairs, but within
                pairs of the same category the order does not
                matter
            numberOfPairOrders= factorial(numMixed)/(factorial(
                numHH)*factorial(numGH)*factorial(numHG)*...
                factorial(numGG));

        %determine ways all mixed pairs can be filled from the
            two groups of H- and G-action choosing individuals /
            divided by the total number of combinations
        typeProb(numHH+1,numHG+1,numGH+1,numGG+1)=...
            numberOfCombinations(2*numHH,numH)*....
            numberOfCombinations(numHG,numG,numHG,(numH-2*numHH
                ))*....
            numberOfCombinations(numGH,numG-numHG,numGH,(numH
                -2*numHH-numHG))*....
            numberOfCombinations(2*numGG,(numG-numGH-numHG))
                *...
            numberOfPairOrders/....
            totalCombinations;
        end
    end
  end
 end
%typeProb contains the probability that all possible combinations
    of HH, HG, GH and HH occur
distributionMixedPairtypeProbs=typeProb;
end
```

This function allows knowledge for each possible population state and each number of mixed pairs. Given this knowledge, it can be calculated how many infected and healthy individuals remain in healthy-healthy and infected-infected pairs, by subtracting all players involved in the mixed pairs.

As there is no prior difference between individuals assumed, the remaining H-choosing and G-choosing individuals are distributed randomly among infected and healthy. This distribution is an instance of a known distribution, namely the hypergeometric distribution. If the number of remaining G-choosers/H-choosers is $n_G$ and $n_H$, respectively, and the number of infected and healthy is $n_i$ and $n_h$ with $n_G + n_H = n_i + n_h$, then the number of healthy G-choosers is distributed as:

$$P(n_{Gh}=k)=f(k;n_G+n_H,n_h,n_G)=\frac{\binom{n_h-k}{k}\binom{n_G+n_H-n_h+k}{n_G-k}}{\binom{n_G+n_H}{n_G}} \tag{44}$$

For example, if $n_G=3$, $n_H=9$, $n_h=2$, and $n_i=10$, then

$$P(n_{Gh}=0)=f(0;12,2,3)=\frac{\binom{2}{0}\binom{12-2}{3-0}}{\binom{12}{3}} \tag{45}$$

$$=\frac{1\cdot\binom{10}{3}}{\binom{12}{3}}=\frac{\frac{10\cdot9\cdot8}{3\cdot2\cdot1}}{\frac{12\cdot11\cdot10}{3\cdot2}}=\frac{120}{220}=\frac{6}{11}, \tag{46}$$

$$P(n_{Gh}=2)=f(2;12,2,3)=\frac{\binom{2}{2}\binom{12-2}{3-2}}{\binom{12}{3}} \tag{47}$$

$$=\frac{1\cdot10}{\binom{12}{3}}=\frac{10}{\frac{12\cdot11\cdot10}{3\cdot2}}=\frac{10}{220}=\frac{1}{22}, \tag{48}$$

and

$$P(n_{Gh}=1)=f(1;12,2,3)=\frac{\binom{2}{1}\binom{12-2}{3-1}}{\binom{12}{3}} \tag{49}$$

$$=\frac{2\cdot\binom{10}{2}}{\binom{12}{3}}=\frac{2\cdot\frac{10\cdot9}{2\cdot1}}{\frac{12\cdot11\cdot10}{3\cdot2}}=\frac{90}{220}=\frac{9}{22} \tag{50}$$

Knowing this distribution and the distribution of mixed pair types it would in theory be possible to iterate through all possible population states, all possible numbers of mixed pairs, all possible distributions of types within these mixed pairs and all possible distributions of players outside the mixed pairs, determine next-round results for each of these states and weight these results by probabilities to determine the next round's distribution. We decided to draw the line at this point for this manuscript, as the use case for this algorithm is of little general interests, and simulation results can easily take the place of the formal analysis.

More importantly, these considerations illustrate the reason why it is unlikely that the full, unrestricted game can be successfully analyzed with this approach. If each unique group of strategies need to be represented as a probability distribution, one might end up easily with 100 distinct strategies. This would imply $2^{100}$ possible population states, most of which would have a positive probability. Iterating through this states alone is not feasible. Take the following code in Listing 13.

Listing 13: Code for estimating the duration of iterating through states alone.

```
tic
nIterations=1*10^9;
for ci=1:nIterations
```

```
    a=1;
end
toc
```

For $n = 10^9$ iterations, the code runs for 2 to 2.1 seconds on the writer's machine. This scales quite linearly. For $n = 10^{10}$ iterations (10,000,000,000 iterations), the code needs 20.5 to 21.5 seconds to compute. $2^{100}$ corresponds to about $1.27 \cdot 10^{30}$, which is $1.27 \cdot 10^{20}$ times the amount of the previous exercise. Using a lower bound of 20 seconds for $10^{10}$ iterations, this amounts to $2.54 \cdot 10^{21}$ seconds to do a simple variable assignment for each iteration. This amounts to $4.23 \cdot 10^{19}$ minutes, $7.04 \cdot 10^{17}$ hours, $2.93 \cdot 10^{1}6$ days, or $8.03 \cdot 10^{13}$ years. If the age of the earth is assumed to be 4.543 billion years ($4.53 \cdot 10^9$), this would be about 17,700 times the age of the earth, just for completing a simple iteration. Thus, while this approach might be feasible to be programmed, we would be skeptical about its practical use in the coming decades.

## E6 Analytical results for the transmission game

The analytical results developed in the preceding section offer opportunities for revisiting some previous results and extending the analysis to new questions. Here, we will start by comparing the new analytical results with previous simulation estimates. We then repose the question of equilibria and "optimal" behavior, but end in a cautionary note (and demonstration) on why these optimality results need to be taken with a grain of salt.

### E6.1 Confirming the results of previous simulations

Showing that simulation results and analytical results align is supportive evidence for claiming that our simulations are correct implementations of the described game and our analytical tools are indeed able to compute their expected results (if one excludes the possibility of a common error in both).

#### *E6.1.1 Revisiting the flattening of the curve*

To demonstrate the equivalence, we choose the earlier results on flattening the curve. In this simulation, we simulated all possible strategies that start with H-actions and from some round on switch to G-actions for the remainder of the game. Depending on the switching point, we found different shapes of infection curves with large differences in final states.

Calculating these results with our analytic tools is straightforward (see Listing 14). The calculations take a fraction of the time for 1,000,000 trials of the games (on most PCs it should take less than an hour).

Listing 14: Top level routine for calculating the results of the HG-simulation.

```
tic
maxPairs=50;
numIndivs=100;
numRounds=25;
probDist=zeros(100+1,numRounds+1);
probDist(8+1,1)=1;
```

```
collectDist=zeros(26,27);
for ci=0:25
    roundH=zeros(25,1);
    if (ci>0)
        roundH(1:ci,1)=1;
    end
for ri=1:numRounds
  newProbDist=zeros(numIndivs+1,1);
  for distCond = 1:numIndivs+1
      if probDist(distCond,ri)>0
          %determine number of mixed pairs for pop with distCond
              -1 infected
          numInfected=distCond-1;
          numHealthy=100-numInfected;
          distMixedPairs=findProbDistMixedPairs(numHealthy,
              numInfected,...
               maxPairs);
          infectionRisk=0.05+roundH(ri,1)*0.2;
          infectDist=calcInfectionDistribution(distMixedPairs,
              infectionRisk,...
                maxPairs);
          maxIncrease=min(distCond-1+maxPairs,numIndivs)-(
              distCond-1);
          newProbDist(distCond:distCond+maxIncrease)=...
               newProbDist(distCond:distCond+maxIncrease)+...
                  probDist(distCond,ri)*infectDist(1:maxIncrease
                      +1,1);
      end
  end
  probDist (1:numIndivs+1,ri+1)= newProbDist;
end

toc
meanDist=[0:100]*probDist;
collectDist(ci+1,1)=ci;
collectDist(ci+1,2:27)=meanDist;
dlmwrite("[FOLDERNAME]/res/HGAnalysis.txt", collectDist,'
   delimiter', "\t",'precision',8);
end
```

Based on these results, we were able to calculate the exact expected values for the simulation results. Figure E13 directly compares simulation results and analytical results. The maximum absolute difference across all plotted values is smaller than 0.00018 (or 0.018%), with no systematic direction of deviation.

### *E6.1.2 Roundwise results*

The results generated by the analysis tools are of course richer than the expected end points for infection numbers. For each round, the tool generates the precise probability distribution for population states. For a selected strategy (an HG-strategy with six initial H-actions), the progression of population state distributions across rounds is shown in its entirety in Figure E14.

The mean of each distribution corresponds to the points in Figure E13 that are located on the seventh line from the bottom (the bottom-most line corresponds to 0 H-choices). Naturally, the longer the game lasts the larger the variance of game states will be (at least for infection rates below 50%). Extremely high values are still possible (starting after round 4, 100 infected players are technically possible), but not probable.

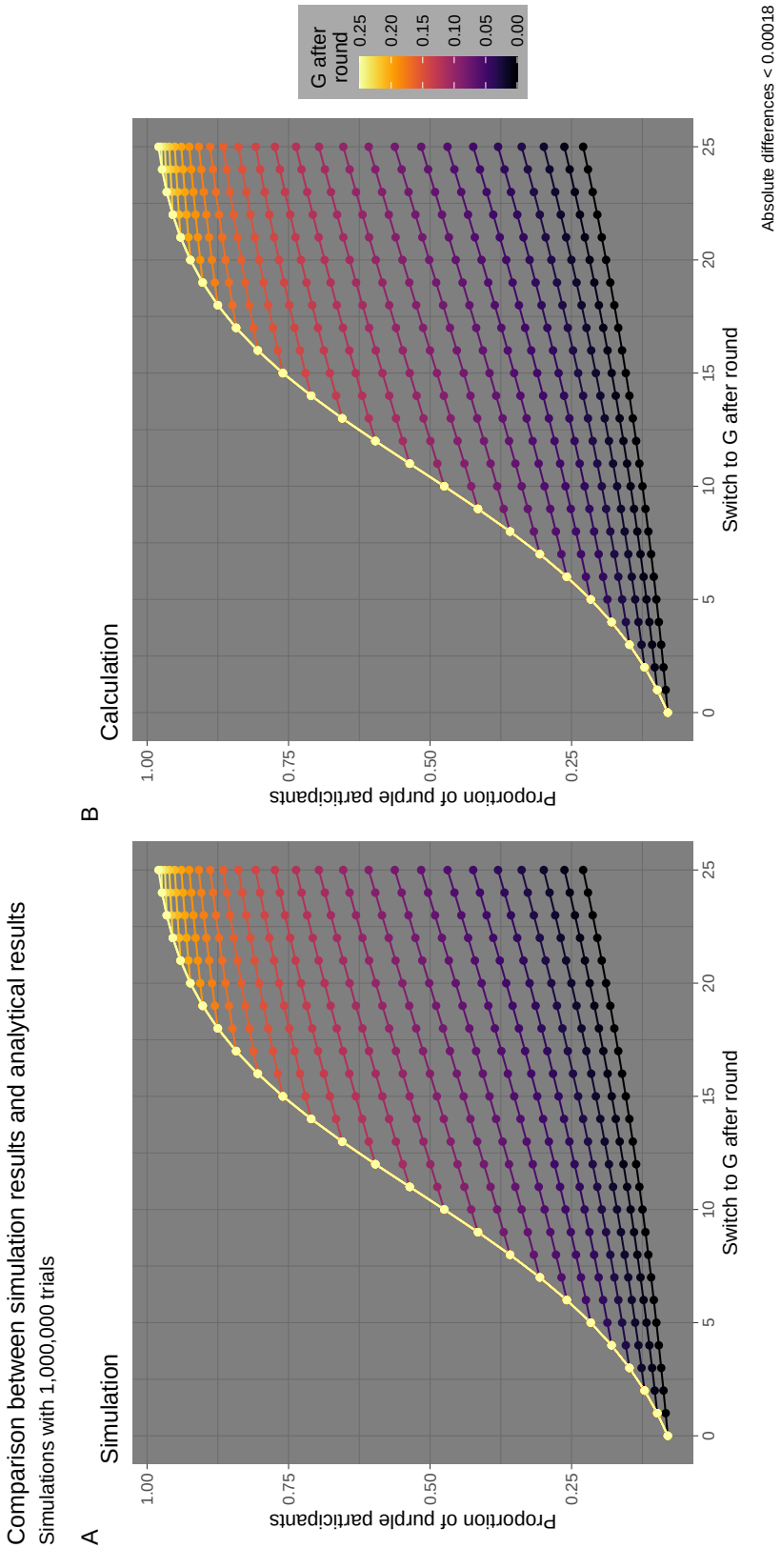


**Figure E13**
*Average proportion of purple participants after each round for each possible collective switching point from H-actions to G-actions: (A) points mark the mean across 1,000,000 simulated games (B) points mark the expected values based on analytical results.*

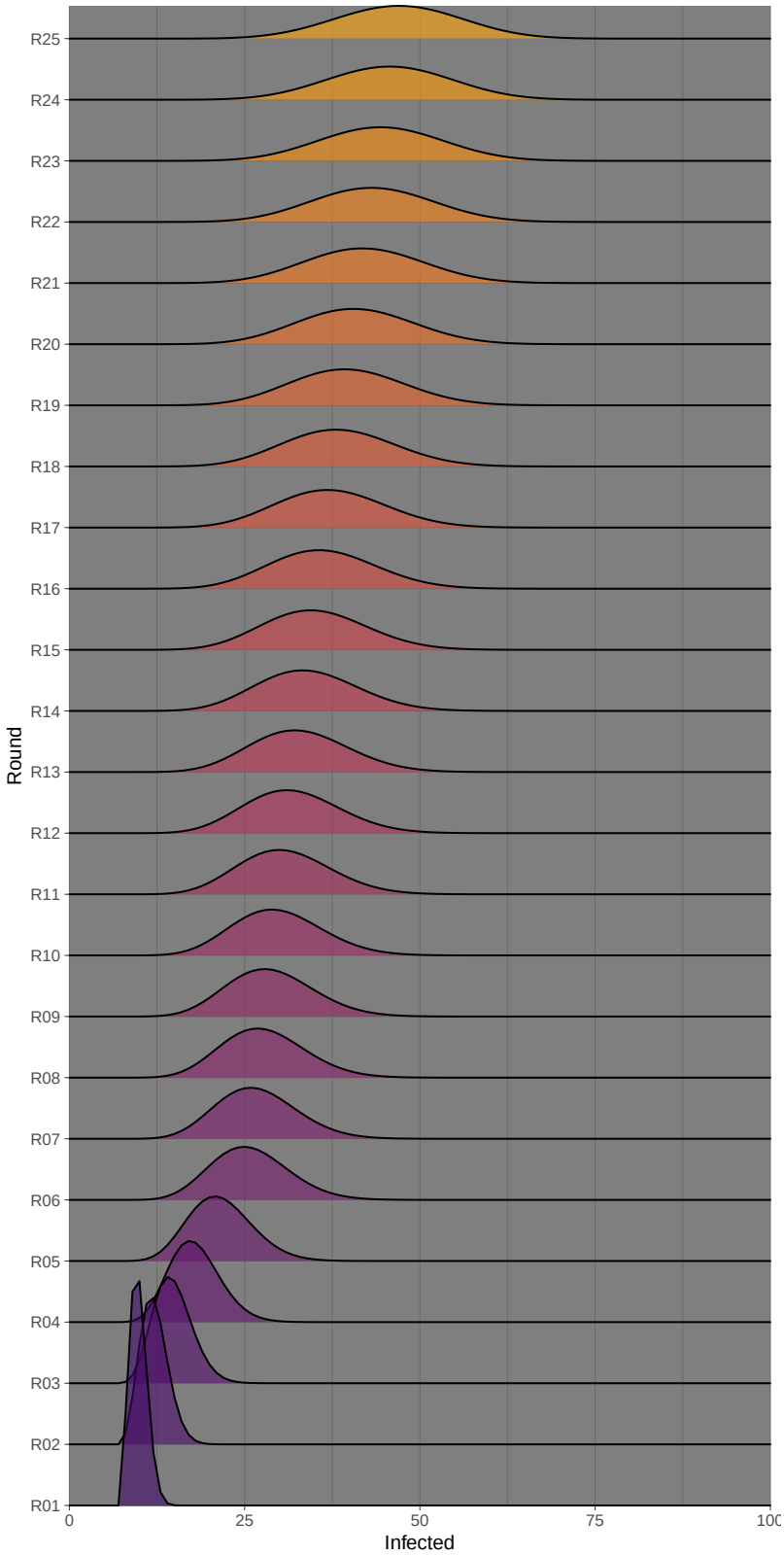


**Figure E14**

*Derived roundwise distributions of computed population states for a HG-strategy with 6 H-actions in the first six rounds and 19 G-actions.*

## E6.2 Collusive equilibrium

We can revisit our previous discussion of the collusive optimum, defined as the the universal strategy with the highest expected outcome. We demonstrated through simulations that HG-strategies dominated strategies of fixed proportions of H-players. We also tested a variety of other strategies adopted by the entire population.

Here, we can use the tools to derive some analytical results that can bolster the confidence in the identified solution. There are two conflicting intuitions in conceptualizing whether it is advantageous to choose risky actions sooner or later in the game: For one, infection rates will increase across round, no matter what players do. This implies that later rounds of the game are more dangerous in increasing the chance to meet infected individuals which makes riskier actions even more risky in later rounds. At the same time, players that are infected early might spread the infection to more players if playing the riskier actions sooner who can themselves infect other players, etc.

### *E6.2.1 Analysis results*

One way to test these intuitions directly against each other is to simulate a game of two rounds with two separate populations that utilize two different strategies. In the first population, everyone chooses action H first and action G second, in the second population the action order is reversed. Both populations will accumulate the same number of points (both individually and collectively), the survival rate alone thus solely determines the game outcome (see Listing 15).

As the outcome of this comparison between populations might rest on the initial state of the populations, we compared the consequences of the two strategies for each possible population state. We computed the difference in survival rates by subtracting the survival rate for the HG-strategy from the survival rate of the GH-strategy. Negative values thus indicate a relative advantage for the HG-strategy.

Listing 15: Code for sytematically comparing the HG-strategy and the GH-strategy.

```
tic
maxPairs=50;
numIndivs=100;
numRounds=2;
probDist=zeros(100+1,numRounds+1);
survivalGH=zeros(101,1);
survivalHG=zeros(101,1);

for statei=0:100
  probDist=zeros(100+1,1);
  probDist(statei+1,1)=1;

  roundH=[1;0];
  resDistHG  = distributionOverRoundsUniform( probDist,1,...
    numRounds,roundH,maxPairs,numIndivs );
  survivalHG(statei+1,1)=1-[0:100]*resDistHG/100;
```

```
17
18    roundH=[0;1];
19    resDistGH   = distributionOverRoundsUniform( probDist,1,...
20      numRounds,roundH,maxPairs,numIndivs );
21    survivalGH(statei+1,1)=1-[0:100]*resDistGH/100;
22  end
23  survivalLoss=survivalGH-survivalHG;
24  summary=[survivalHG,survivalGH,survivalLoss];
25  dlmwrite("[FOLDERNAME]/res/survivalHG-GH-Diff.txt",...
26      summary,'delimiter','\t','precision',8);
27  toc
```

There is not a single population state for which the difference between the survival rates is positive (see also the uppermost curve in Figure E16). Differences are not large with a minimum at -0.016%. The difference is zero for extreme values: When there is no infection or everyone if infected, strategies cannot make any difference. When all but one individual are infected, the timing of the H-action makes no difference, as both partners are guaranteed to be infected independent of strategy.

### *E6.2.2 Implications of results*

This result is more powerful than it might appear at first glance: Given that the possible result for a strategy depends solely on the number of H-actions, this result implies that a strategy is superior to any strategy that is generated by switching two adjacent actions when the first action is an H-action and the second action is a G-action.

Consider two strategies $S_{HG}$ and $S_{GH}$, with

$$(\exists i)(s_{GH,i} = 0 \wedge s_{HG,i} = 1 \wedge s_{GH,i+1} = 1 \wedge s_{HG,i+1} = 0) \wedge \\ (\forall j)((j \neq i \wedge j \neq i+1) \supset s_{HG,j} = s_{GH,j}), i \in [1,24], j \in [1,25] \tag{51}$$

Both strategies will enter round $i$ in the same state, and go through the same actions in any subsequent round. An extremely high infection rate may occur, but there is no strategy possible that would result in a guaranteed extreme case (even for pure H-strategy the distribution includes lower infection numbers). The result above now shows that strategy $S_{HG}$ will enter the subsequent part of the game with a lower average infection rate, resulting in a lower final infection rate.

More importantly, for any possible strategy with $k$ H-actions that is not a HG-strategy, the corresponding HG-strategy can be reached by repeatedly switching G-actions with subsequent H-actions. Take the string "HHHHGHGGHHHGGHGGGGGHGGGGG" and see the 25 required steps in Figure E15. Any strategy that is not an HG-strategy will by necessity have G-actions preceding H-actions and thus can be transformed in the described way by moving the H actions to earlier rounds.

### *E6.2.3 Testing the implications*

Based on the result shown above we know that each of these swaps will increase the survival rate of a population that moves from the previous strategy to the new one. To demonstrate this consequence, we repeated the earlier analysis for HG and GH (one swap

| Transformation steps |
|---|
| HHHHGHGGHHHGGHGGGGGHGGGGG |
| HHHHHGGGHHHGGHGGGGGHGGGGG |
| HHHHHGGHGHHGGHGGGGGHGGGGG |
| HHHHHGHGGHHGGHGGGGGHGGGGG |
| HHHHHHGGGHHGGHGGGGGHGGGGG |
| HHHHHHGGHGHGGHGGGGGHGGGGG |
| HHHHHHGHGGHGGHGGGGGHGGGGG |
| HHHHHHHGGGHGGHGGGGGHGGGGG |
| HHHHHHHGGHGGGHGGGGGHGGGGG |
| HHHHHHHGHGGGGHGGGGGHGGGGG |
| HHHHHHHHGGGGGHGGGGGHGGGGG |
| HHHHHHHHGGGGHGGGGGHGGGGG |
| HHHHHHHHGGGHGGGGGGHGGGGG |
| HHHHHHHHGGHGGGGGGGHGGGGG |
| HHHHHHHHGHGGGGGGGGHGGGGG |
| HHHHHHHHGHGGGGGGGGHGGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHGGGGGGGGGGHGGGG |
| HHHHHHHHHHGGGGGGGGGGGGGG |

**Figure E15**

*Example for a transformation of a strategy into an HG-strategy by repeatedly swapping G and H. The pair of actions that is swapped in the subsequent row is shown in red.*

necessary) for the strategy pairs HGG and HGG (two swaps), HGGG and GHHH (three swaps) and HGGGG and GGGGH (four swaps). The results are presented in Figure E16.

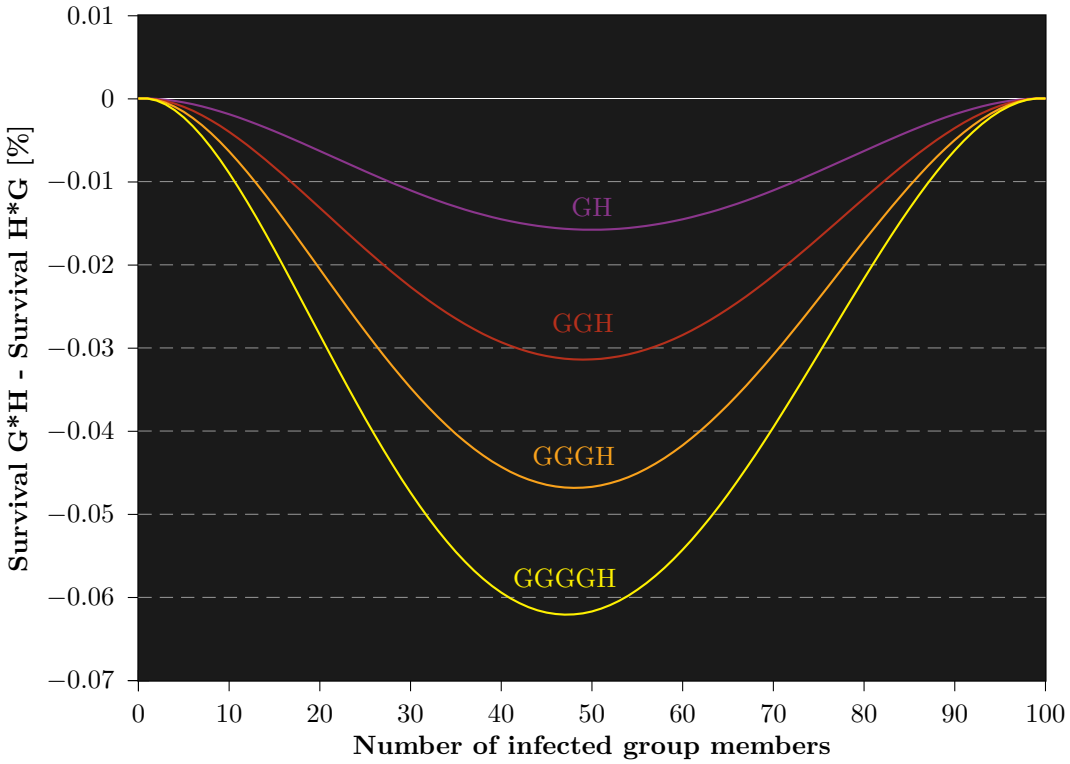


**Figure E16**
*Difference in survival rates for players with strategies starting with a single H action or ending with it: Games last two (GH) to five (GGGGH) rounds and the difference in average survival rates is plotted for each possible population state in the first round of the game. Negative values indicate that the H*G strategy leads to higher survival than the G*H strategy.*

As expected based on the previous explanations, the difference between strategies increases, the more swaps are necessary to move a trailing H action to the front.

Taken these results together, an HG-strategy with six H-actions and 19 G-actions seems to be the optimal strategy if played by every population member. We will discuss some caveats after dealing with the individual level.

### E6.3 Search for the individual equilibrium

The previous results hold for a population with a unified strategy. On the individual level, there is no reason to assume that this strategy is chosen over an alternative that allows the individual to attain higher outcomes. In contrast, in a Nash equilibrium, no single player would be able to achieve a better outcome by deviating from their strategy. The search for a possible Nash equilibrium is therefore more involved than the search for the collusive optimum.

### *E6.3.1 Individual HG-patterns*

First, we explore whether a similar result can be shown for individual-level strategies as what we found for universal strategies. To do that, we need to show that individuals cannot benefit from choosing GH-patterns over HG patterns. On this level, we need to conceive of the population as independent from the individual.

In a first step, we adapt the previous simulation to the individual level by assuming a population that chooses G-action all of the time. As before, we consider games of two, three, and four rounds, and this time we also consider a game of eight rounds. As outcome, we consider the difference in survival probability (and therefore expected outcome) for individuals that choose a strategy starting with a single H-action and those ending with a single H-action (see Listing 16).

Listing 16: Code for systematically comparing the HG-strategy and the GH-strategy on the individual level.

```
tic
maxPairs=50;
numIndivs=100;
numRounds=3;
savedResultsSurvival=zeros(100,2);
rowNumber=0;
for indivi=1:2
  for groupi=0:99
      out=groupi
      rowNumber=rowNumber+1;
      probDist=zeros(99+1,2,numRounds+1);
      probDist(groupi+1,1,1)=1;
      roundHgroup=[0;0;0];
      if indivi==1
         roundHfocus=[1;0;0];
      else
         roundHfocus=[0;0;1];
      end
      for ri=1:numRounds
         newProbDist=zeros(numIndivs,2);
            for distCond = 1:numIndivs
                  if (probDist(distCond,1,ri)+probDist(distCond,2,
                     ri)>0)
                     numInfectedGroup=distCond-1;
                     numHealthyGroup=99-numInfectedGroup;
                     [probDistMixedPairs,probFocusInPair] =
                        findProbDistMixedPairsEQ(numHealthyGroup,
                        numInfectedGroup,...
                       maxPairs);
                     infectionRiskGroup=0.05+roundHgroup(ri,1)
                        *0.2;
```

```
                    infectionRiskInter=0.05+roundHgroup(ri,1)
                        *0.1+roundHfocus(ri,1)*0.1;
                    infectDist=calcInfectionDistributionEQv2(
                        probDistMixedPairs,probFocusInPair,...
                            infectionRiskGroup,infectionRiskInter,
                                maxPairs);
                    maxIncrease=min(distCond-1+maxPairs,numIndivs
                        -1)-(distCond-1);
                    newProbDist(distCond:distCond+maxIncrease,1)
                        =...
                        newProbDist(distCond:distCond+maxIncrease
                            ,1)+...
                        probDist(distCond,1,ri)*infectDist(1:
                            maxIncrease+1,1);
                    newProbDist(distCond:distCond+maxIncrease,2)
                        =...
                      newProbDist(distCond:distCond+maxIncrease
                            ,2)+...
                      probDist(distCond,1,ri)*infectDist(1:
                          maxIncrease+1,2)+...
                      probDist(distCond,2,ri)*infectDist(1:
                          maxIncrease+1,4);
                end
            end
        probDist(1:numIndivs,1:2,ri+1)= newProbDist;
    end
    finalDist=probDist(1:numIndivs,1:2,numRounds+1);
    groupSum=sum(finalDist,2);
    colSum=sum(finalDist);
    avIndiv=colSum(1,2)*100;
    avGroup=100*([0:99]*groupSum)/99;
    survIndiv=100-avIndiv;
    survGroup=100-avGroup;
    pointsGroup=sum(roundHgroup)*32+200;
    pointsIndiv=sum(roundHfocus)*32+200;
    savedResultsSurvivalGroup(groupi+1,indivi)=survGroup;
    savedResultsSurvivalIndividual(groupi+1,indivi)=survIndiv;
  end
end
dlmwrite('[FOLDERNAME]/res/saveResultsGroupHGGvsHGG-GHG.txt',
   savedResultsSurvivalGroup,'delimiter','\t','precision',8)
dlmwrite('[FOLDERNAME]/res/saveResultsIndivHGGvsHGG-GHG.txt',
    savedResultsSurvivalIndividual,'delimiter','\t','precision'
        ,8)
toc
```

FigureE17 presents the results of the new analysis. Again, there is a consistent advantage for the focus individual to place the H-action at the beginning of the sequence. The advantage again grows with the number of swaps necessary to change the sequences into each other.

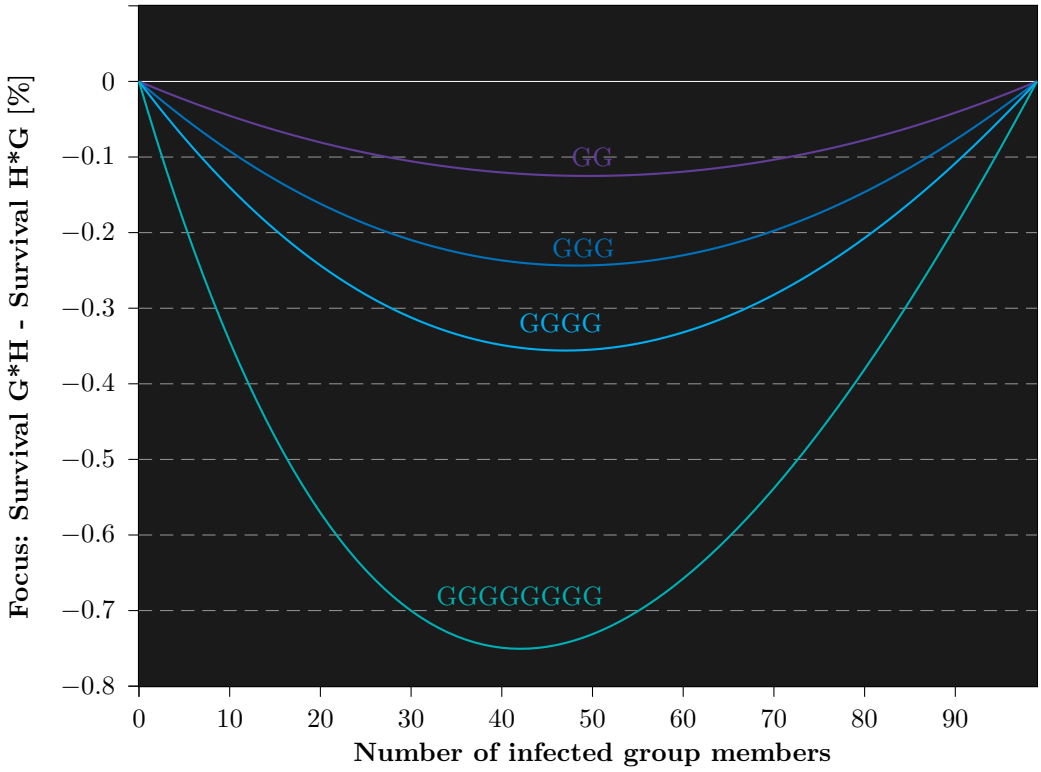


**Figure E17**

*Difference in survival rates for focus individuals with strategies starting with a single H action or ending with it: Games last two (GG), three (GGG), four (GGGG), or eight (GGGGGGGG) rounds and the difference in average survival rates is plotted for each possible group infection state in the first round of the game. Group members play G-actions in every round. Negative values indicate that the H*G strategy leads to a higher survival rate for the focus individual than the G*H strategy.*

### *E6.3.2 Varying group behavior*

Yet, these results are not sufficient to deduce that HG-patterns always lead to higher survival rates than alternative patterns with the same number of H-actions, as all simulations assumed G-actions by group members in all rounds. To test the generalizability of the results we calculate the differences between the focus individual choosing an HG- versus GH-pattern, while varying the group behavior across all possible choice-patterns for the group members: GG, HH, HG, and GH. Figure E18 presents the results for this analysis.

For three of the four conditions,the previous result hold. Naturally, they hold for GG (this was shown before) and they also hold for HH (the difference is even larger). The

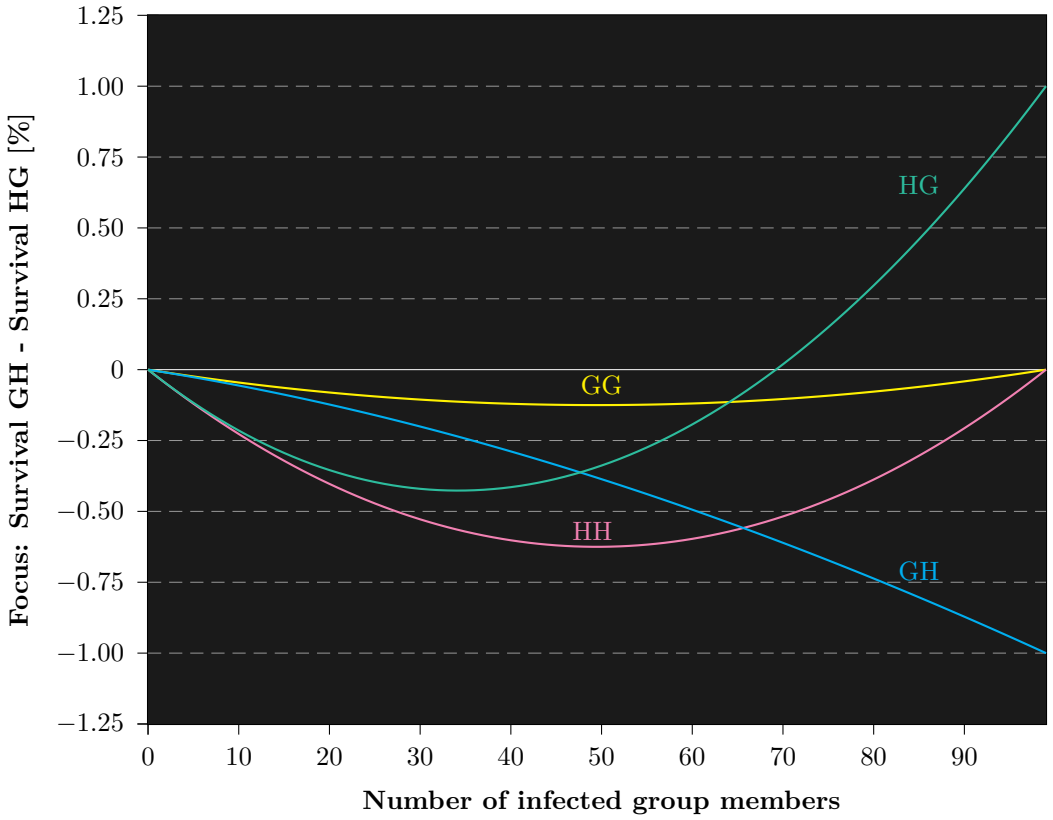


**Figure E18**
*Difference in survival rates for focus individuals with GH- and HG-strategies: Games last two rounds and in four conditions, group members play one of the four possible patterns (GG, HH, HG, and GH). The difference in average survival rates for the focus individual is plotted for each possible group infection state in the first round of the game. Negative values indicate that the HG-strategy leads to a higher survival rate for the focus individual than the GH-strategy, positive value denote the opposite.*

results for HG and GH show a more interesting pattern: With a larger number of infections in the population, an anti-coordination effect becomes stronger: the focus-individual is rewarded for choosing a pattern contrary to the group's behavior. That being said, for any group with an infection rate below 70 out of 99, HG remains the better individual response, even when the group members choose an HG-pattern as well. An infection rate of 70% is rarely reached in games and then only in late rounds of the game. We will constrain our best response analysis therefore to HG-patterns.

### *E6.3.3 Best possible response analysis*

The best possible response analysis is based on the premise that group behavior is in equilibrium only if there is no individual incentive to deviate from it. If the best response for an individual is to copy the population behavior, a Nash equilibrium is identified. As

stated before, we assume that every individual chooses among HG-strategies (even though there might be rare cases, when individuals might benefit to be contrarian in later rounds, but those cases do not seem relevant for competitive strategies). If we call HG-$k$ an HG strategy with $k$ H actions, then there are 26 different HG-$k$-strategies from HG-0 to HG-25. HG-0 and HG-25 are (non-competitive) strategies without switching between actions.

To approach the problem, we use the analytical tools described above to measure individual outcomes for focus individuals that may or may not deviate from group behavior. Specifically, for each relevant HG-$k$-strategy, we observe the relative success of individuals that choose to play strategies in proximity to the group strategy, varying the switching point from $k-3$ if possible to $k+3$ if possible. So, for the HG-12-strategy played by the group individuals are analyzed that play HG-9, HG-10, HG-11, HG-12, HG-13, HG-14, and HG-15. This includes matching the group's behavior at HG-12. The code for this analysis is in Listing 17.

Listing 17: Code for the best-response analysis to universal group behavior.

```
tic
maxPairs=50;
numIndivs=100;
numRounds=25;
saveResultsSurvivalGroup=zeros(200,6);
saveResultsSurvivalIndiv=zeros(200,6);
saveResultsOutcomeGroup=zeros(200,26);
saveResultsOutcomeIndiv=zeros(200,26);
rowNumber=0;
for groupi=0:25
  out=groupi
  for indivi=max(groupi-3,0):min(groupi+3,25)
    outi=indivi
    rowNumber=rowNumber+1;
    probDist=zeros(99+1,2,numRounds+1);
    probDist(8+1,1,1)=92/100;
    probDist(7+1,2,1)=8/100;
    roundHgroup=zeros(25,1);
    roundHfocus=zeros(25,1);
    if groupi>0
      roundHgroup(1:groupi,1)=1;
    end
    if indivi>0
       roundHfocus(1:indivi,1)=1;
    end
    for ri=1:numRounds
       newProbDist=zeros(numIndivs,2);
       for distCond = 1:numIndivs
         if (probDist(distCond,1,ri)+probDist(distCond,2,ri)>0)
           numInfectedGroup=distCond-1;
```

```
            numHealthyGroup=99-numInfectedGroup;
            [probDistMixedPairs,probFocusInPair]
                findProbDistMixedPairsEQ(numHealthyGroup,
                numInfectedGroup,...
                 maxPairs);
            infectionRiskGroup=0.05+roundHgroup(ri,1)*0.2;
            infectionRiskInter=0.05+roundHgroup(ri,1)*0.1+
                roundHfocus(ri,1)*0.1;
            infectDist=calcInfectionDistributionEQv2(
                probDistMixedPairs,probFocusInPair,...
              infectionRiskGroup,infectionRiskInter,maxPairs);
            maxIncrease=min(distCond-1+maxPairs,numIndivs-1)-(
                distCond-1);
            newProbDist(distCond:distCond+maxIncrease,1)=...
              newProbDist(distCond:distCond+maxIncrease,1)+...
                  probDist(distCond,1,ri)*infectDist(1:maxIncrease
                      +1,1);
            newProbDist(distCond:distCond+maxIncrease,2)=...
              newProbDist(distCond:distCond+maxIncrease,2)+...
                  probDist(distCond,1,ri)*infectDist(1:maxIncrease
                      +1,2)+...
                  probDist(distCond,2,ri)*infectDist(1:maxIncrease
                      +1,4);
          end
        end
        probDist (1:numIndivs,1:2,ri+1)= newProbDist;
    end
    finalDist=probDist(1:numIndivs,1:2,numRounds+1);
    groupSum=sum(finalDist,2);
    colSum=sum(finalDist);
    avIndiv=colSum(1,2)*100;
    avGroup=100*([0:99]*groupSum)/99;
    survIndiv=100-avIndiv;
    survGroup=100-avGroup;
    pointsGroup=sum(roundHgroup)*32+200;
    pointsIndiv=sum(roundHfocus)*32+200;
    saveResults(rowNumber,1)=groupi;
    saveResults(rowNumber,2)=indivi;
    saveResults(rowNumber,3)=survGroup;
    saveResults(rowNumber,4)=survIndiv;
    saveResults(rowNumber,5)=survGroup/100*pointsGroup;
    saveResults(rowNumber,6)=survIndiv/100*pointsIndiv;
    dlmwrite('[FOLDERNAME]/saveResults.txt', saveResults,'\t','
        precision',8)
end
```

```
    toc
end
```

To confirm the analytical results and also to validate previous simulation results, we also conducted a simulation covering the same set of combinations of individual and group strategies (with 1,000,000 trials each). The simulation results are presented in Figure E19, the analytical results in Figure E20.

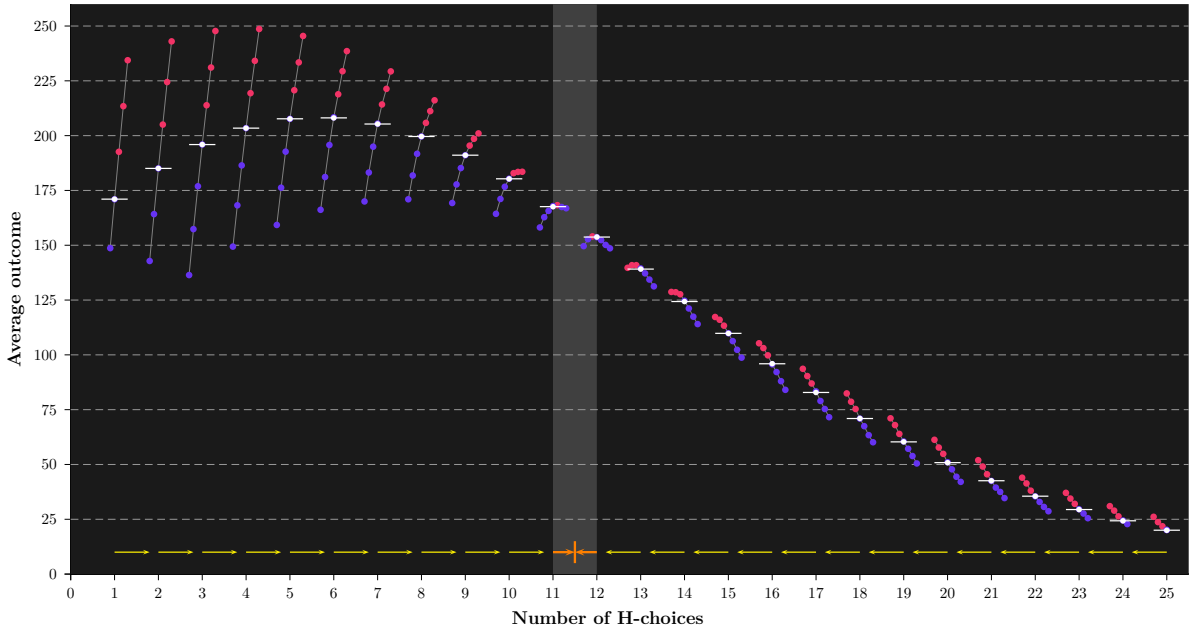


**Figure E19**

*Simulation results for the analysis of HG-strategy combinations: Each point represents the average outcome for a combination of HG-k-strategies for group and focus individual. Points that are linked to the same group strategy are connected by lines with the x-value for the middle point of each group corresponding to the k-value for the group. Focus strategies are varied from $k-3$ $k+3$ where possible from left to right, with the average group result for the strategy match shown as a white line. Individual results that were higher than this group result are shown in red, results that were lower in blue. The yellow arrows at the bottom shows how maximizing individuals would change the group strategy relative to each starting point. The are in lighter shade spans the pair of group strategies that would move groups towards each other.*

Individuals that deviate from the group strategy in the same direction will either consistently perform better or consistently worse than the group in the observed set of results. On the one hand, this means that none of the group strategies can be stable, there will always be an incentive to deviate. At the same time, the pattern of improvements is systematic, and group strategies are driven towards a $k$ value of 11 or 12. There remains a small incentive for individuals to move to a HG-12-strategy if the group plays HG-11 and vice versa, but there is no incentive to choose a different strategy than that. In any case, any possible equilibrium will differ from the collusive optimum in this case.

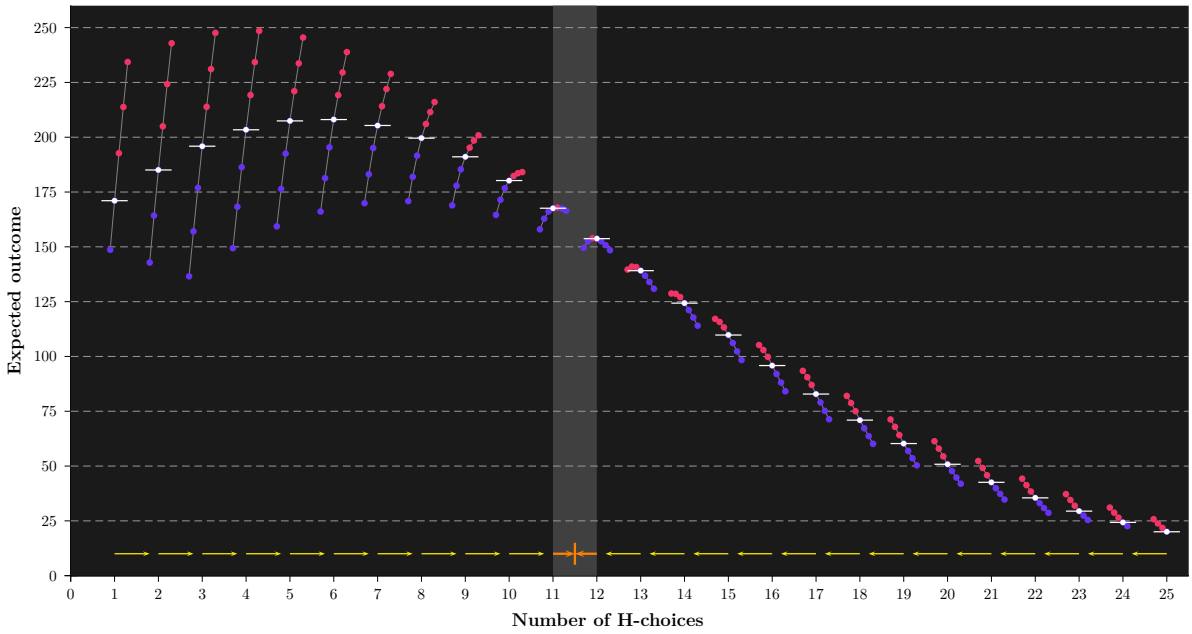


**Figure E20**

*Analytical results corresponding to the simulation results above: deviations are small and unsystematic.*

### *E6.3.4 Considering a mixed-strategy equilibrium*

The pattern of results points at the possibility of a mixed-strategy equilibrium. For a given group strategy, individuals profit from a higher degree of risk-taking, if the level of risk-taking in the group is low. In contrast, lowering the degree of risk-taking is advised for high level of risk-taking in the group. In the first case it seems relatively safe to go for a higher score, in the second case the costs in the form of losing all points through infection are too high. The situation is the same for HG-11 and HG-12, but by shifting single individuals from HG-11 to HG-12, the level of risk-taking in the population can be fine-tuned. If we can identify a state for which individuals can not profit from switching strategies, a mixed-strategy equilibrium will be found. This could be conceptualized as every individual choosing between HG-11 and HG-12 with a fixed probability corresponding to the ideal proportion of HG-11 vs. HG-12-players.

We programmed a simulation, in which individuals play identical strategies in rounds 1–11 and 13-25. Only in round 12, we let two groups of players choose different actions (one G, the other H). We systematically varied the number of players choosing H. Listing 18 shows the top-level code for conducting this simulation.

Listing 18: Code for systematically varying the choice proportions in round 12.

```
tic;
popSize=100;
numberRounds=25;
passiveRisk=0;
```

```
interactionRisks=[0.05,0.15,0.25];
bfixedPerc=1;
initialPerc=8;
bonusG=8;
numTrials=10000000;
simType="SwitchHGnashSpecific";
bconditionalOutcomes=1;
bSingleSave=0;
bMeanSave=1;
bStandardDeviationSave=1;
bPercentileSave=0;
for simParameter =91:99
  disp(int2str(simParameter))
  saveFolder="[FOLDERNAME]";
  saveName=strcat("resNashSpecificNew-", int2str(simParameter));
  [savePopState,saveOutcomeState,saveSumOutcomes,...
    saveInfectionRate,saveActionRate,saveConditionalOutcomes]=...
    transmissionPlaySimulationBasev03_07sep21(...
   popSize,numberRounds,passiveRisk,interactionRisks,bfixedPerc
      ,...
   initialPerc,bonusG,bonusH,numTrials,simType,simParameter,...
   bconditionalOutcomes);
   complete  = saveReducedSimulationResults_v02(...
     saveFolder,saveName,numberRounds,...
     saveSumOutcomes,saveInfectionRate,...
     saveConditionalOutcomes,...
     bconditionalOutcomes,bMeanSave,bStandardDeviationSave,...
     bPercentileSave);
   toc
end
toc
```

One should note that the differences between different variants in this simulation are quite subtle. In a first step we varied the number of H-players in increments of 10. Figure E21 gives an overview. In a first round of simulations, most points seemed to lie on a linear function that would imply that the relative advantage of H-players systematically and linearly decreases with the number of H-players. One point deviated, but in a replication of this condition was located much closes to the line. As the variance in results was too high for identifying the exact cutting point, we increased the number of trials to 10,000,000, and conducted simulations in increments of 1 between 70 and 100.

Again, the results were roughly situated on a straight line. The remaining variance makes it difficult to judge the exact location of the true point in which the line crosses zero, but it is likely to be in the range of 81–83. In other words, if about 82 players choose to play H in the 12th round, and the remaining 18 players play G, none of the players would profit from switching actions uni-laterally. This is a mixed-strategy equilibrium.

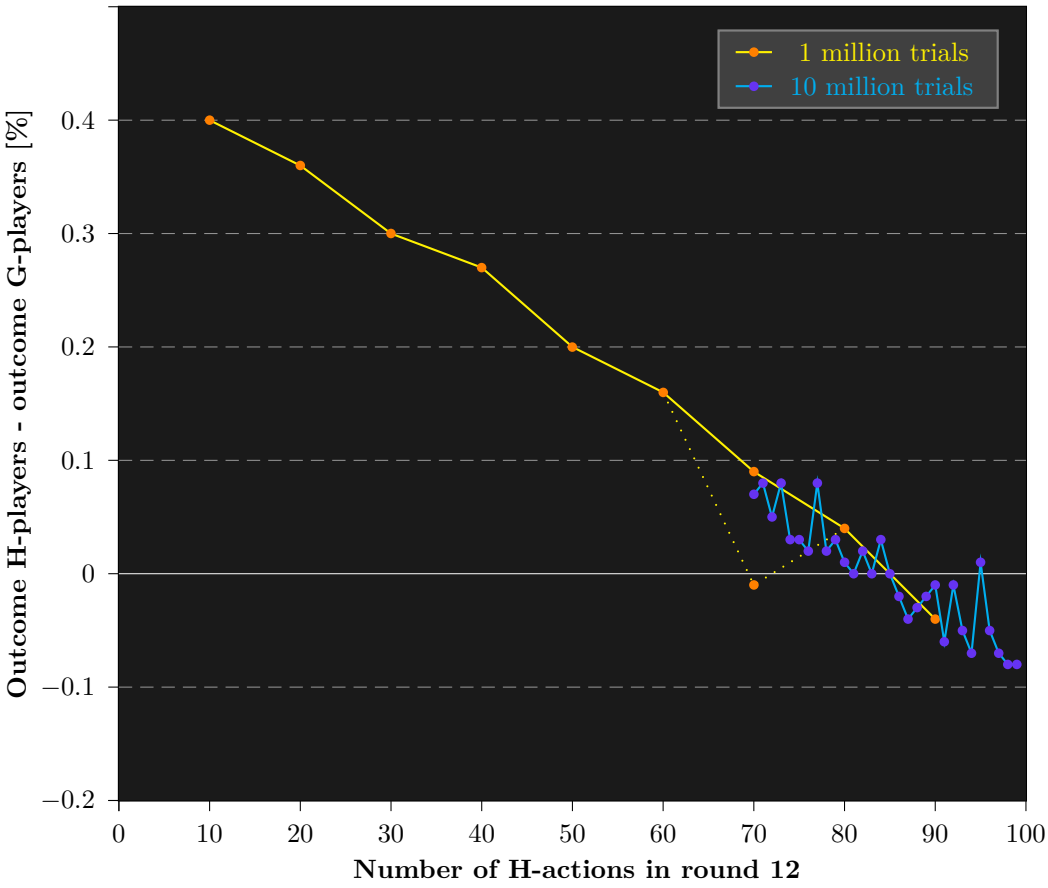


**Figure E21**

*Simulation results for the analysis of different action proportions in round 12: Each point represents the difference in average outcomes between a group of individuals choosing action H and a group of individuals choosing action G in round 12, while all players choose action H in rounds 1–11 and action G in rounds 13–25. A first set of simulations (shown as orange points) with 1,000,000 trials each varied the number of players choosing H in increments of 10. A seeming outlier at 70 (connected with dotted lines) led to a single replication closer to the trend of the other points. For values between 70 and 99, we conducted a second set of simulations with 10,000,000 trials, shown in blue.*

### E6.4 Some caveats in interpreting the results

We stopped short of increasing the trial numbers further to determine the exact cutting point for a reason: The argument can be made that this exact point has little practical relevance, but also little theoretical relevance. To be fair, this analysis uncovered interesting findings about the interplay of strategies, but there is little practical relevance considering how vastly different most participant' strategies were from the discussed strategies.

In addition, there is a theoretical justification why this should not come as a surprise even when assuming rational players with complete insight into the simulation mechanics.

In contrast to many (but certainly not all) economic games, the transmission game's nature is fundamentally stochastic and points can be converted into cash with a given probability that is less than 1. The risk of losing all payoffs increases with the number of points scored in the game (riskier actions are worth more points, after all). The analysis so far considered individuals to be risk-neutral expected-point maximizers, which is an unrealistic assumption for human players.

### *E6.4.1 Introducing a utility function*

The analysis of choices between gambles differentiates between sure gains and stochastic outcomes. From its inception, expected utility theory has implied risk aversion, which is a direct consequence of assuming decreasing marginal gains when it comes to evaluating the value of a given amount of a resource. A flexible group of utility functions (Saha, 1993) is defined as exponential power functions of the form:

$$u(x) = \theta - e^{-\beta \cdot x^{\alpha}} \tag{52}$$

The parameters determine the shape of relative and absolute risk aversion. Here, we combine increasing relative risk aversion and decreasing absolute risk aversion (Arrow, 1971), and we implement this function by choosing the parameters as $\theta = 1$, $\alpha = 0.8$ and $\beta = 0.01$. This function is shown in Figure E22.

As the function value $u(x)$ of an average value $x$ is higher than the average of function values (or $u(x) > \frac{u(x+k)+u(x-k)}{2}$), a sure gain is preferred over equal chances to receive $k$ units more or less. Riskier actions in the transmission game result both in higher expected game scores and a higher risk of losing everything. It was for this exact reason, that our previous analysis of equilibria identified the pareto-optimal set instead of a single solution. Different degrees of risk aversion will thus lead to different solutions.

### *E6.4.2 Repeating the analysis for expected utility*

Our analytical tools allow us to repeat the previous expected-value analysis for expected utilities, assuming the utility function defined above. The results for the new analysis are presented in Figure E23.

In this case, we find a pure-strategy Nash equilibrium at HG-7. The figure also demonstrates that the collusive optimum is shifted towards HG-2. In general, any utility function that implies risk aversion will move the equilibria towards lower number of $k$ (only risk-seeking individuals would prefer higher numbers). An added difficulty in interpreting this solution is that most studies find a good amount of heterogeneity in individuals' risk preferences (Levy, 1994).

In this respect, it is reassuring to find the strong relationship between risk preference (as measured with the general risk item) and game behavior: Participants with higher risk aversion tend to shift their behavior in the normatively appropriate direction.

### *E6.4.3 Other utility-considerations*

Risk preference is certainly only the tip of the ice-berg. Participants have been found to be motivated by considerations of equality and fairness (Fehr & Schmidt, 1999), their differential weighting of outcomes for self and others (Balliet et al., 2009), the goal

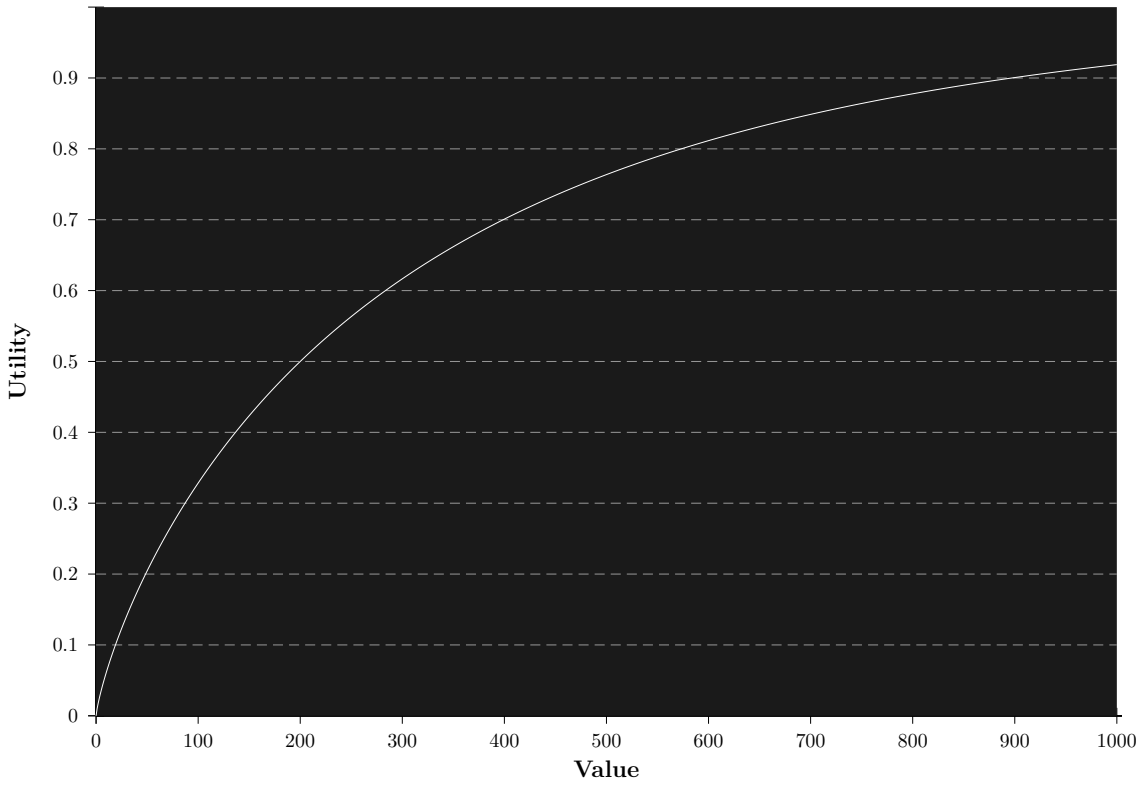


**Figure E22**

*Utility function that transforms objective game outcomes into utilities: the function is the expo-power function with* $\theta = 1$*,* $\alpha = 0.8$ *and* $\beta = 0.01$*.*

of maximizing relative and absolute outcomes (Woike & Hafenbrädl, 2020), feelings or responsibility (De Cremer & Van Lange, 2001), levels of trust in others (Balliet & Van Lange, 2012) and many other factors.

Our results show that social value orientation is linked to participants' choice of strategy, and that cognitive reflection is associated with choosing strategies from the HG-$k$ group. It is for these reasons that non-economists are often skeptical about the value of equilibrium analysis with questionable assumptions about participants' motivation and cognitive abilities. We would argue that the analytical tools offer insight into the mechanics of the simulation and the consequences of action, but in the context of this study see our major contribution in learning more about participants' motivation and willingness to change their behavior through interventions.

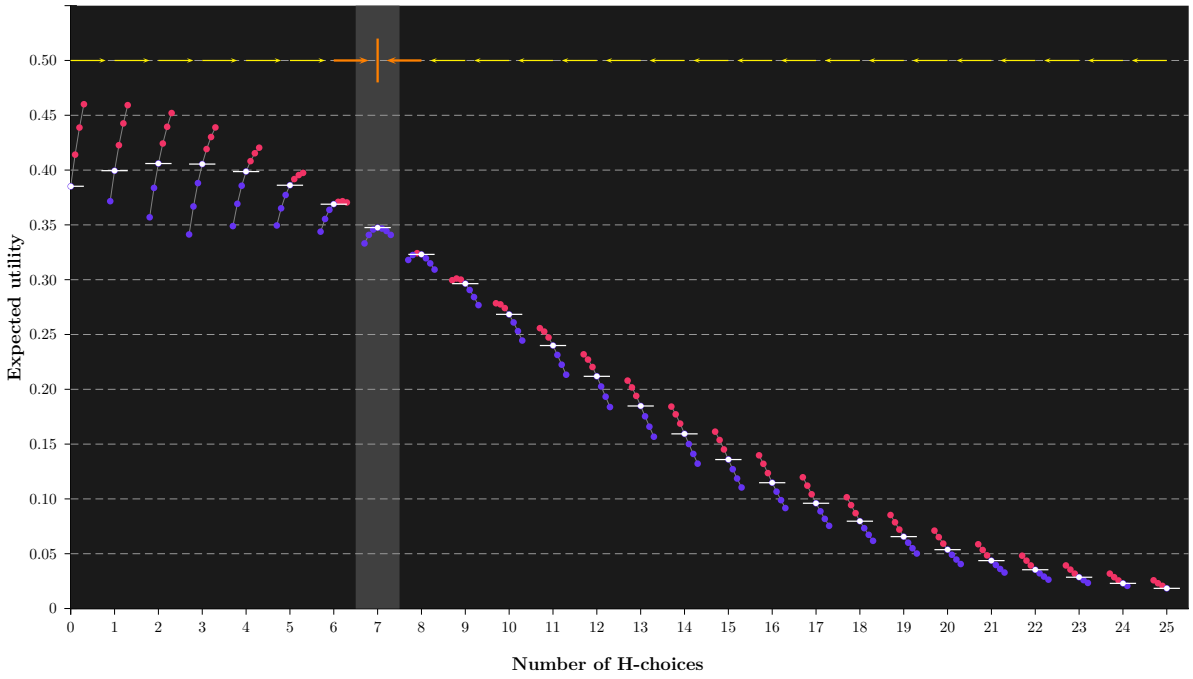


**Figure E23**

*Analytical results for expected utility when combining HG-strategies: Each point represents the expected utility for a combination of HG-k-strategies for group and focus individual. Points that are linked to the same group strategy are connected by lines with the x-value for the middle point of each group corresponding to the k-value for the group. Focus strategies are varied from $k-3$ $k+3$ where possible from left to right, with the average group result for the strategy match shown as a white line. Individual results that were higher than this group result are shown in red, results that were lower in blue. The yellow arrows at the bottom shows how maximizing individuals would change the group strategy relative to each starting point. The are in lighter shade marks the point, where the focus individual does not profit from deviating from the group strategy.*

# Part III
# Supplementary Materials for Study 1

## E7 Study 1: Selected Materials

*Note: The full set of materials is found in document **SM1** (Study Materials Study 1, see Section E1).*

### E7.1 Transmission game instructions

Here, we omit the comprehension checks that had to be passed to proceed from one page to the next (16 checks in total, see **SM1**).

#### *E7.1.1 Page 1/5*

[*See Figure E24.*]

In the next part, you will play a game with 99 other participants (there are 100 players in this game, including you).
The other participants might not play the game at the same time as you do, so you will not have to wait for others.
Rest assured that the other players are other participants on Prolific. We never deceive participants in our instructions or invent participants who do not exist.
We will determine the results of the game, after we have collected the responses of all players.

Next, we will explain the rules of the game.

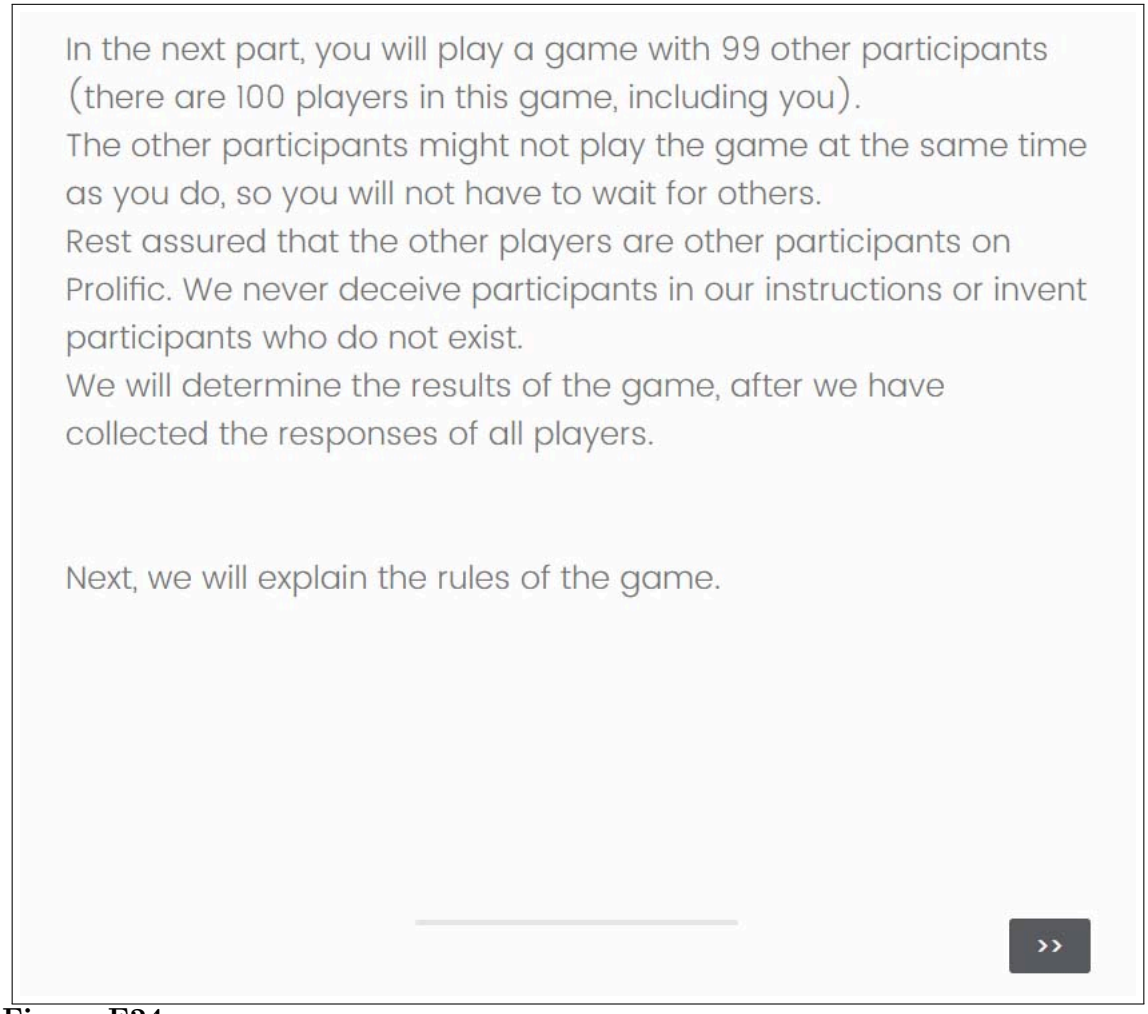


**Figure E24**
*Game Instructions Page 1*

### *E7.1.2 Page 2/5*

[*See Figure E25.*]

The game lasts 25 rounds.

In each round you will make one decision. Your decisions will affect your own bonus payment and can affect other players' bonus payments.

Every player has a color, either blue or purple.
You and every other player all start as blue players.

Some players may change from blue to purple in the course of the game (we explain below how this can happen).
Once a player has changed to purple, they stay purple for the remainder of the game.

It is not possible to change back to blue.

Like all the other players, you will only know your color at the beginning (blue).
Once the game has started, and throughout the game, nobody will know their own color, nor will they know any other player's color.

In the course of the game, you will earn points.

At the end of the game, blue players will receive a bonus payment based on the number of points they have earned across all rounds.
Each point is worth a fixed amount of bonus money for a blue player.

In contrast, purple players will NOT receive a bonus payment for this game. Their points are not converted into money at the end of the game.

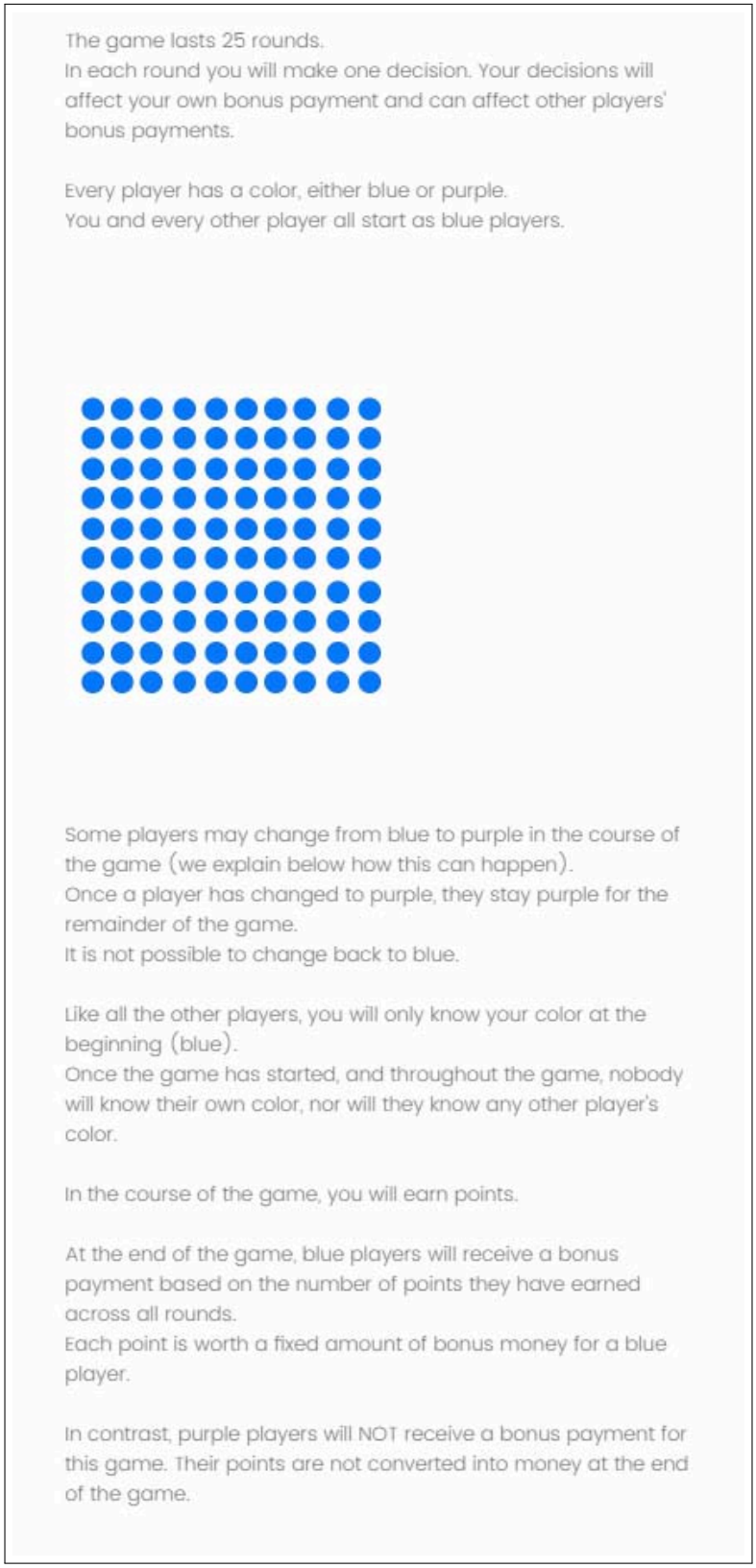


**Figure E25**
*Game Instructions Page 2*

### *E7.1.3 Page 3/5*

[*See Figure E26.*]

At the start of the game, just before the first round, eight randomly chosen players will change from blue to purple.
No player will know whether they have changed color or not: After the start of the game, no player knows their color.

In each round of the game, each player will be paired with one of the other 99 players (there is an equal chance to be paired with any one of them).
Both will decide on one action: either G or H. These decisions will affect your bonus payment.

A player who chooses action G will receive 8 points.
A player who chooses action H will receive 40 points.

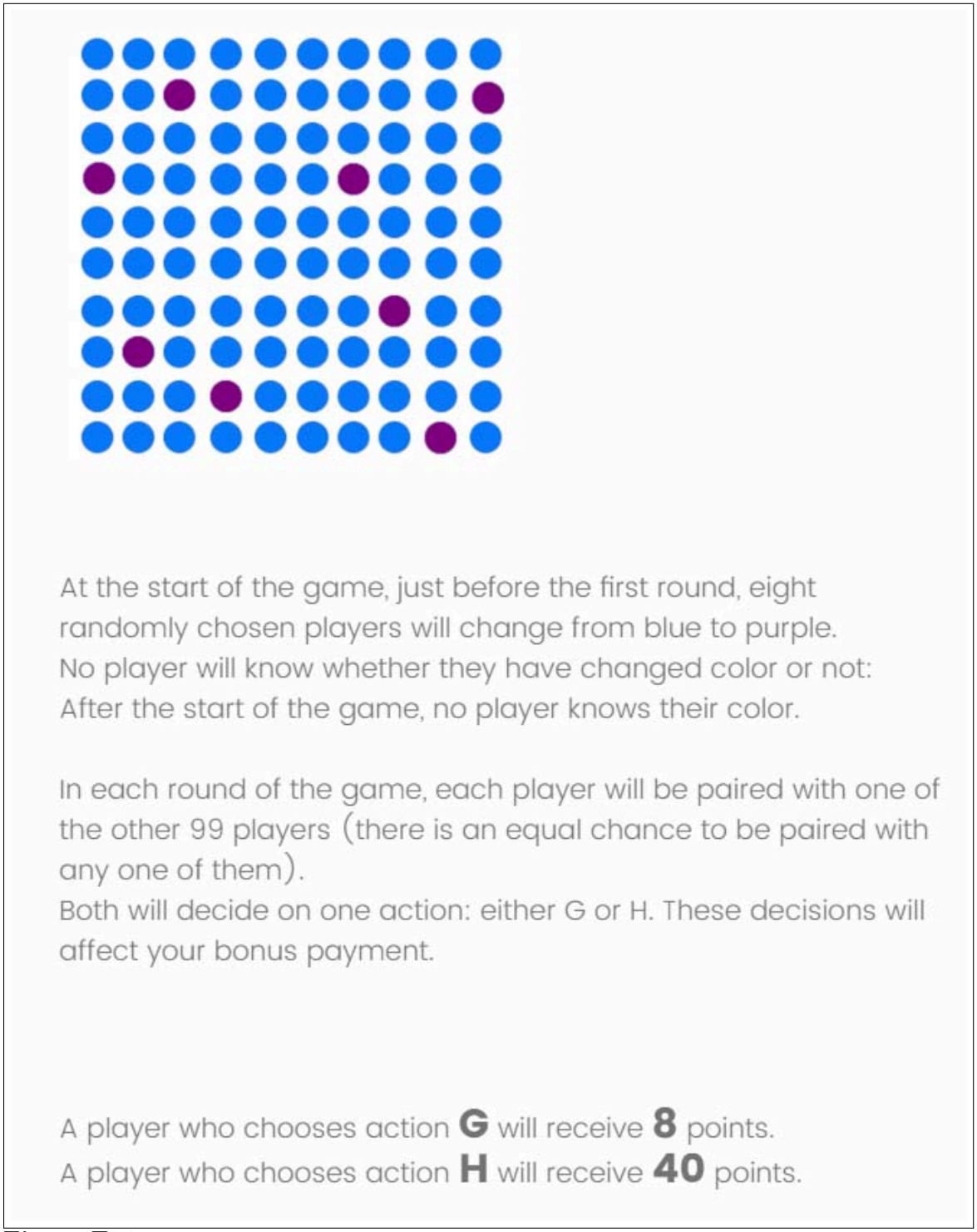


**Figure E26**
*Game Instructions Page 3*

### *E7.1.4 Page 4/5*

[*See Figure E27.*]

If two blue players or two purple players are paired, neither player will change color in this round.

What happens if a blue player is paired with a purple player?
In this case, the blue player may change into a purple player.
The probability of this change depends on the actions that both players choose in this round.

If both choose action G (8 points), there is a 5% chance that the blue player changes into a purple player in that round.
This means, 1 in 20 blue players would change on average in this situation, 19 of 20 would stay blue.

If one of the two players chooses action H (40 points) instead, the chance is increased (by 10%) to 15%.
This means, 3 in 20 blue players would change on average in this situation, 17 of 20 would stay blue.

If both players choose action H, the chance is increased (by another 10%) to 25%.
This means, 5 in 20 blue players would change on average in this situation, 15 of 20 would stay blue.

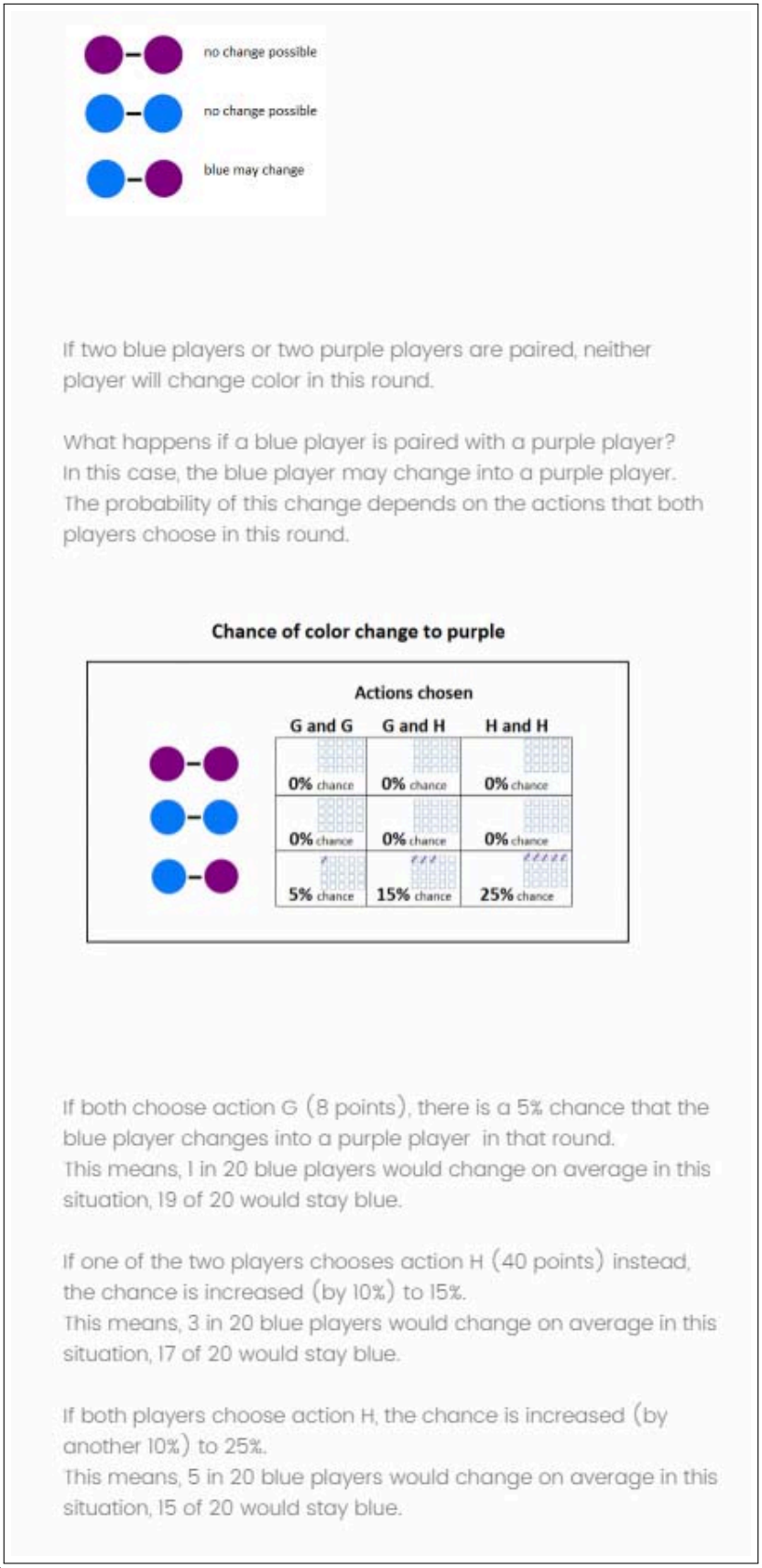


**Figure E27**
*Game Instructions Page 4*

#### *E7.1.5 Page 5/5*

[*See Figure E28.*]

As you know, you can receive 8 points or 40 points in each of the 25 rounds.
How much would you earn per point if you are a blue player at the end of the game?

Points are converted into bonus money at an exchange rate of:
**10 points equal 5p**.

So, 100 points are worth 50p and 200 points are worth £1 of possible bonus payment.
Each choice of action G is potentially worth 4p in the end, each choice of action H is potentially worth 20p.

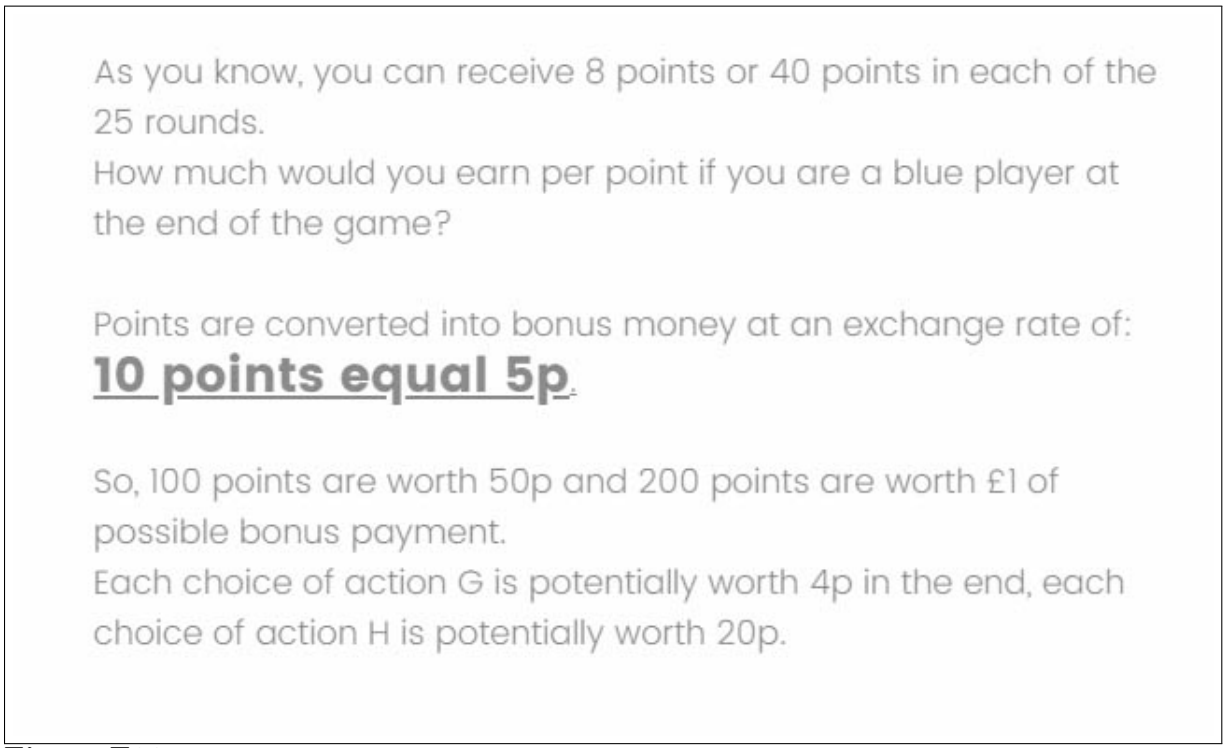


**Figure E28**
*Game Instructions Page 5*

### E7.2 Transmission Game

#### *E7.2.1 Introduction*

[*See Figure E29 and Figure E30.*]

You will now play the game. Note that every other player in your game will have the same information as you have. They will also have seen the same instructions as you have.

This is a brief summary of the rules:

1. **At the start of the game** all 100 players (yourself included) are blue.
2. **Before the first round only**, eight randomly selected players change into purple players.
3. **During each round**:
   - Players are paired at random.
   - Each player chooses an action and receives points (8 points for G, 40 points for H).
   - A blue player that is paired with a purple player may change into a purple player (with a probability between 5% and 25%, see the figure below).
   - A blue player that is paired with a blue player never changes color, no matter which actions are chosen.
4. **After 25 rounds, at the end of the game**, the payoff across all rounds for blue players will be converted into bonus money. Purple players will not receive bonus money for this task.

You will not have to wait for the other players (they may play the game before or after you).
We will collect your decisions and the decisions of 99 other players, then play out the game once to determine the results.
Afterwards, we will inform you about your color at the end of the game in a message when you receive your bonus payment for the entire survey.

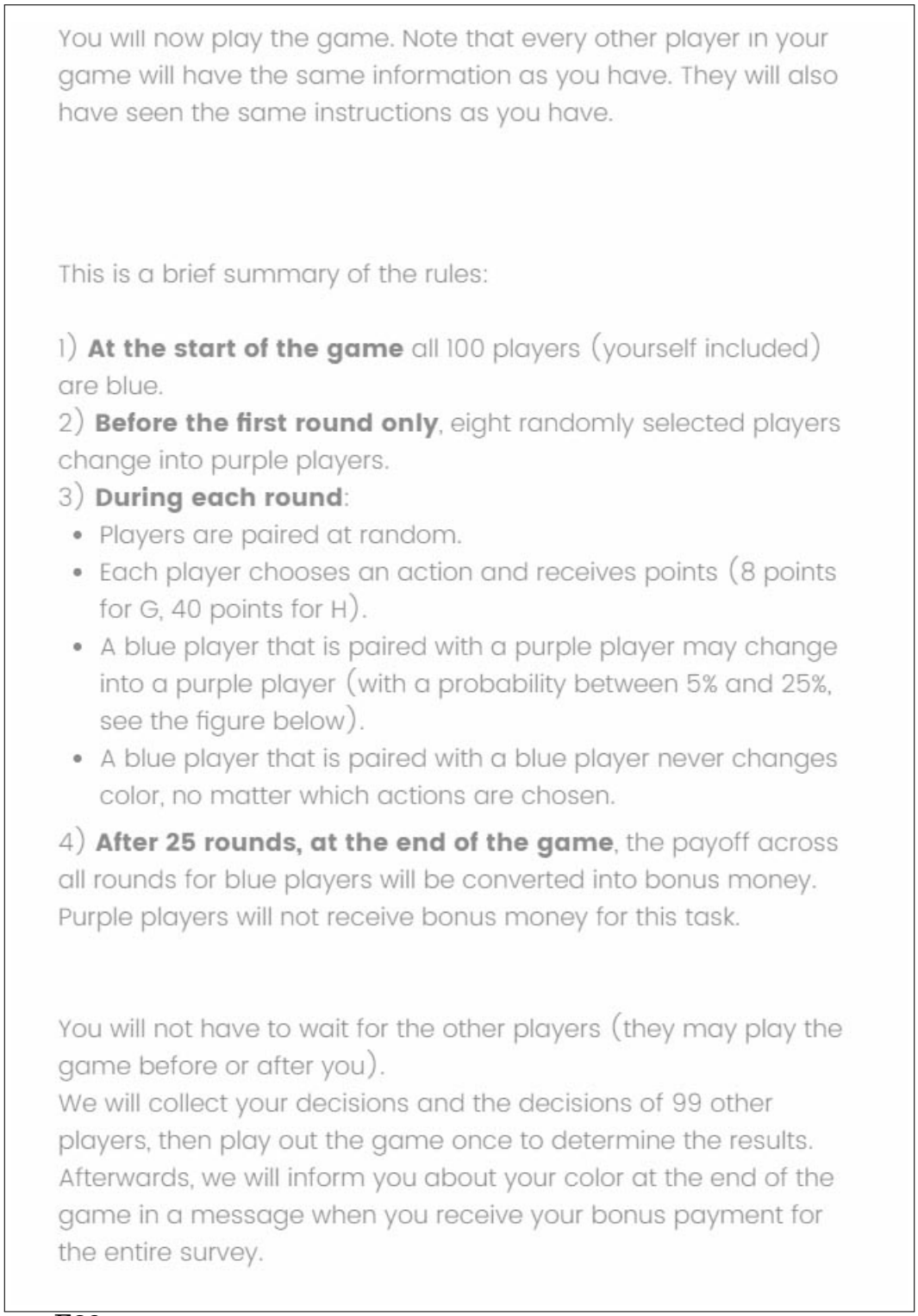


**Figure E29**
*Game Introduction*

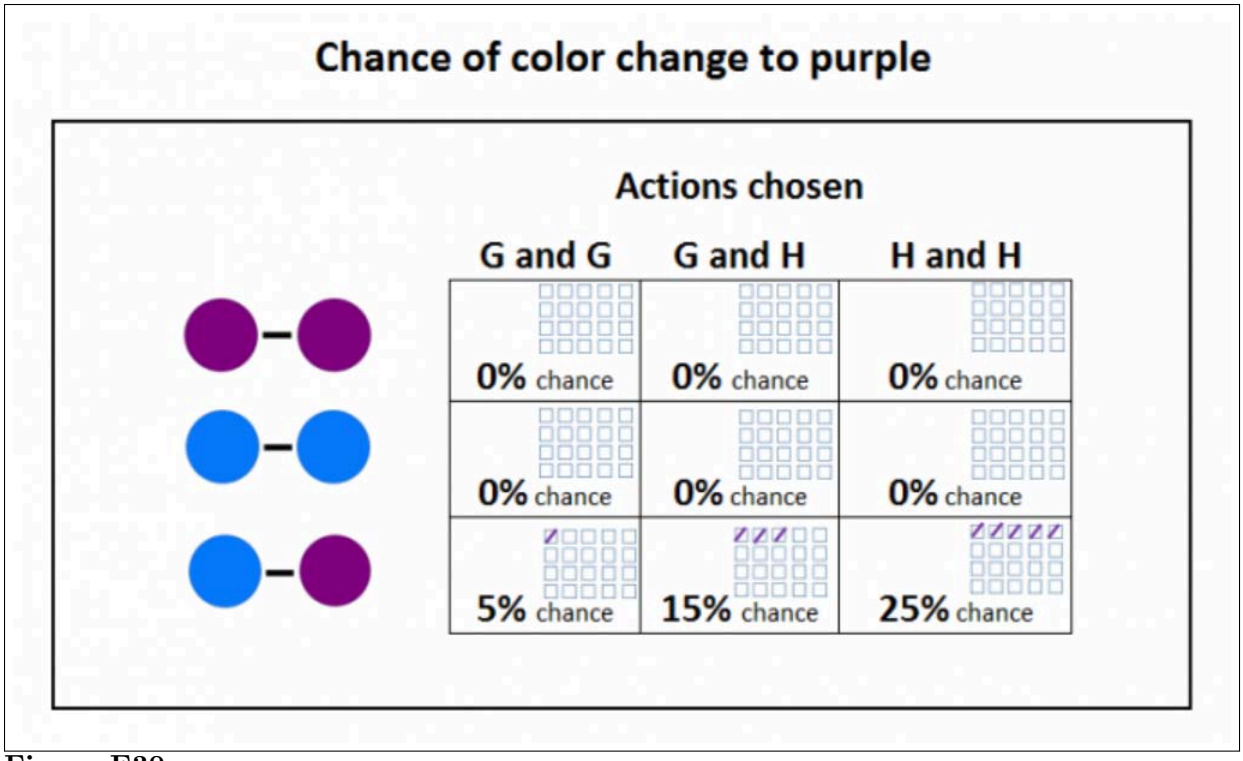


**Figure E30**

*Game matrix (presented below game introduction)*

### *E7.2.2 Rounds*

[*See Figure E31 and Figure E32.*]

Round [*current round*]/25
Points gained so far: [*current number of points*] (+ [*points gained in previous round*])
Possible bonus: £[*current bonus*] (+£[*bonus amount gained in previous round*])
Points will be converted into bonus money after the final round (for blue players only).

Which of the two actions do you choose?

- Action G: I receive 8 points. (8)
- Action H: I receive 40 points. (40)

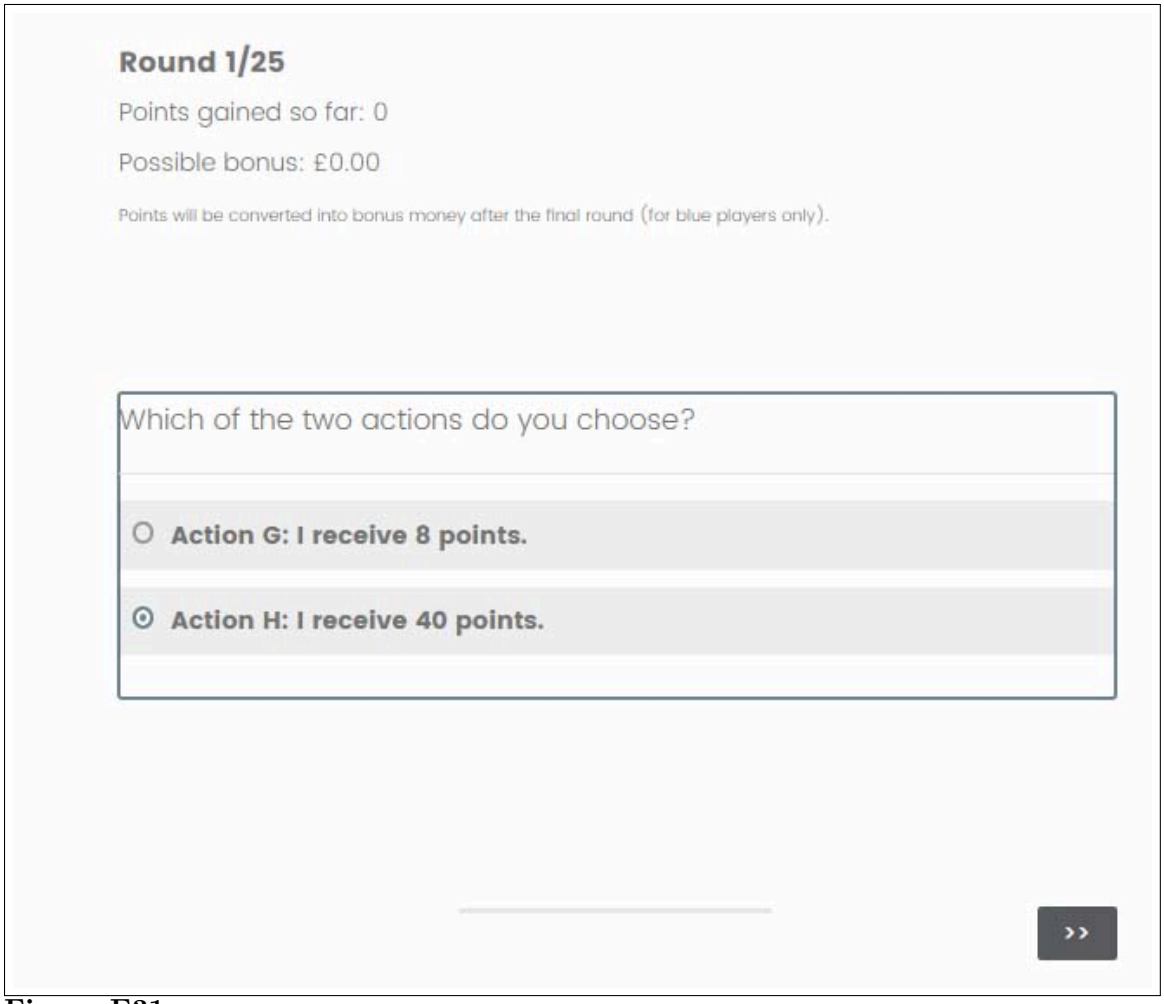


**Figure E31**
*Game round 1 (example)*

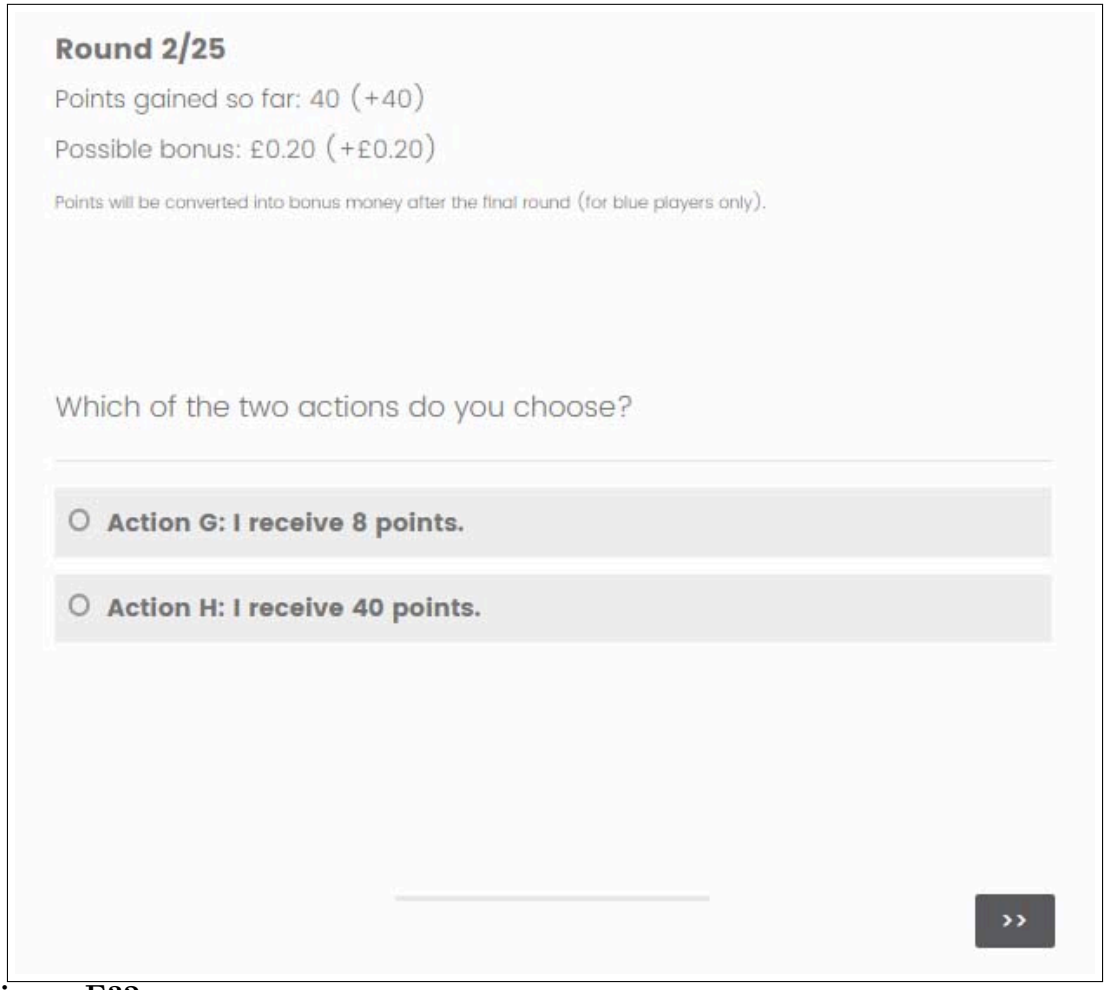


**Figure E32**
*Game round 2 (example)*

### *E7.2.3 Final results*

[*See Figure E33.*]

End of game
Points gained: [*final number of points*]
Possible bonus: £[*final bonus amount*]
Points will be converted into bonus money (for blue players only).

**Figure E33**
*End of game (example)*

### E7.3 Measures and additional variables used in Study 1

A complete description of items and measure statistics can be found in **SM1** (all screen texts and screenshots) and **VM1** (distributions, scale statistics, scale formats, etc.).

#### *E7.3.1 Demographics*

Participants indicated their gender, age, education level, household income in 2019, current employment status, and activity on MTurk (if applicable).

#### *E7.3.2 Cognitive measures*

**Cognitive Reflection Test (CRT).** Participants responded to four items measuring cognitive reflection (Frederick, 2005). Three items were taken from the CRTt (Woike, 2019), the final item from the CRT 2 (Thomson & Oppenheimer, 2016).

**Exponential reasoning.** This scale was developed by the authors and measures the ability to estimate exponential growth.

**Berlin Numeracy Test.** The four numeracy items comprising the Berlin numeracy test (Cokely et al., 2012) were included in the survey.

**Subjective Numeracy.** A very brief, three-item measure of subjective numeracy (McNaughton et al., 2015) was administered.

#### *E7.3.3 Political preferences and attitudes*

**Climate Change Skepticism.** Climate change skepticism was measured with five items taken from Lewandowsky et al. (2015) and one from Lewandowsky, Oberauer, et al. (2013).

**SECS.** The Social and Economic Conservatism Scale (Everett, 2013) consists of 12 items in 2 subscales: social conservatism and economic conservatism.

**Single-item measures.** Participants were asked about their voter registration status, their political position and political party affiliation, and the role of religion in their lives.

### *E7.3.4 Personality scales*

**Brief HEXCACO inventory (BHI).** The 24-item Brief HEXACO Inventory (De Vries, 2013) offers scales for all six personality factors:

1. H: Honesty-Humility (4 items)
2. E: Emotionality (4 items)
3. X: eXtraversion (4 items)
4. A: Agreeableness (4 items)
5. C: Conscientiousness (4 items)
6. O: Openness to Experience (4 items)

**H60-scale.** In addition, participants completed the full Honesty-Humility scale of the 60-item HEXACO-60 scales (based on the HEXACO-PI-R) in the self report form (Ashton & Lee, 2009).

**Psychological reactance scale.** We used a refined, 11-item version of the Hong psychological reactance scale (Hong & Faedda, 1996).

**Oxford Utilitarianism Scale.** This scale is published in Kahane et al. (2018) and contains two subscales: impartial beneficence (OUS-IB, 5 items) and instrumental harm (OUS-IH, 4 items).

**Dark Core of personality.** A 16-item short measure of the dark core of personality (Moshagen et al., 2019) was part of the survey.

### *E7.3.5 Personality single-item measures*

**General Risk Taking.** The general risk question has been used in the German Socio-Economic Panel (SOEP), and has been demonstrated to predict risky behavior inside and outside the laboratory (Dohmen et al., 2011).

**Intertemporal choice.** The time preference measure is used as a "qualitative measure of patience" in Falk et al. (2018, p. 1654).

**Impatience.** This brief measure of patience was part of the German Socio-economic Panel (SOEP) and validated in Vischer et al. (2013).

**Impulsiveness.** This brief measure of patience was part of the German Socio-economic Panel (SOEP) and discussed in Vischer et al. (2013).

**Trust.** The trust question was part of the global preference survey, as described in Falk et al. (2018).

#### *E7.3.6 Economic games and social preferences*

**Social Value Orientation.** Participants completed the six-item version of the ring measure of social value orientation (Murphy et al., 2011).

**Dice game: measure of dishonesty.** Participants completed 15 rounds of a computerized version of the dice game introduced in Kanngiesser et al. (2020). Each dice choice is considered as an item of the scale.

**Dictator Game.** Participants played a simple dictator game (see, e.g., Engel, 2011) and indicated their expectation for other participants' behavior.

**Money Burning Game.** Participants played a money burning game (Zizzo, 2003), in which they could reduce another participant's bonus money at no personal gain. They were also asked about their expectations about other participants' behavior.

#### *E7.3.7 Brief health measures*

**General Health.** A general health question was taken from Centers for Disease Control and Prevention (CDC) (2000).

**Subjective Life Expectancy.** The subjective life expectancy was measured on a slider scale similar to the method used in Mittal et al. (2020).

#### *E7.3.8 COVID-19 related measures*

**COVID-19 related worries.** This 5-item scale measured concerns about the consequences of COVID-19 for self, others and society.

**Compliance with counter-measures.** Three items measured compliance with measures against the spread of the virus.

**Attitude towards social distancing.** Six items measured attitudes towards social distancing.

**COVID-19 related misinformation.** Ten items measured the acceptance of false information about COVID-19. Some of the items were taken from Pennycook, McPhetres, Bago, et al. (2020b). Two other items were adapted from Geldsetzer (2020).

**Willingness to accept tradeoffs.** Six items measured the willingness to accept tradeoffs between health outcomes for the vulnerable population and other outcomes.

**Further measures.** Participants also responded to questions about changes to their job situation—similar to the question used in Dennis et al. (2020)—, about knowing other people testing positive for the virus or having died due to its consequences, and their willingness to get vaccinated following the development of a vaccine considered medically safe. Participants were asked if they had ever been tested for COVID-19 and the result.

#### *E7.3.9 Postquestionnaire*

The postquestionnaire asked for participants' motivation in the game, and asked for estimates concerning the likelihood of ending the game as a purple player, the number of purple players in several selected game rounds, and the number of H-actions in the entire group in selected rounds. This was complemented by several open-format questions, asking for comments on the game, asking for associations with the game experience, for hypothetical messages to other players choosing cooperative or competitive actions, for explanations of chosen strategies and for general comments.

## E8 Study 1: Selected Results

*Note: The full set of results is found in document **SA1** (Supporting Analyses Study 1, see Section E1). Variables and measures are described in document **VM1**.*

### E8.1 Behavior in the transmission game

A sensible statistic for summarizing behavior in the transmission game is the number of risky choices made throughout the game. The number of risky choices, $n_r$, is linearly related to the final score after round 25 ($S_{25} = 8 \cdot 25 + 32 \cdot n_r$). Figure E34 shows the distribution of final scores across the 100 participants, demonstrating considerable variance in the sample.

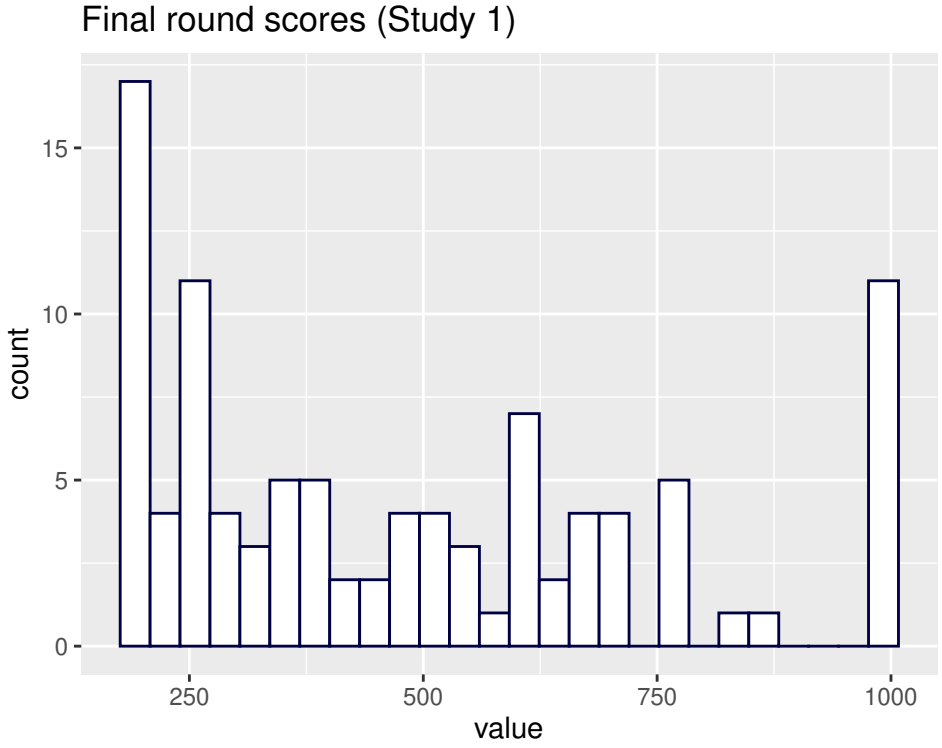


**Figure E34**
*Histogram of scores after round 25 (n=100).*

Figure E35 shows the proportion of H-choices in each round of the game. The majority of participants chooses action H in the first round, and this percentage decreases over the course of the game, yet stays above 25% for each single round.

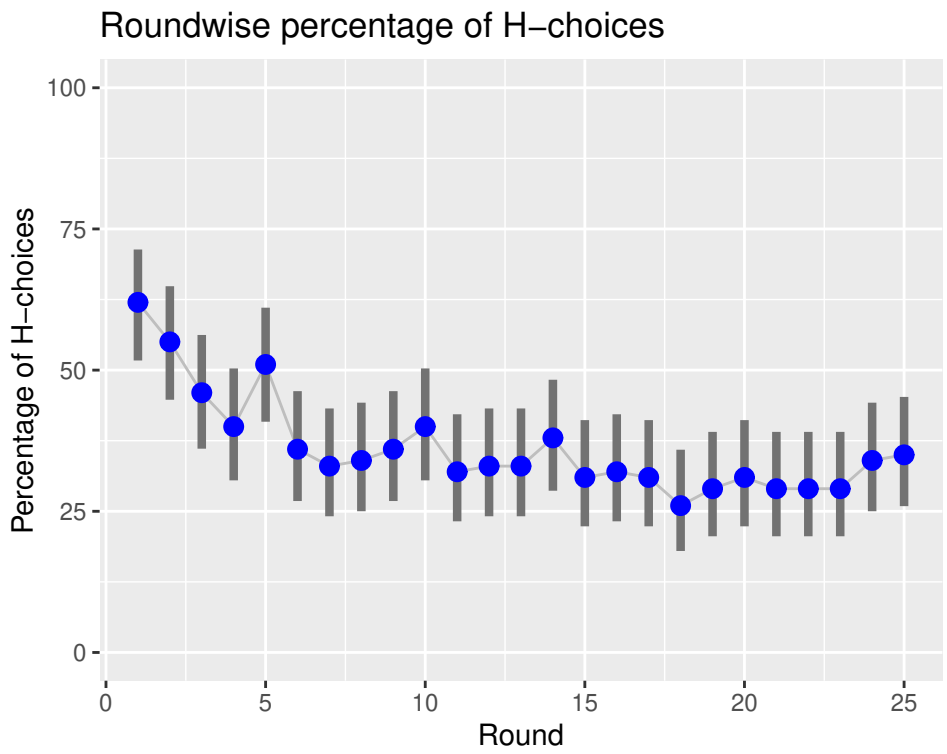


**Figure E35**
*Proportion and 95% CI of proportion of H-choices for each round of the game.*

## E8.2 Postquestionnaire results

### *E8.2.1 Game motivation*

The general postquestionnaire asked for participants' motivation in the game. The postquestionnaire included the following questions (measured on a scale from 1–5, the name at the start of each line is only used in the figure below):

- MaxBonus: I wanted to make as much bonus money as possible.
- MoreBonus: I wanted to make more bonus money than other players.
- Responsible: I felt responsible for other players.
- MakeSwitch: I wanted to make other players switch color.
- DoNotCare: I did not care at all what happened in this game.

The distribution of postquestionnaire responses and their correlation with the game score are summarized in Figure E36. Most participants were focused on achieving the maximum bonus possible, but less so on achieving more than other participants. There was substantial variance in the feeling of responsibility for other participants, but few participants wanted to make other participants switch color or indicated that they did not care about the game result.

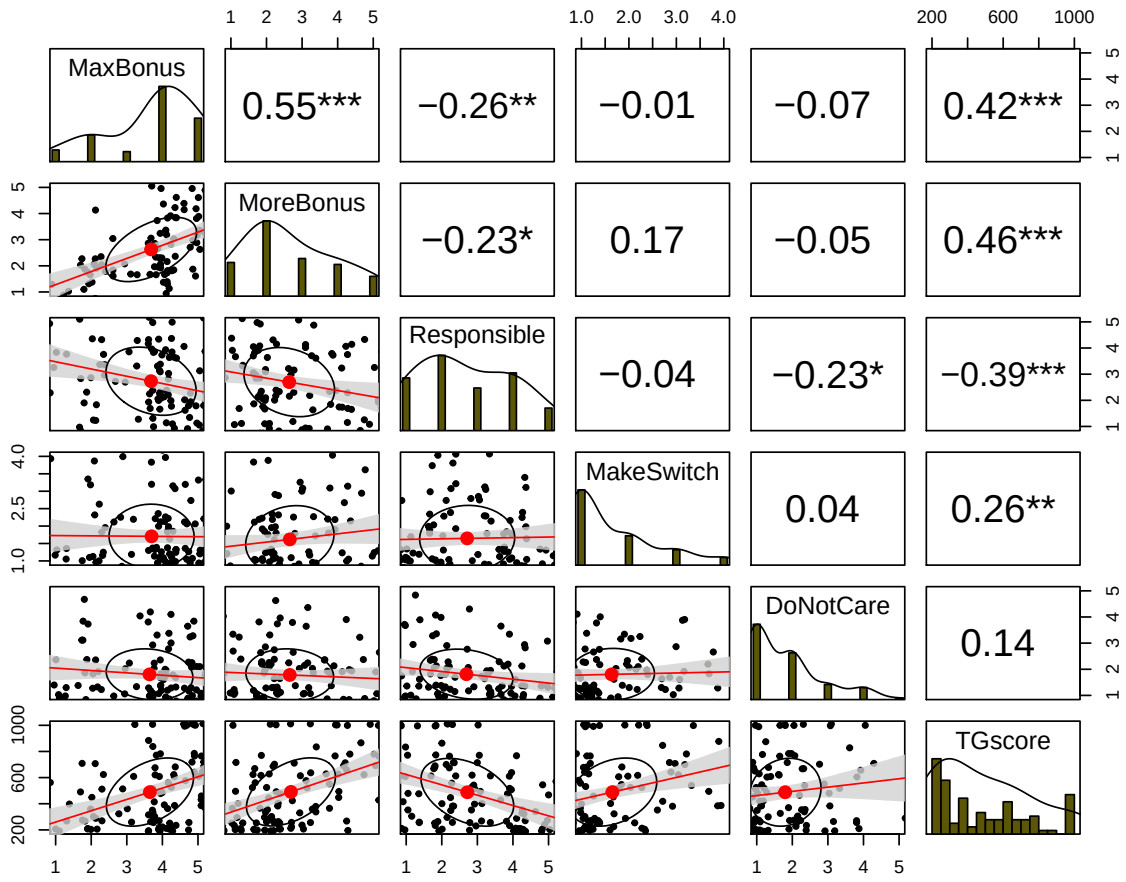


**Figure E36**
*Histograms of postquestionnaire-responses and game score with scatterplots and correlation coefficients.*

The indicated goals were clearly consistent with the observed game behavior. Participants who indicated that they wanted the maximum bonus and those who wanted to have more than others tended to have higher game scores. In contrast, those feeling responsible for others chose the H-action less often. A somewhat weaker relationship is observed between the intention to make others switch and H-actions, and only a weak correlation is seen between disinterest in game results and the game score.

### *E8.2.2 Strategy descriptions*

The distribution of responses (see Figure E34) shows that a substantial number of participants consistently chose the same strategy throughout the game, *AllG* or *AllH*. Asked about their strategy when playing the game, a typical response for an *AllG*-player was "I stuck with the safest option always, for fear of turning purple". The focus for *AllH*-players was different, for example: "I always chose H because it was the highest potential payoff".

Players choosing each action (G and H) at least once fell into two distinct categories. Some participants indicated that they wanted to switch between options to achieve a balanced outcome. In deciding when to switch they did not offer clear guiding principles (*Switch*-players). One typical explanation was: "I switched between picking H and G to even out my playing field, I wanted the most bonus money but also didnt want to end up being purple so I took risks when needed". A more sophisticated strategy involved choosing a single switching point, after starting the game with one or more H-actions (*HG*-players). We show an independent justification for this strategy in our simulations, but several participants clearly expressed their intuition: "I picked the risky actions at the start of the game in order to built up more points as there is a lower amount of purple players but as the game continue it was safer to pick the less risky options (8 pt as opposed to 40 pt) since its likely that the number of purple players would increase as well. Due to the risk of turning purple, it was more important to me to aim for lower pts than to risk losing all pts."

More of these explanations are compiled in **VM1**. In general, strategies seemed to correspond well with the objective game features, validating that our instructions succeeded in conveying the game rules.

#### *E8.2.3 Further qualitative data*

See document **VM1** for a selection of responses regarding associations with the transmission game in open-format answers. The majority of participants (about 80%) did not associate the game with aspects of COVID-19. Some participants offered analogies in the intended direction: "COVID-19 and Social Distancing. In a blue and purple pairing, if both choose G there is less likely to be a spread of purple. If one chooses G and the other H, then theres still a chance of spread but less of it. If both choose H, the risk is highest."

When asked to send hypothetical messages to other players choosing either action G or H, participants differed in their response to both groups. Some participants were grateful to G-players ("Thank you for choosing the option that benefi[t]ed everybody!") and berated H-players for being either greedy, reckless, or at least "bold". Others were more neutral and did not send negative messages. Some players send positive messages to H-players and were puzzled by G-choices ("Are you not worried about money?").

Asked about attitudes about the COVID-19 health crisis, responses reflected arguments from the entire political spectrum, including clear misinformation. Answers also showed a high degree of uncertainty and experienced ambiguity ("I started out very afraid and over time feel more numb about it. Its hard to know whats true with so many opposing views flooding the media. I cannot stay home forever. We wear masks, use sanitizer and remain socially distant.").

#### *E8.2.4 Estimates of other participants' behavior*

Participants estimated the percentage/number of purple players for six different rounds (1, 5, 10, 15, 20, 25). Figure E37 summarizes their estimates: Median responses predict a close to linear increase from the initial purple proportion (8%) to over 70% purple players[9].

[9]The outliers indicate players who most likely gave estimates for G-actions.

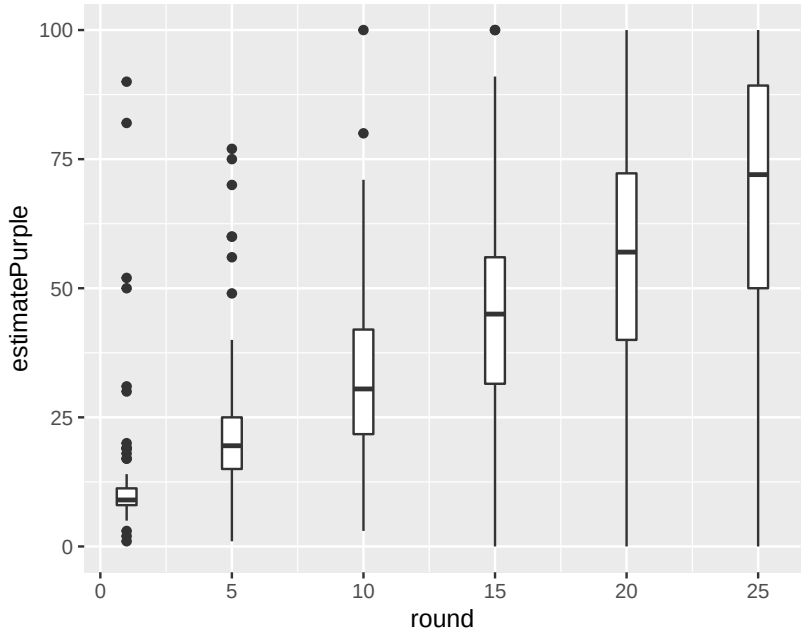


**Figure E37**
*Boxplots of estimates of the number of purple players after rounds 1, 5, 10, 15, 20, and 25.*

Median estimates for the number of players choosing the H-action in each round showed a clearly decreasing trend starting at 80% and ending at about 30% (see Fig. E38). These numbers tended to over-estimate the actually observed round proportions.

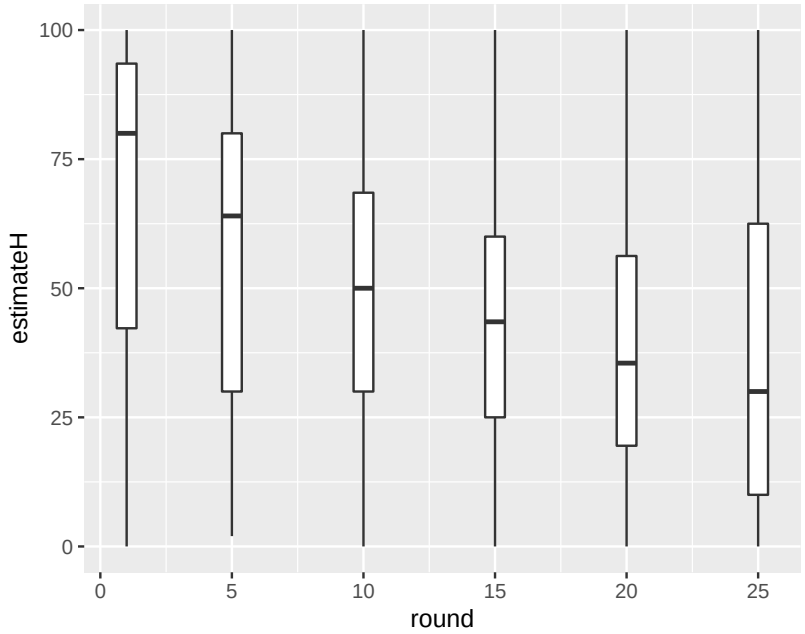


**Figure E38**
*Boxplots of estimates of the number of H-actions in rounds 1, 5, 10, 15, 20, and 25.*

### E8.3 Simulation results based on Study 1 data

We used the simulation code presented in **SCR** to play out the game based on participants' choices. We simulated the game once to determine payoffs (in accordance with our instructions), but we simulated the game 1,000,000 times to generate a solid base for estimating the likely game dynamics.

The distribution of rates of purple players in each round is visualized in Figure E39.

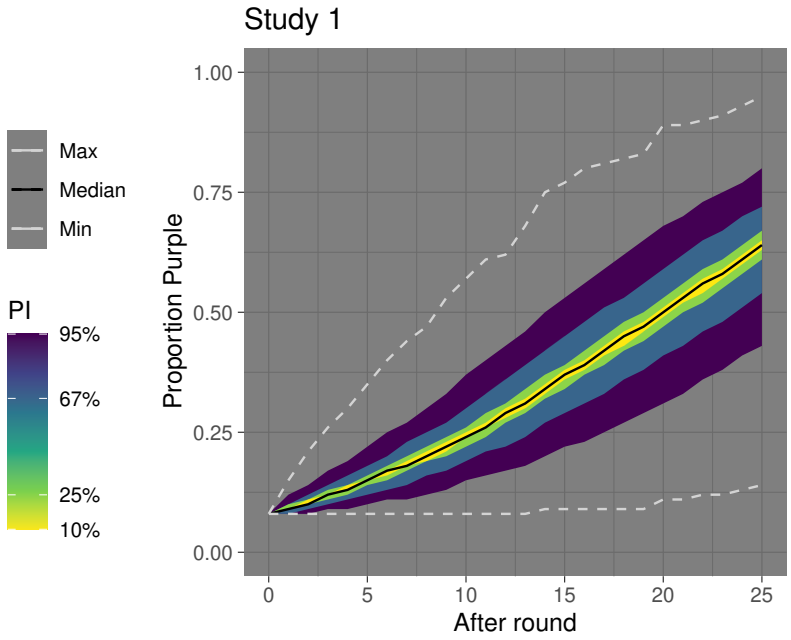


**Figure E39**

*Simulation results for games based on participants' choices in Study 1: The inner line corresponds to the median result across 1,000,000 simulation trials, shaded areas contain the inner 10%, 25%, 67% and 95% of all results.*

In most simulation runs, the majority of all players was purple at the end of the game. The increase in cases accelerated after a slower start, as expected from previous simulations. Comparing the graph with predicted infection rates, participants seemed to have over-estimated actual purple rates. However, considering their over-estimation of H-choices, this may be better interpreted as an underestimation of the consequences of H-actions.

### E8.4 Predictors of game results

The goal of Study 1 was to identify predictors of game results using a wider array of plausible scales and variables. Thus, we advise caution in interpreting these findings without replication (as attempted in Study 2). The range of considered variables is exhaustively presented in **SM1**, **VM1**, and **SA1**. Here, we reduce the complexity by presenting only relationships with variables that we identified as promising in these analyses.

These variables, and their relationship with game results, are presented in Figure E40.

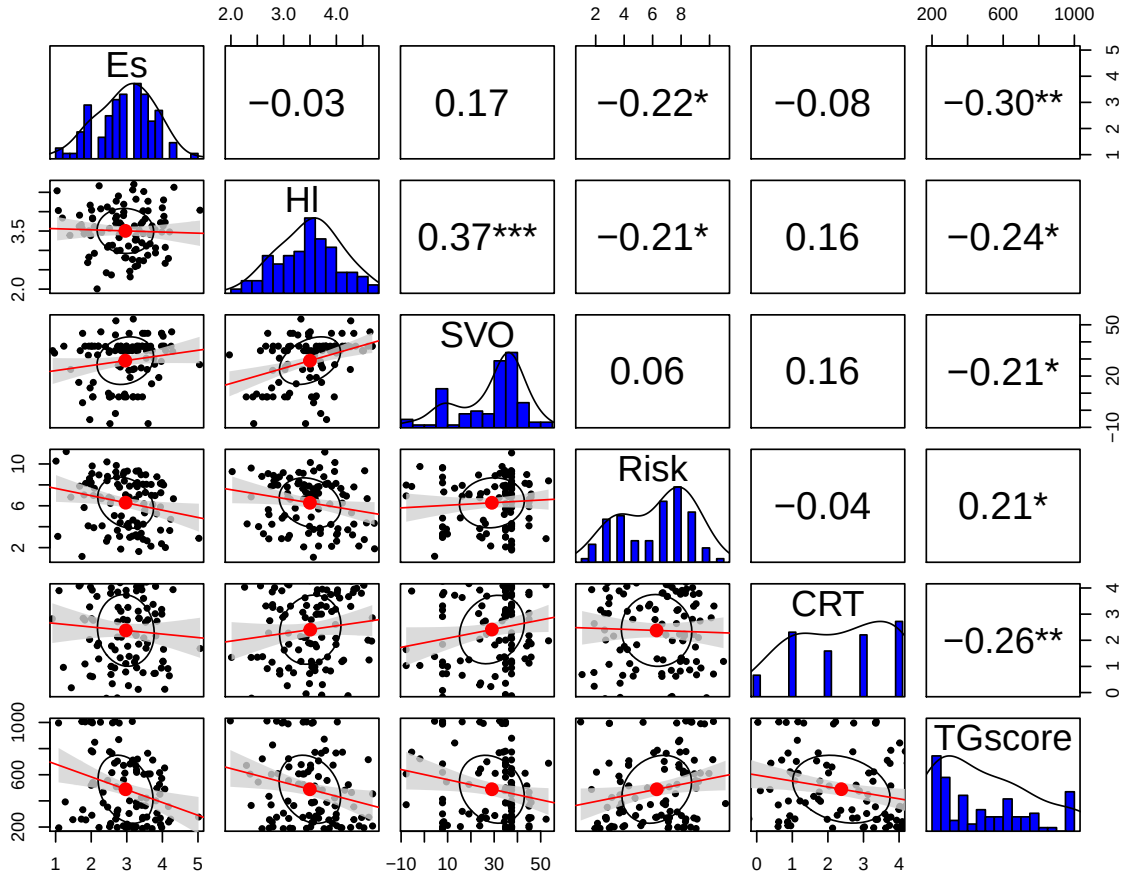


**Figure E40**

*Predictors of game behavior: Histograms, correlations and scatterplots for variables predicting the game score (TGScore), which are, in order, the brief HEXACO-Emotionality scale (Es), the longer HEXACO-Honesty/Humility scale (Hl), the angle computed for social value orientation (SVO, higher values indicate altruistic choices), the general risk item (Risk), and the score in the cognitive reflection test (CRT).*

### E8.5 Political measures and COVID-19

As described in **SA1**, there were low correlations between game scores and responses related to COVID-19. Likewise, there was only a low (positive) correlation between game scores and conservatism. Figure E41 demonstrates that COVID-19 reactions were strongly interlinked with political attitudes.

Conservatism and preference for the incumbent's winning the election were highly correlated with climate skepticism, negative attitudes towards social distancing, endorsement of untrue statements about COVID-19, and acceptance of tradeoffs sacrificing members of vulnerable populations for personally desired outcomes. There was a weaker, yet substantial, negative correlation with the degree of worrying about the consequences of the virus and compliance with counter-measures. In general, most of these measures were highly inter-correlated.

There are further interesting findings about the relationship of COVID-19 scales with economic games and personality variables in **SA1**.

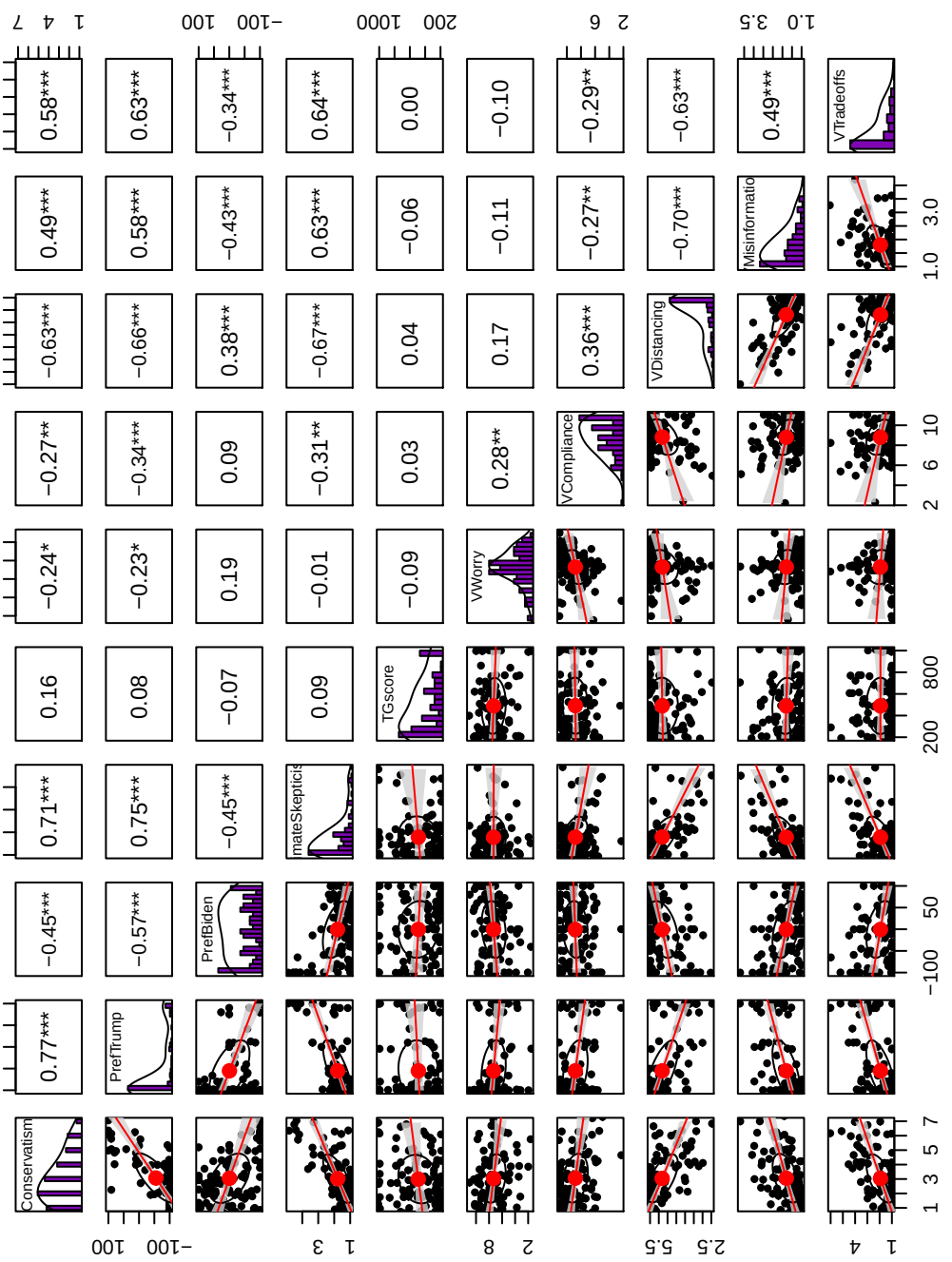


**Figure E41**
*Political attitudes and COVID-19 scales: Histograms, correlations and scatterplots for political attitudes and COVID-19 scales. The variables are Conservatism, expected happiness with President Trump (PrefTrump) and Joe Biden (PrefBiden) winning the coming election, climate skepticism, the game score (TGscore), Worries about COVID-19 (VWorry), compliance with measures against the virus (VCompliance), attitudes towards social distancing (VDistancing), endorsement of objectively false statements about COVID-19 (VMisinformation), and the willingness to accept tradeoffs between human lives and personally desired outcomes (VTradeoffs).*

# Part IV
# Supplementary Materials for Study 2

## E9 Study 2: Interventions

In this section, we give some background information on the choice and design of interventions in Study 2. We implemented five different interventions in Study 2: a vicarious learning condition, two norm-based conditions (descriptive norms and injunctive norms), and two boosting interventions (chain reaction and simulator).

### E9.1 Vicarious Learning

In this condition, participants could learn about the consequences of player decisions by observing the outcomes of a group of previous participants. In particular, in each game round participants saw the number of purple players observed in the corresponding round of a previous game. This is an instance of vicarious learning (Bandura & Walters, 1977): Even though the behavior leading to a growing rate of color change is not disclosed, it is evident that only risky actions (H) can lead to such an increase. Observing the increase of purple players in Study 1 thus serves the function of vicarious punishment (in the sense of observing negative consequences of the model's behavior). A participant who underestimates the consequences of risky actions could adjust their expectations based on these numbers.

The beginning of the health crisis, when nations watched each others' performance to learn how best to respond to the virus (Baumgartner et al., 2020; Gatto et al., 2020; Liang et al., 2020), provided opportunities for learning from vicarious consequences. Negative examples of outbreaks served as cautionary tales to inhibit future risk-taking, and people became gradually aware of the risks posed by COVID-19 (Wise et al., 2020). Graphs of rising case numbers are commonly featured in print news, television news, on online forums, etc. We decided to confront participants in Study 2 with the progressive increase of purple participants in Study 1. To maintain independence between players within a group and still keep the design deception-free, we informed participants about the source of the data.

We presented a typical, step-wise increase of purple players based on simulations of games using the player actions in Study 1. This information was presented as a graph (see section E10 for screenshots of all conditions). We chose a linear axis for better comprehensibility (Romano et al., 2020), and also included a sentence describing the state of the game together with tooltips displaying numbers for each data point (compare Fansher et al., 2020).

### E9.2 Norm-based interventions

Two of our interventions were norm-based interventions. We follow the distinction made by Cialdini and Goldstein (2004) between norms that inform about behavior others typically engage in (descriptive norms) and norms that inform about behavior others typically approve or disapprove of (injunctive norms; Bicchieri & Dimant, 2019; Cialdini et al., 1990; Schultz et al., 2007).

#### *E9.2.1 Descriptive norms*

The first norm-intervention provided participants with descriptive norms based on actions chosen in Study 1. In each round, participants read information about how many participants had chosen action G in that round in Study 1. If a participant who chooses the risky option believes that everyone behaves like them, this will be easily disconfirmed by this information. Instead, starting from round 2, the majority of participants in Study 1 chose the safe and cooperative action. This type of interventions has been used in attempts to convince people that a majority of others follow social distancing rules (Kitamura & Yamada, 2020; Leary & Dvorak, 2020; Lees et al., 2020; Van Bavel et al., 2020), but there are many examples of applications well before the current health crisis. Descriptive-norm interventions have been applied to handwashing behavior (Dickie et al., 2018), punctuality of paying taxes (Halpern, 2015), using stairs instead of elevators (Burger & Shelton, 2011), reducing meat consumption (Sparkman & Walton, 2017) or college student drinking (Huh et al., 2015; Lewis & Neighbors, 2006a, 2006b; Whitt et al., 2019).

#### *E9.2.2 Injunctive norms*

The second norm-intervention focused on judgments of appropriate behavior. Injunctive-norm interventions go beyond simple instructions of what to do by establishing morally relevant reasons for some behavior and against a different behavior. For example, an intervention by (Cialdini, 2003) asked park visitors not to remove petrified wood from the park in order to preserve its natural state. Preservation adds a moral connotation to an act of removing elements from the environment. Similarly, researchers have utilised injunctive-norm interventions to increase hand-washing hygiene among health care workers (Grant & Hofmann, 2011) by connecting a lack of hand hygiene to patient harm.

Government messages to the public often invoke injunctive norms. For example, the National Health Service (NHS) posters in the UK at the beginning of the crisis used the tripartite instruction: "Stay Home—Protect the NHS—Save lives". This slogan connects the decisions available to the decision-maker (i.e., staying home or going out) to beneficial outcomes with an increasingly moral evaluation. In comparison, the decision not to stay home (and therefore not to protect the NHS and not to help save lives) appears clearly inappropriate. Perceiving a moral basis in one's behavioral choices has reinforced behavior in accordance with these choices in Luttrell et al. (2016) and Seale et al. (2020).

In response to the current health crisis, several injunctive-norm interventions have been implemented by researchers to test variants of messages for increasing protective behaviors. These studies haven often taken the form of surveys and used a change in stated intentions to do (or not to do) something to evaluate the success of the intervention (Capraro & Barcelo, 2020; Everett et al., 2020; Favero & Pedersen, 2020; Jordan et al., 2020; Pfattheicher et al., 2020) or reported compliance with protective measures (Gouin, 2020; Kuiper et al., 2020). Our paradigm allows to test the effect of injunctive norms on incentivized behavior.

We distilled some of the findings in theoretical and survey studies (Everett et al., 2020; Globig et al., 2020; Jordan et al., 2020; Kuiper et al., 2020; Michie et al., 2020; Moore et al., 2020; Pfattheicher et al., 2020; Seale et al., 2020; Yoeli & Rand, 2020) to create an intervention that presents G-choices as an unambiguously appropriate behavior, arguing

that G choices avoid harms to both players themselves and other players. In addition to providing these explanations at the start of the game, we summarized them in a single sentence ("Choose action G to protect your and other players' bonus money."), which we prominently displayed close to the action options.

### E9.3 Boosting interventions

Finally, we added two boosting interventions in Study 2. Recently, behavioral science has moved into the focus of policy makers' attention for successfully demonstrating the concept of "nudging" (Thaler & Sunstein, 2009): By making use of assumed biases in decision making or exploiting motivational inconsistencies nudges change decision making situations to achieve desired choices. Importantly, decision makers should not perceive nudges as choice restrictions or sanctions, but as a gentle push that they might not even notice (Halpern, 2015). Nudges try to make decision easy, the desired choice attractive and use social influence to steer decisions. Educative nudges try to achieve outcomes assumed to be in the true interest of the decision maker and help to counter-act self-defeating impulses (Hertwig, 2017).

In contrast, boosting interventions deal with identified weaknesses in a different way. Instead of exploiting them, boosts aim to remediate them by strengthening decision-making competence and assisting individuals in making good, autonomous decision (Hertwig & Ryall, 2020). Boosting interventions offer consciously experienced decision aids (see also Lunn et al., 2020). Our boosting interventions do not take the form of fear-inducing appeals (Bauch & Earn, 2004; Wood, 2000), but target possible sources of misunderstanding the game situation so that decision-makers are empowered to align their choices with their preferences. We implemented two boosting interventions (chain reaction and simulator).

Our two boosting interventions aimed at improving the ability of participants to extrapolate the exponential spread of the transmission. A recent study found that people implicitly assumed linear growth of virus transmission (Lammers et al., 2020), which can result in either underestimating future prevalence based on historical data (or overestimating past prevalence based on present data). Underestimating the absolute amount of spread is likely to bias estimates of risk downwards, resulting in objectively riskier choices than the decision maker would make given accurate beliefs (compare de Neys et al., 2020).

#### *E9.3.1 Chain reaction*

We found inspiration for our first intervention in newspaper articles using simple visual demonstrations of non-linear transmission. Roberts (2020b) used a vertical binary tree diagram[10] to simulate a situation in which an infected person infects two people in the succeeding layer. Preventing a single infection on the first level of the tree following the root thus averted the infection of $2^5 = 32$ people on the sixth level. The graphical tree representation makes these quantities easy to understand. Evershed and Ball (2020) used a circular representation of a tree to offer similar support for forecasting the spread, varying the infection rates for three transmission steps. Similarly, Roberts (2020a) demonstrated the impact of a single prevented case over time in a stacked area graph[11]. Our intervention

[10]The diagram is credited to the science graphics editor Jonathan Corum.

[11]This graph is credited to the epidemiologist Britta Jewell.

used a tree format, showing the consequences of an early color change over the rounds of the game. We developed a multilevel tree in steps and thus demonstrated the far-reaching consequences of early risk-taking. The complete intervention is presented in section E10, the JavaScript code is contained in **SM2**.

#### *E9.3.2 Simulator*

Simulations heavily featured in the early scientific response to the epidemic. We discussed epidemiological modelling in section E2. It is not always easy though, to communicate simulation outcomes to the general public. Again, we found several interactive examples in (online) newspaper articles, and some of these simulations were quite involved and complex in their presentation. For example, Evershed and Ball (2020) built on their transmission chain story and offered an outbreak simulation with four continuous parameters ($R_0$, fatality rate, susceptibility, isolation rate). The parameters could be manipulated by the user via sliders and results were shown as a graph, circular population diagram, and as a text summary. Fox et al. (2020) used a similar representation of populations as composed of colored dots and went through a sequence of scenarios without user input (scenarios can be restarted, but will always show the same sequence of numbers).

The simulation in Stevens (2020b) represented individuals by randomly moving dots and illustrated the consequences of barriers (representing travel restrictions) or decreases in movement speed (social distancing) on proximity-based transmission rates, indicated by a graph and numbers. A related simulation was used in Stevens (2020a), which was complemented by slider-based calculations. Simulations in Kristof and Thompson (2020) were replaced by calculations and thus involved no random components, but were controlled by six continuous slider and an additional binary tab, and presented an area diagram with three plotted variables.

Our own intervention is closest to the variant in (Stevens & Muyskens, 2020): Two simulation panels allowed for repeated exploration of a limited number of scenarios (two in the first, and four in the second panel). Each click of the button resulted in a graph-representation of 30 simulation runs. All 30 result lines were presented simultaneously (together with 30 results for a second scenario). Both features are helpful for exploring variation within and between conditions, but this requires the willingness of the observer to engage with this added complexity.

We simplified our simulator and decided to show only one simulation run at a time. We offered a total of five scenarios with different splits of the population into AllH-players and AllG-players with 99%, 75%, 50%, 25% and 1% AllH-players, respectively. After each button click, the progression of the number of purple players across rounds was displayed, with a sentence summarizing the final state. To encourage the experience of variance within simulation scenarios, each of the five scenarios had to be experienced at least three times (with no upper limit for exploration). This approach is close to the experience sampling approach in Kaufmann et al. (2013). We present the output diagrams in section E10.

## E10 Study 2: Selected Materials

*Note: The full set of materials is found in document **SM2** (Study Materials Study 2, see Section E1).*

## E10.1 Transmission Game: Condition-specific instructions

### *E10.1.1 Instructions: Injunctive norms condition*

[*See Figure E42.*]

We would like to highlight some consequences of your choices:

You will have noticed that you can obtain a **personal benefit** by choosing action H (=40 points) over action G (=8 points).
But be aware that this benefit in points comes at a **potential cost paid by others, and even by yourself**.
If you are a purple player and choose action H, there is a higher chance that a blue player you are paired with will turn purple (thereby losing their bonus money for the game).
If you are a blue player paired with a purple player and you choose action H, then you will have a higher chance of turning purple yourself (thereby losing your bonus money for the game).

**All points scored across the rounds are worth nothing to a purple player.**

Choosing action H may therefore endanger your own bonus money and the bonus money of other participants.
Choosing action G protects others and yourself from changing into purple players and losing bonus money.

In brief:
**Choose action G to protect your and other players' bonus money.**

### *E10.1.2 Screenshots: Injunctive-norms*

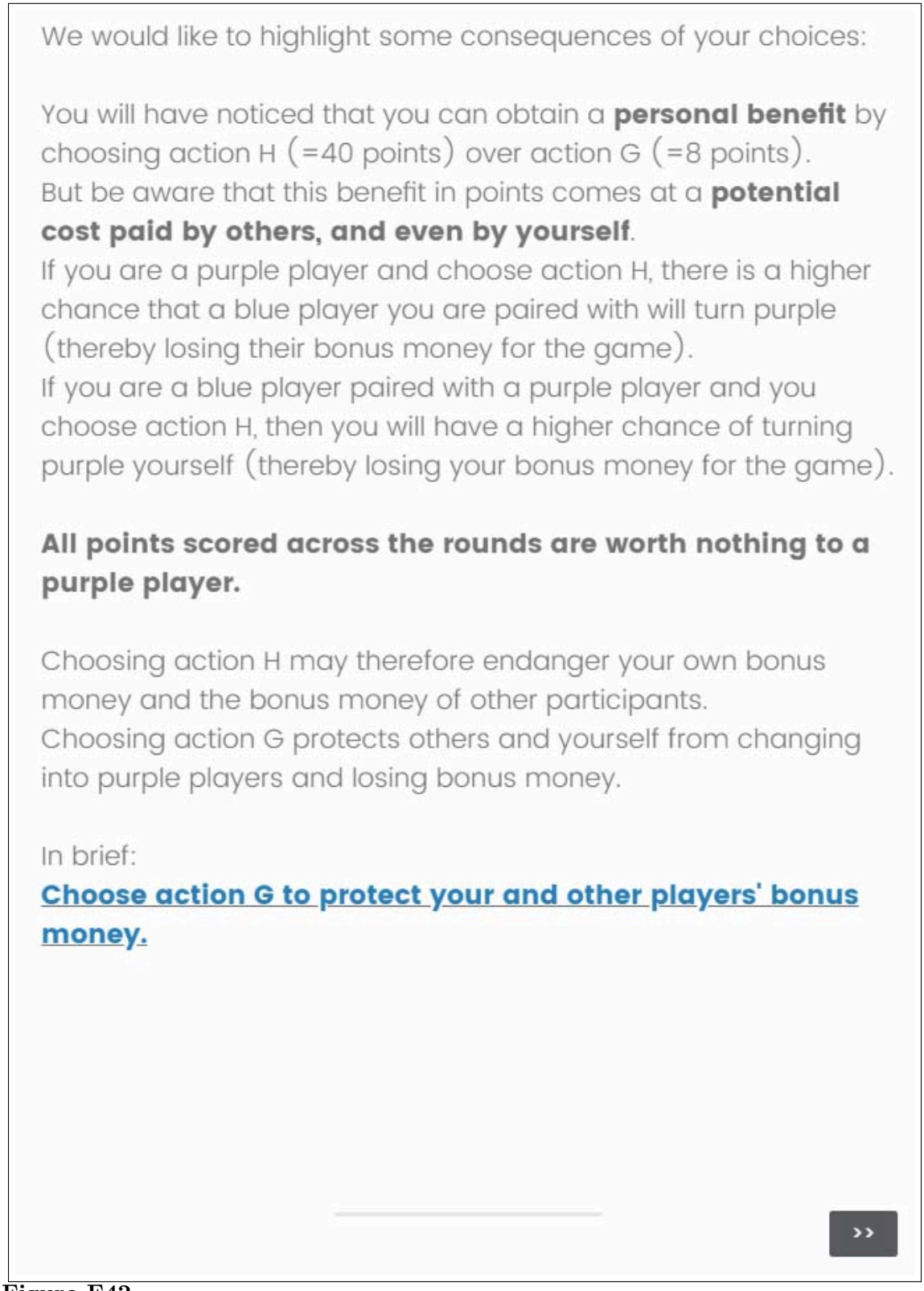


**Figure E42**
*Instructions (injunctive norms)*

### *E10.1.3 Instructions: Chain reaction condition*

[ *These instructions consist of nine pages, each featuring an image of the game development. All images can be seen in the screenshots, presented in order after the text collection.* ]

**Page 1/9.** We would like to show you an example of how a single change of player color can result in a chain reaction of color changes.
In this example, one player begins the first round as a purple player and is paired with a blue player.

**Page 2/9.** Now imagine that—as a consequence of both players' actions—the blue player changes to purple.
There is now one more purple player in the game who will not receive a bonus payment at the end. But the effects of this change do not end here.

**Page 3/9.** The purple player will meet 24 other players over the course of the game. It is possible that none of these 24 other players change color.
But if only one of the two players chooses action H in each round, the chance of changing color would be 15% in each round (and it would be 25% if both choose H).

**Page 4/9.** In this example, 5 of the 24 players change color to purple.
In total, there are now six more players who have changed color as a consequence of the first pair of actions.
But the effects of the initial change do not end here.

**Page 5/9.** Each of the five purple players who changed color (as a consequence of being paired with the player who changed first) will meet other players during the remaining rounds of the game.

**Page 6/9.** In this example, 19 more players (in addition to the previous 6) change color due to these interactions.

**Page 7/9.** In total, there are now 25 players who have changed color as a consequence of the first pair of actions. But the effects of the initial change do not end here.

**Page 8/9.** Each of the 19 players at the end of the previous chain will meet other players in the remaining rounds of the game, and so forth.

**Page 9/9.** Many more players may change color (and may lose their bonus for this game) as a consequence of the very first pair of choices.

Players can change their color for other reasons as well and some of the dots above might actually be the same player (since two players may meet the same third player in different rounds)—but this does not change the lesson of this example:

**In this game, if one player changes color, other players (potentially many players) may also change color in a chain reaction.**

The chance of changing color is higher when at least one player chooses action H and highest when both choose action H.

### *E10.1.4 Screenshots: Chain reaction*

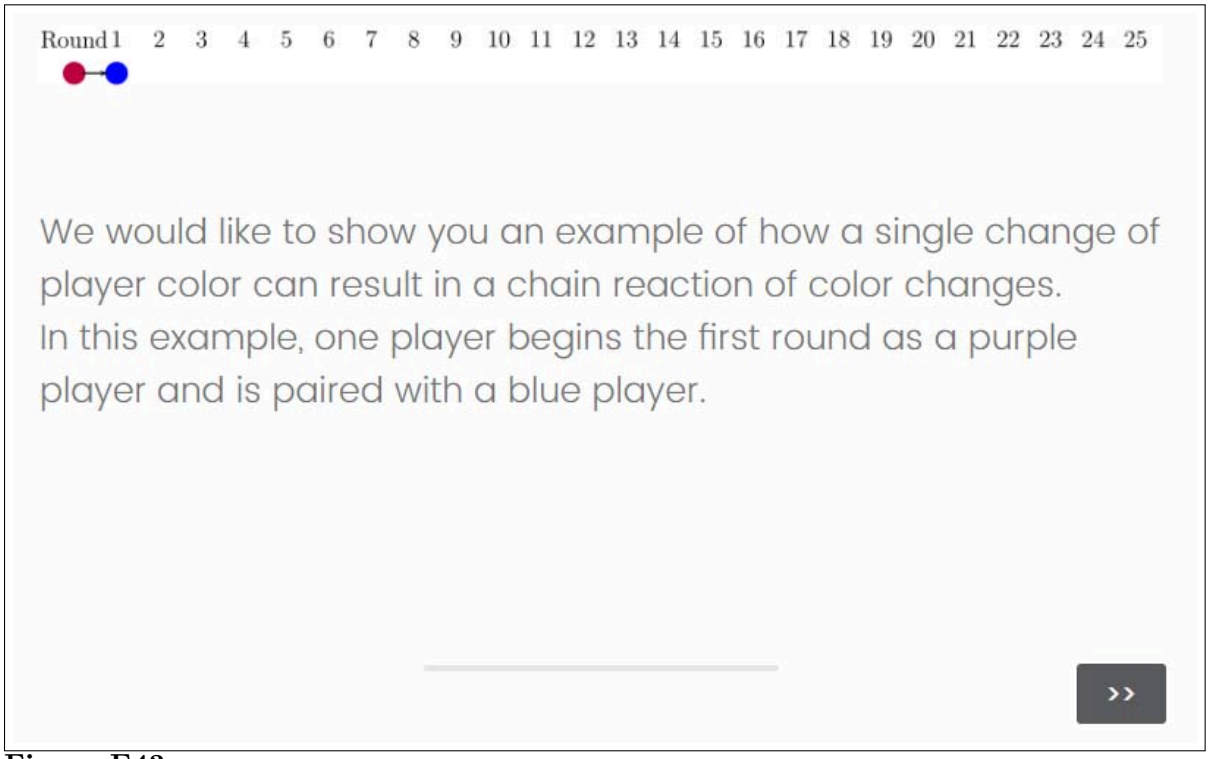


**Figure E43**
*Instructions (chain reaction, page 1/9)*

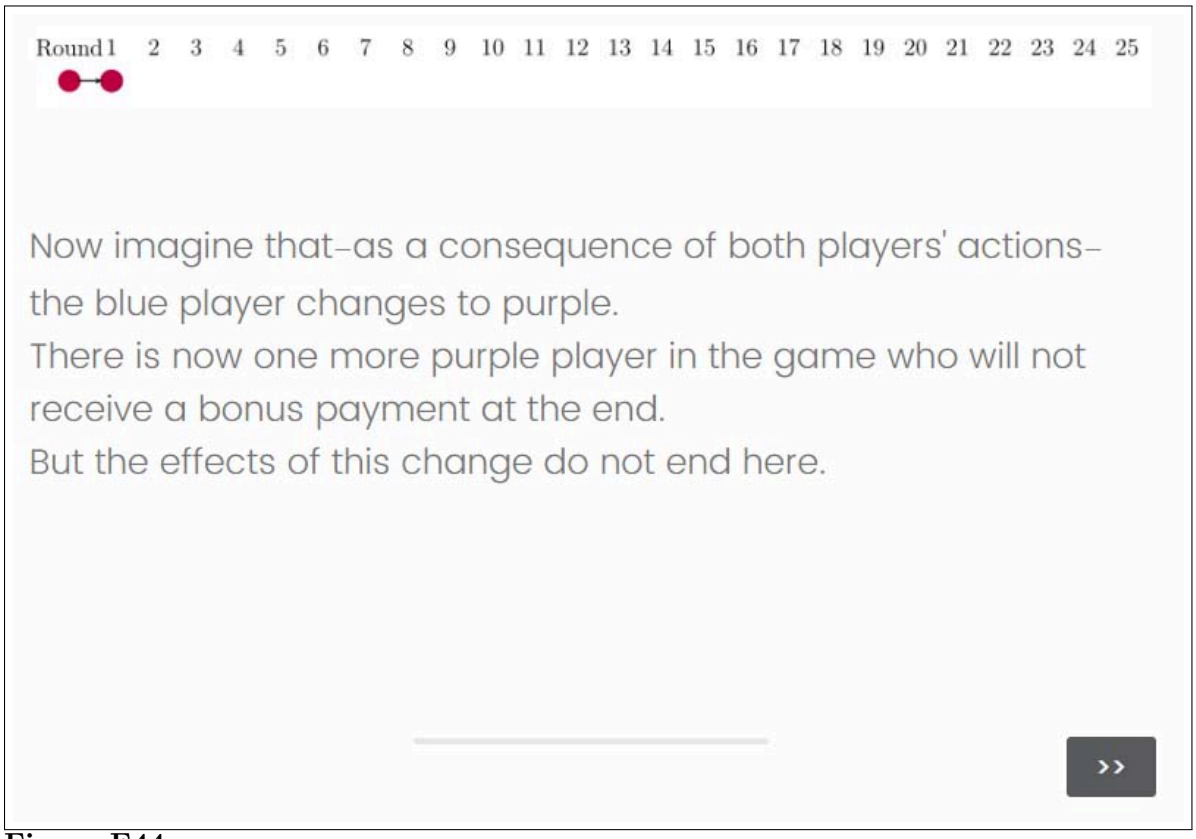


**Figure E44**
*Instructions (chain reaction, page 2/9)*

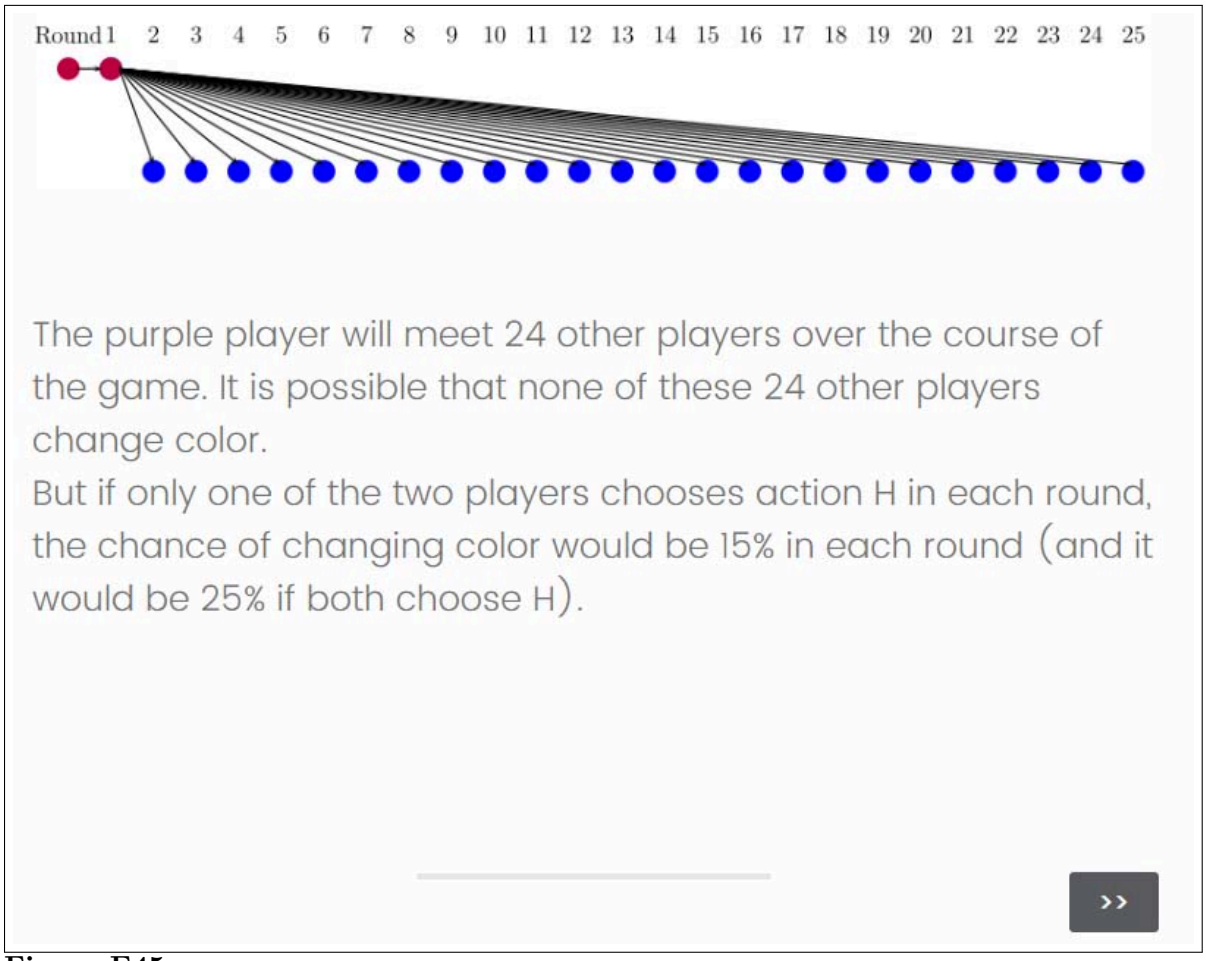


**Figure E45**
*Instructions (chain reaction, page 3/9)*

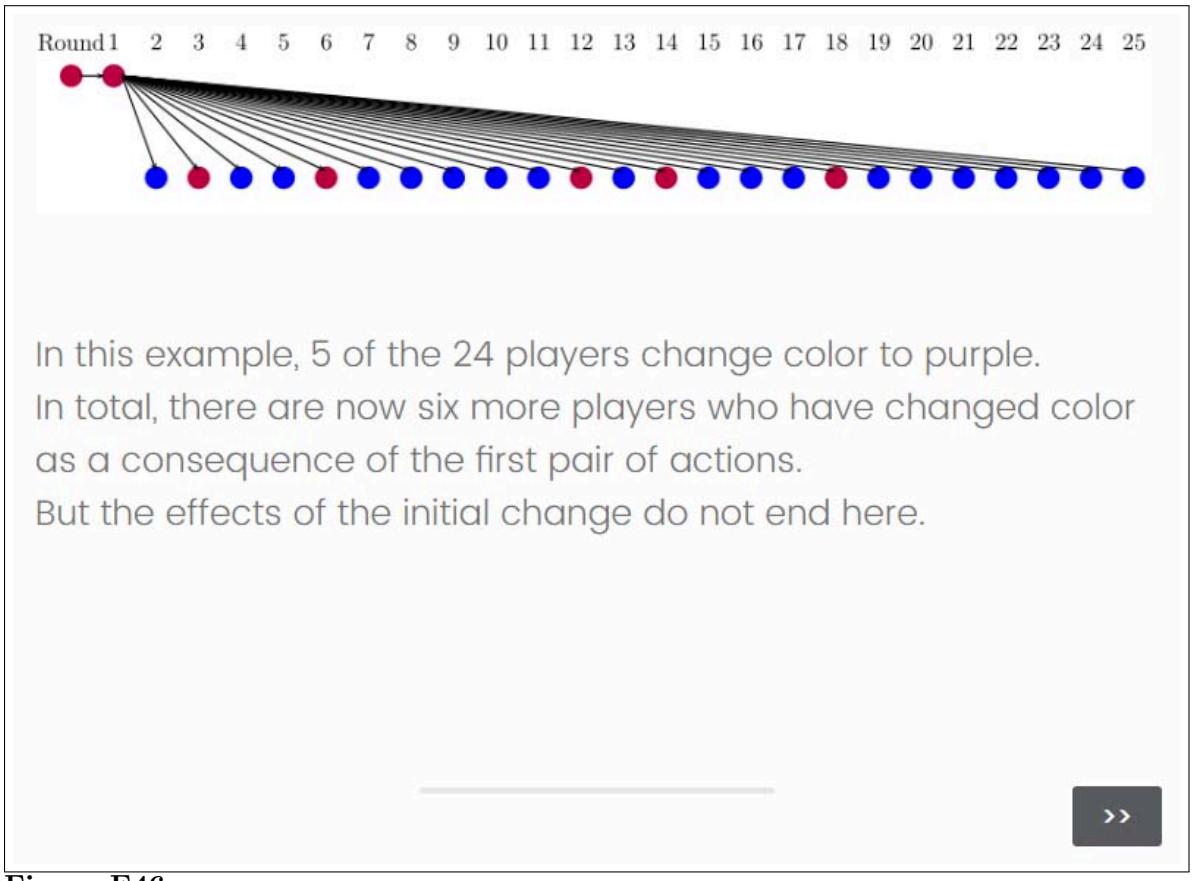


**Figure E46**
*Instructions (chain reaction, page 4/9)*

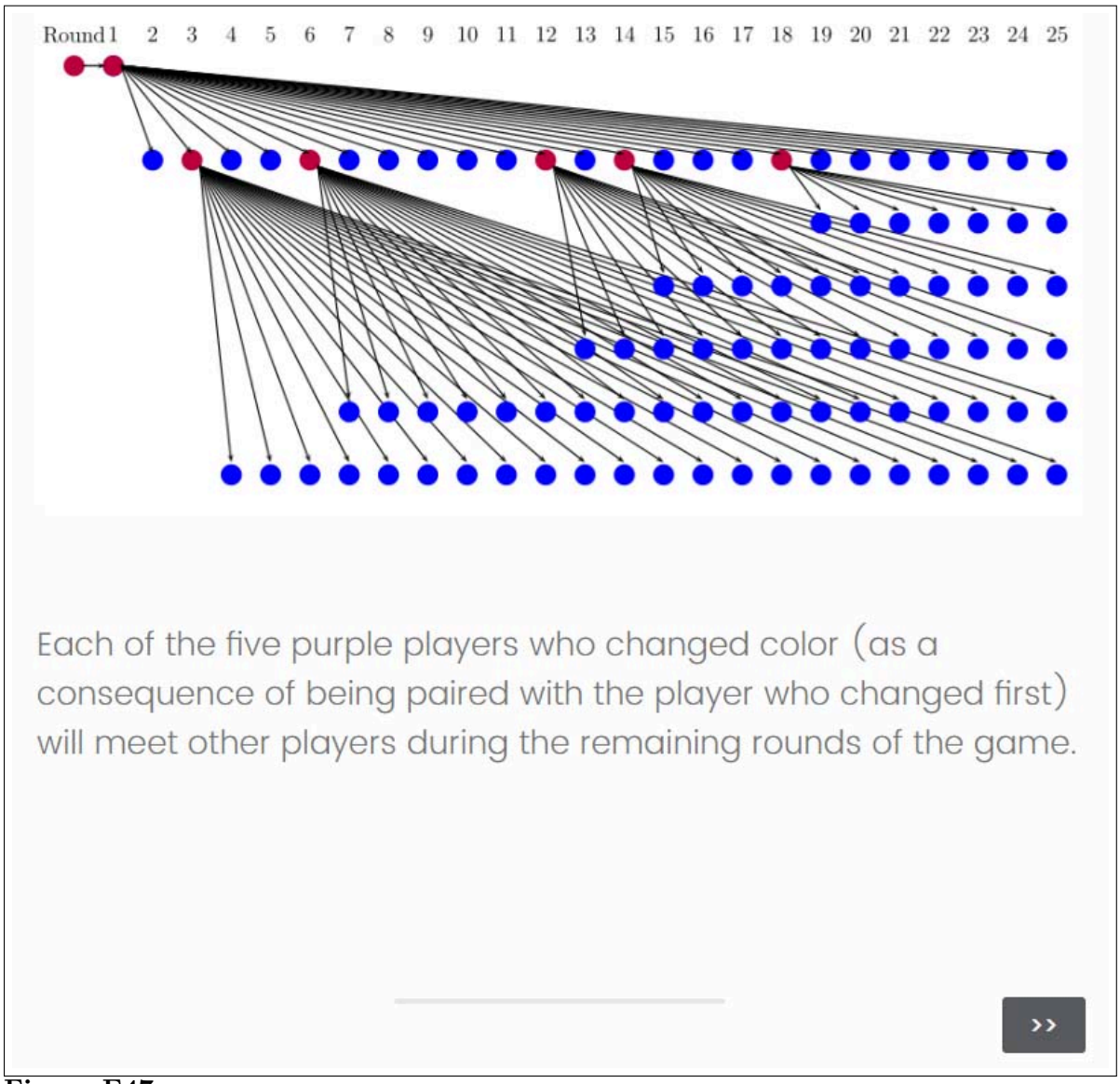


**Figure E47**
*Instructions (chain reaction, page 5/9)*

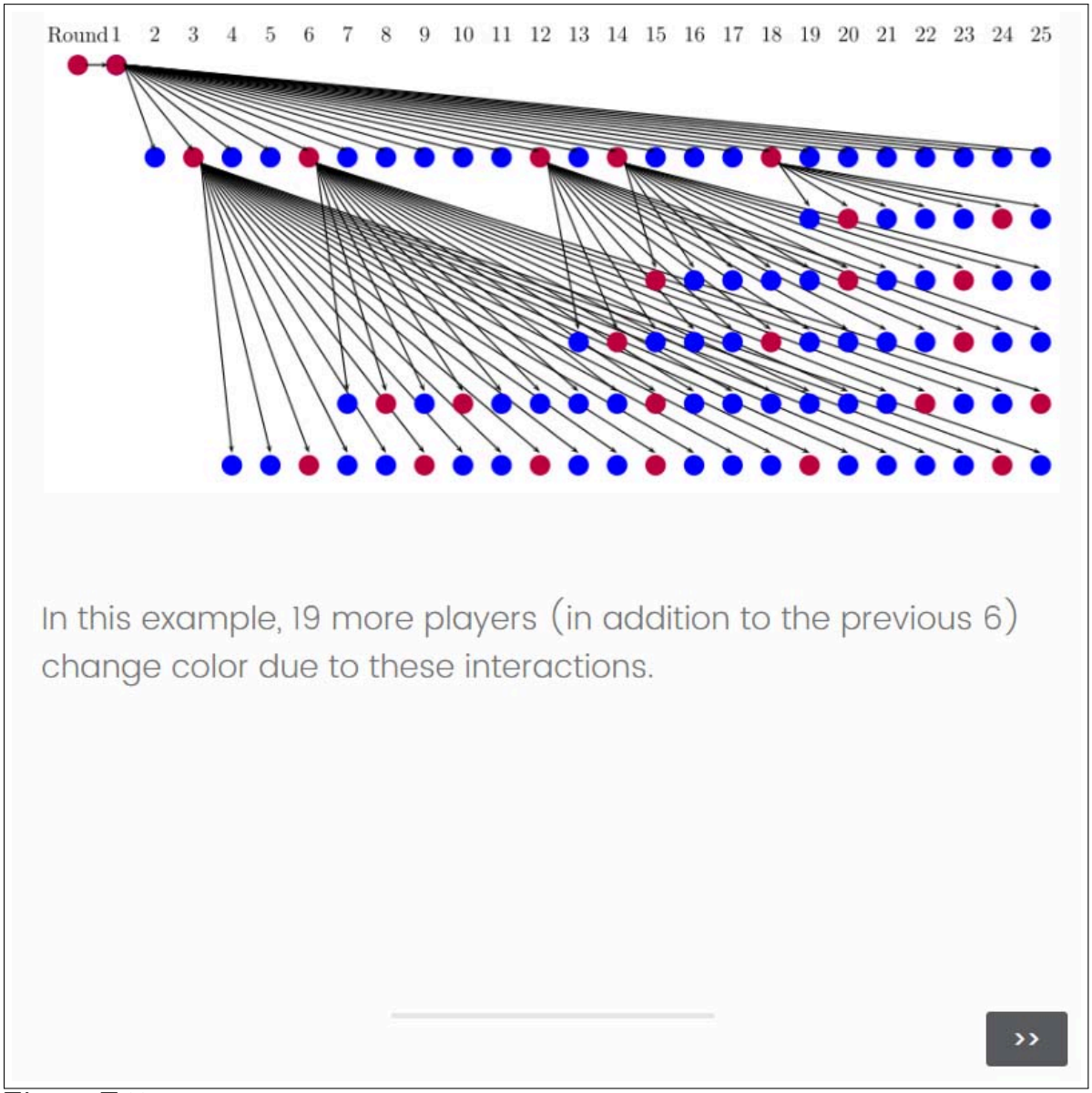


**Figure E48**
*Instructions (chain reaction, page 6/9)*

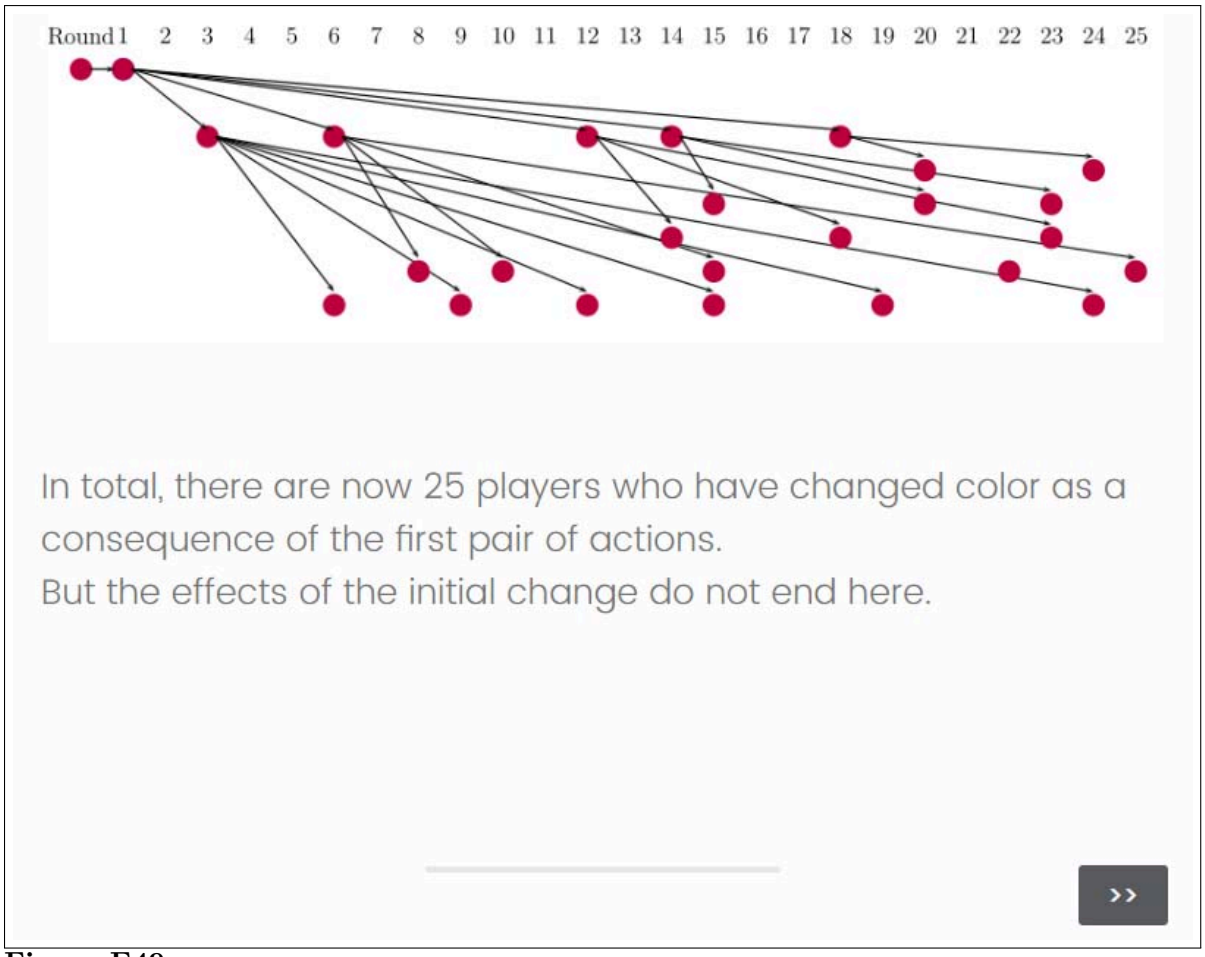


**Figure E49**
*Instructions (chain reaction, page 7/9)*

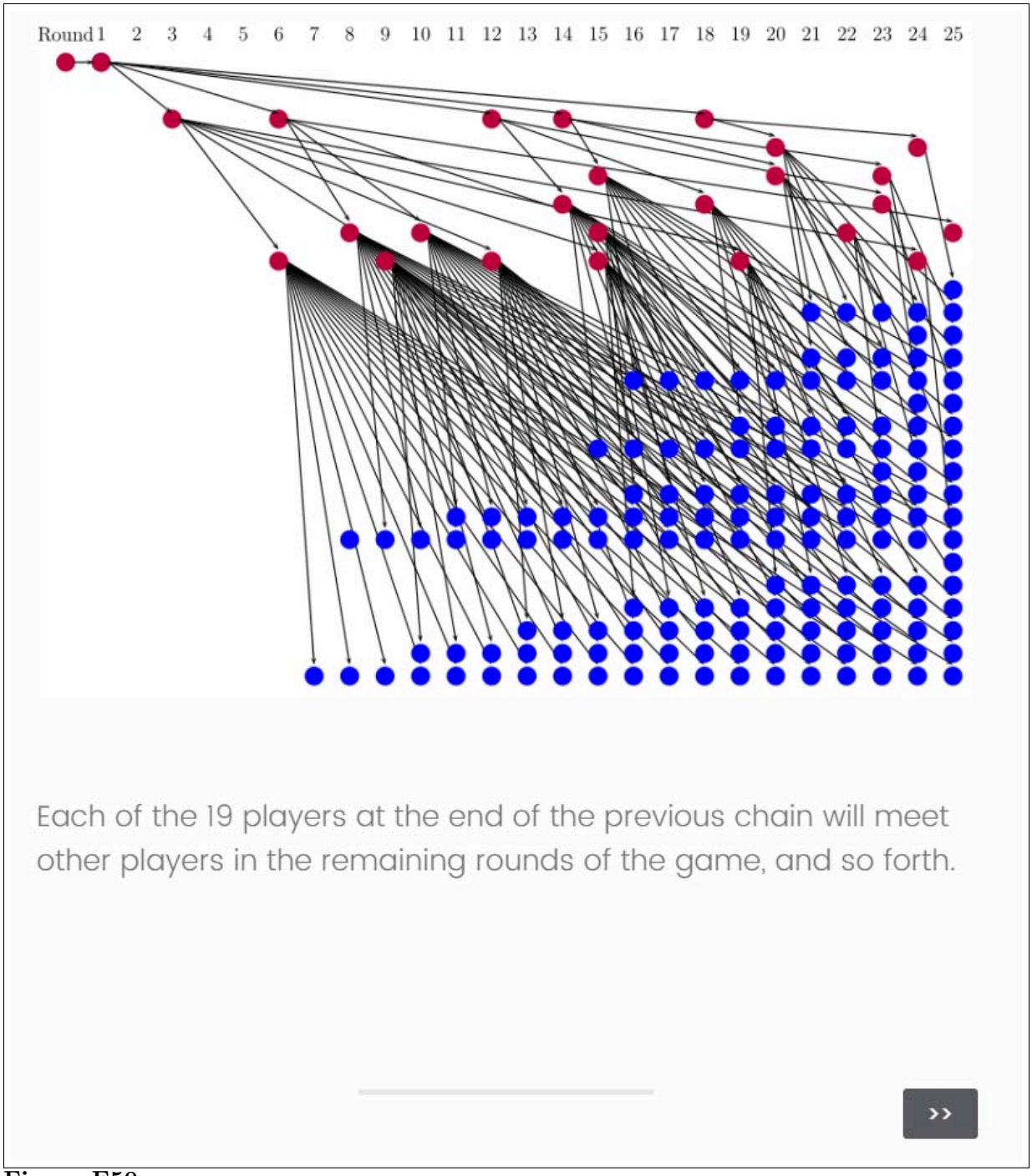


**Figure E50**
*Instructions (chain reaction, page 8/9)*

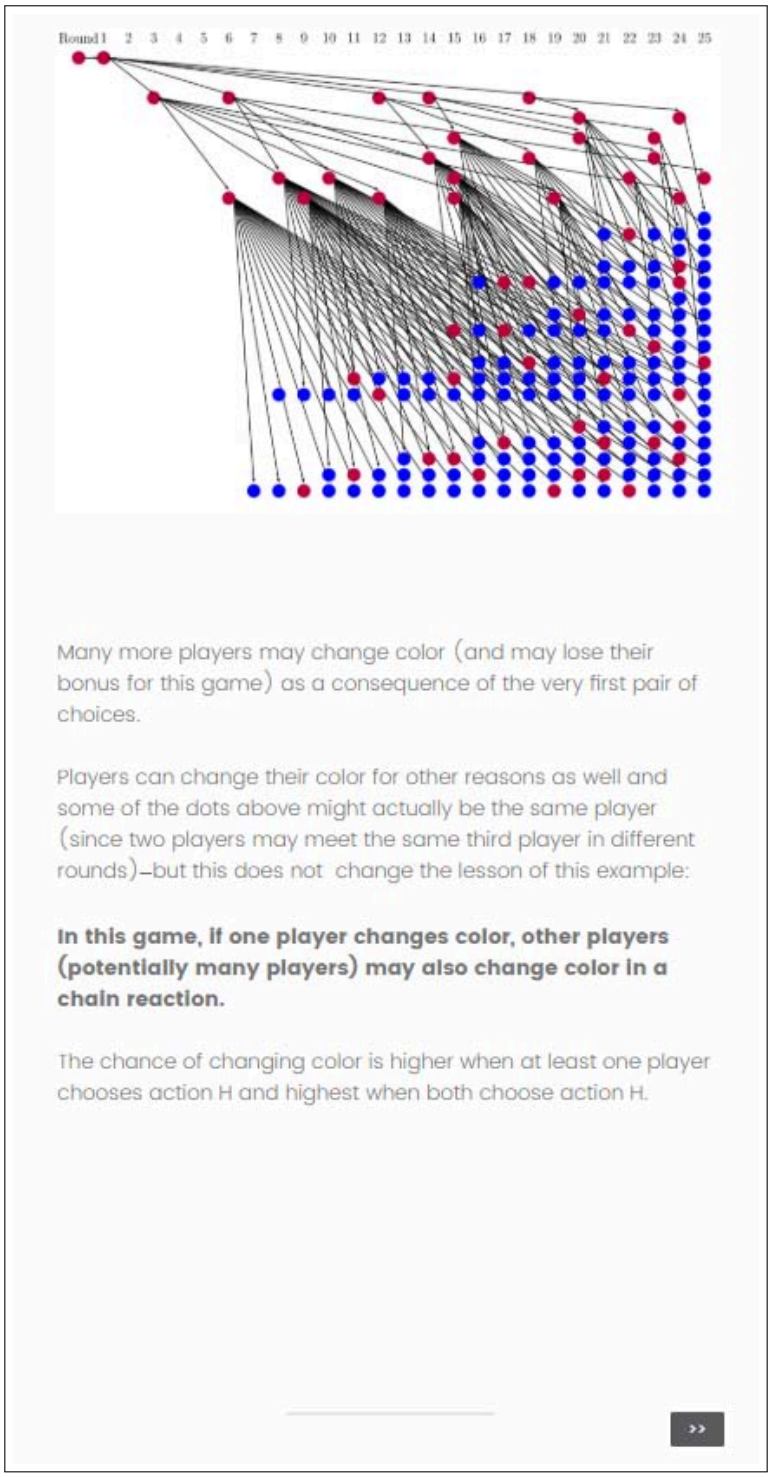


**Figure E51**
*Instructions (chain reaction, page 9/9)*

### *E10.1.5 Instructions: Simulator condition*

**Page 1/2.** Before starting the game, you will see simulations of how the game could play out.
These simulations explore the relationship between player decisions and color changes.
There are many different ways that 100 players in this game could decide, and it would take a long time to study all of them.

The simulations therefore focus on games in which each player follows a simple rule: They either choose action H all the time or they choose action G all the time.
You will be able to observe five game types:

1. **99% H: 99** of the 100 players always choose **H** (and **1** player always chooses **G**).
2. **75% H: 75** of the 100 players always choose **H** (and **25** players always choose **G**).
3. **50% H: 50** of the 100 players always choose **H**(and **50** players always choose **G**).
4. **10% H: 25** of the 100 players always choose **H** (and **75** players always choose **G**).
5. **1% H: 1** of the 100 players always chooses **H** (and **99** players always choose **G**).

You will find the simulator on the following page.

**Page 2/2.** You can simulate games by clicking on the buttons below the graph. The button label tells you which type of game is simulated. For example, the "75% H" button will simulate a game in which 75 players always choose option H and 25 players always choose option G for all 25 rounds, as described before.

Each simulation result will be shown in the graph below.

All the way on the right (on top of the y-axis), you can see how many of the 100 players changed to purple after the 25th round. The higher the point, the more players are purple (between 0 and 100). All these players would not receive any bonus money if the game were real, not simulated.

All the way to the left (above the 0 on the x-axis), you can see how many of the 100 players start the first round as purple players (always 8).

In between, you can see how many players are purple at the end of each round.

As you know, outcomes in the game partly depend on chance. Therefore, two games of the same type might result in different numbers of purple players. By simulating a game type several times, you can experience how different these numbers can be.

You can simulate each game type as often as you like. Each time you will see an entirely new game from start to finish. Each game starts with 100 blue players and the

eight purple players are determined at random. Different players will be paired and actions may result in different consequences.

You need to click each button at least three times before you can continue to the next page. By clicking on the same button, you can observe how the same type of game can result in different consequences when played several times. By clicking on different buttons, you will be able to compare different game types to see how game results depend on player behavior.

Each button click will simulate a new game (the order of clicks does not matter).

[ *The simulator is presented below this page, see screenshot.* ]

#### *E10.1.6 Screenshots: Simulator*

[ *The instructions are shown in Figure E52, Figure E53, and Figure E54. The simulator part in Figure E54 reacts to button clicks by simulating single games with the specified rate of risk-seeking decision makers, followed by an animated transition of the results graph showing the number of purple players across rounds in the simulated game. The simulations are done on the participant's computer individually. The screenshots contain examples for simulations with 99% (see Fig. E55) and 25% (see Fig. E56). There are more examples in* ***SM2****, together with the JavaScript code for the simulator. Numbers shown in button captions are counted down with each click, and the continue button is shown once all counters reach zero. There is no enforced upper limit on the number of simulations per participant.* ]

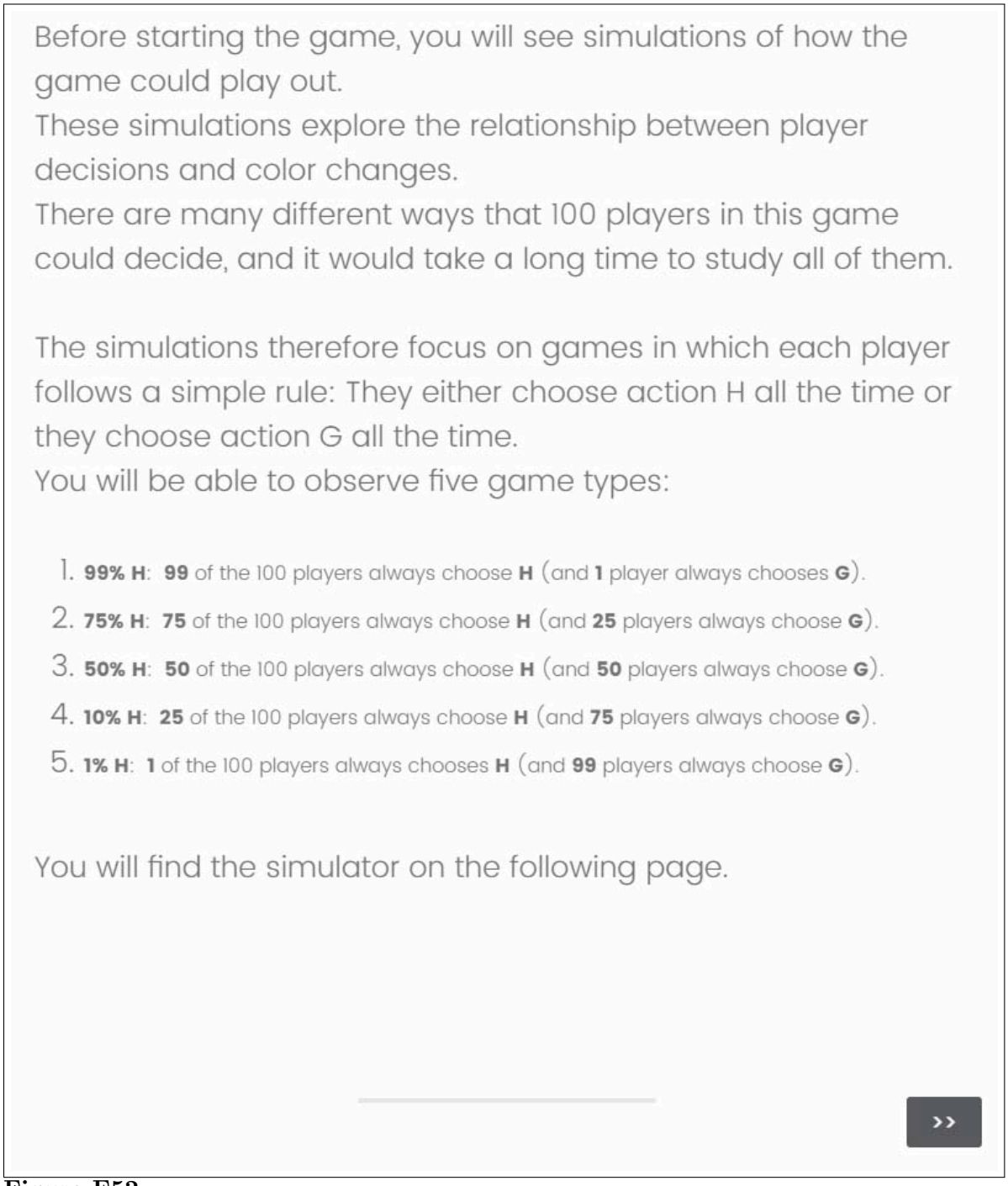


**Figure E52**
*Instructions (simulator, page 1/2)*

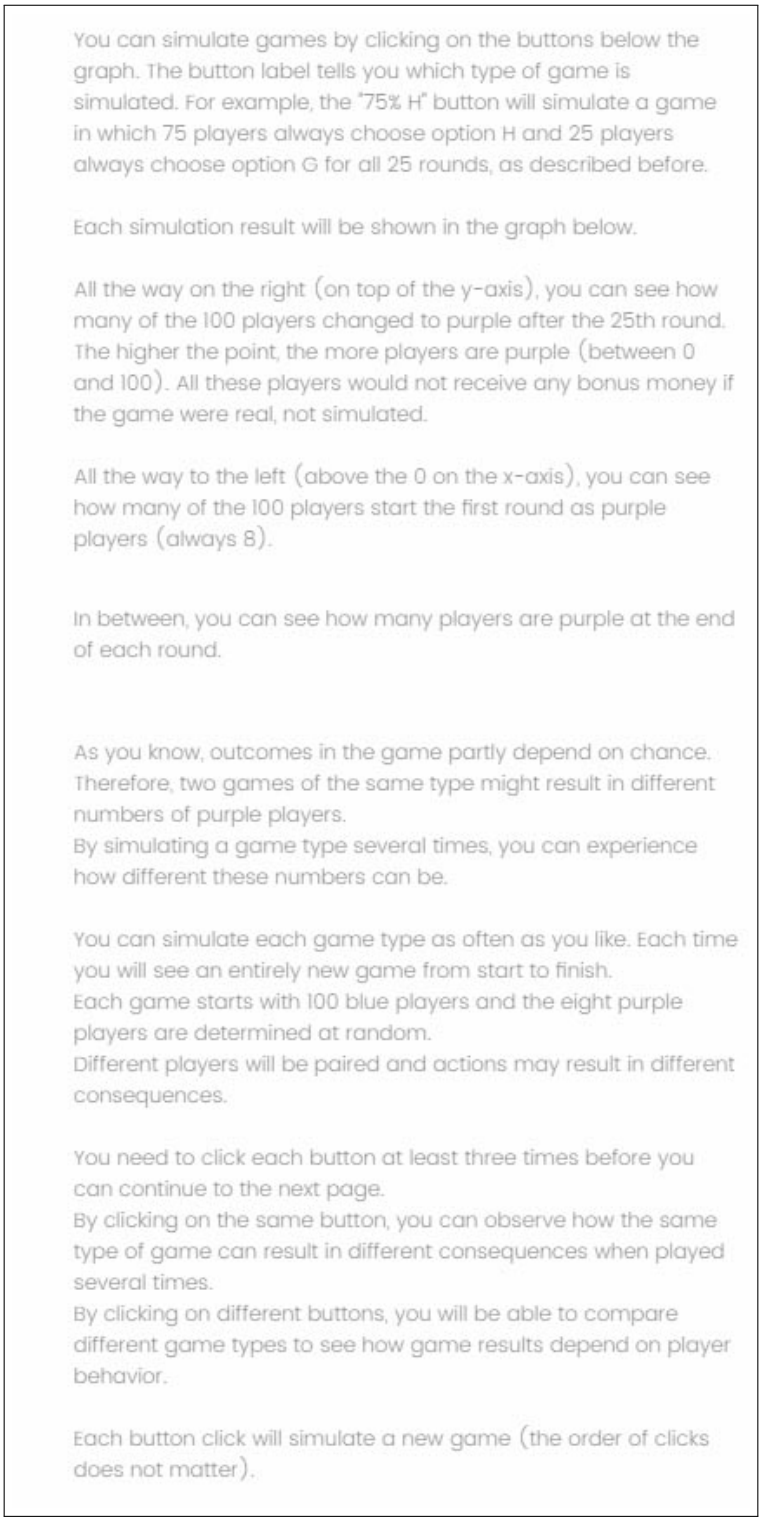


**Figure E53**
*Instructions (simulator, page 2/2, upper part)*

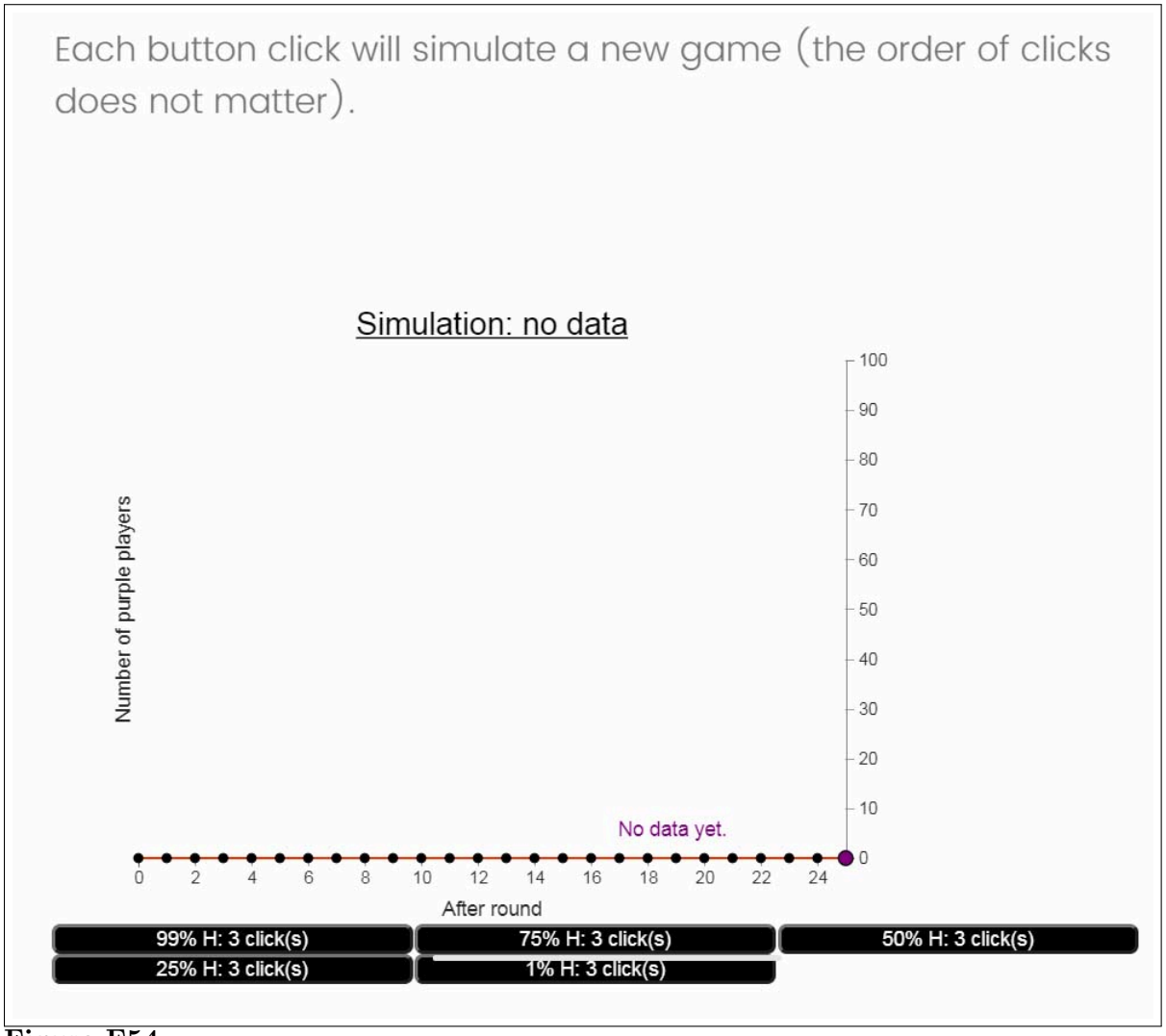


**Figure E54**

*Instructions (simulator, page 2/2, lower part)*

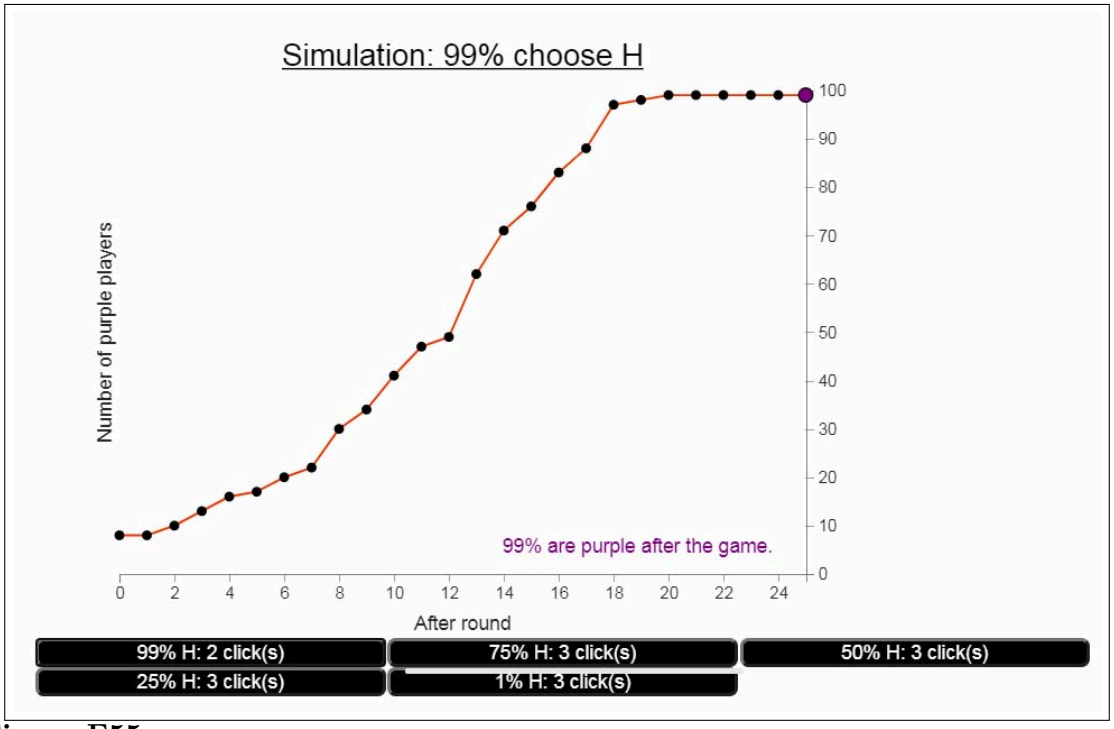


**Figure E55**

*Instructions (simulator, example of 99% H results)*

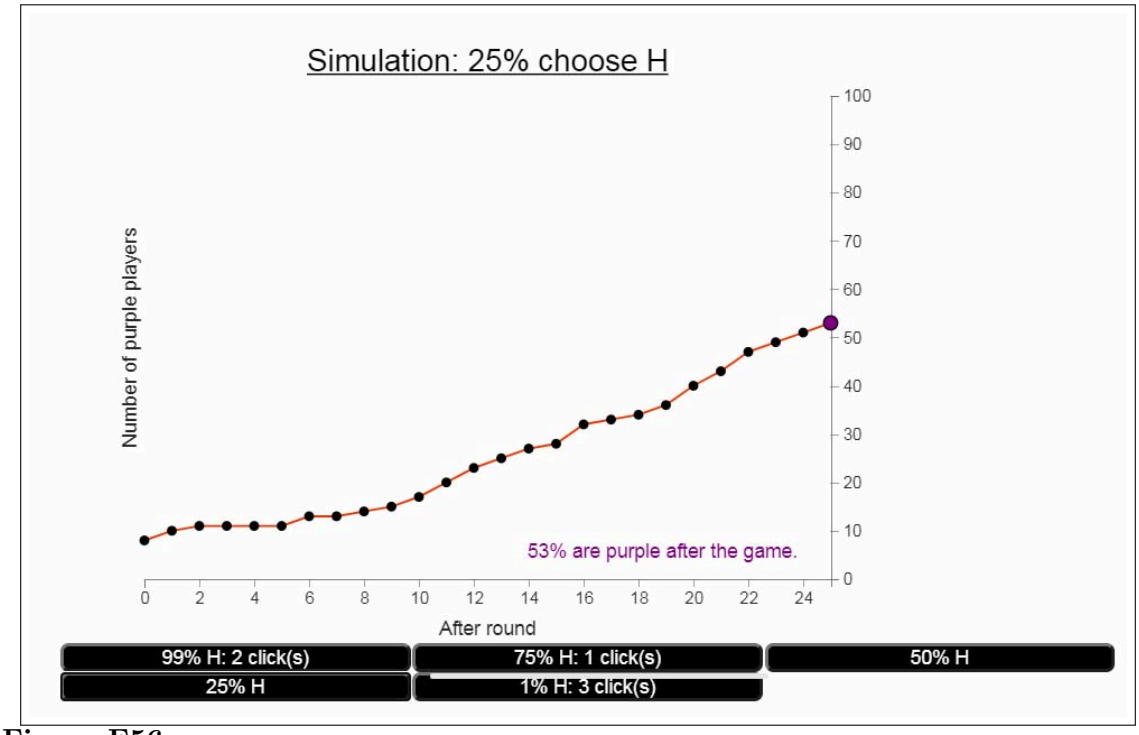


**Figure E56**

*Instructions (simulator, example of 25% H results)*

## E10.2 Condition-specific parts in the game

### *E10.2.1 Introduction: Addition in the vicarious learning condition*

[*The following text was shown below the standard introduction part (see Fig. E57).*]

In the course of the game, we will inform you about a group of participants that played the game earlier.

Before each decision, you will be able to see how many players in the previous group were purple at the beginning of each round of their game.
A graph will show you these numbers up to the round you are about to play.

This number may differ from the number of purple players in your own game.

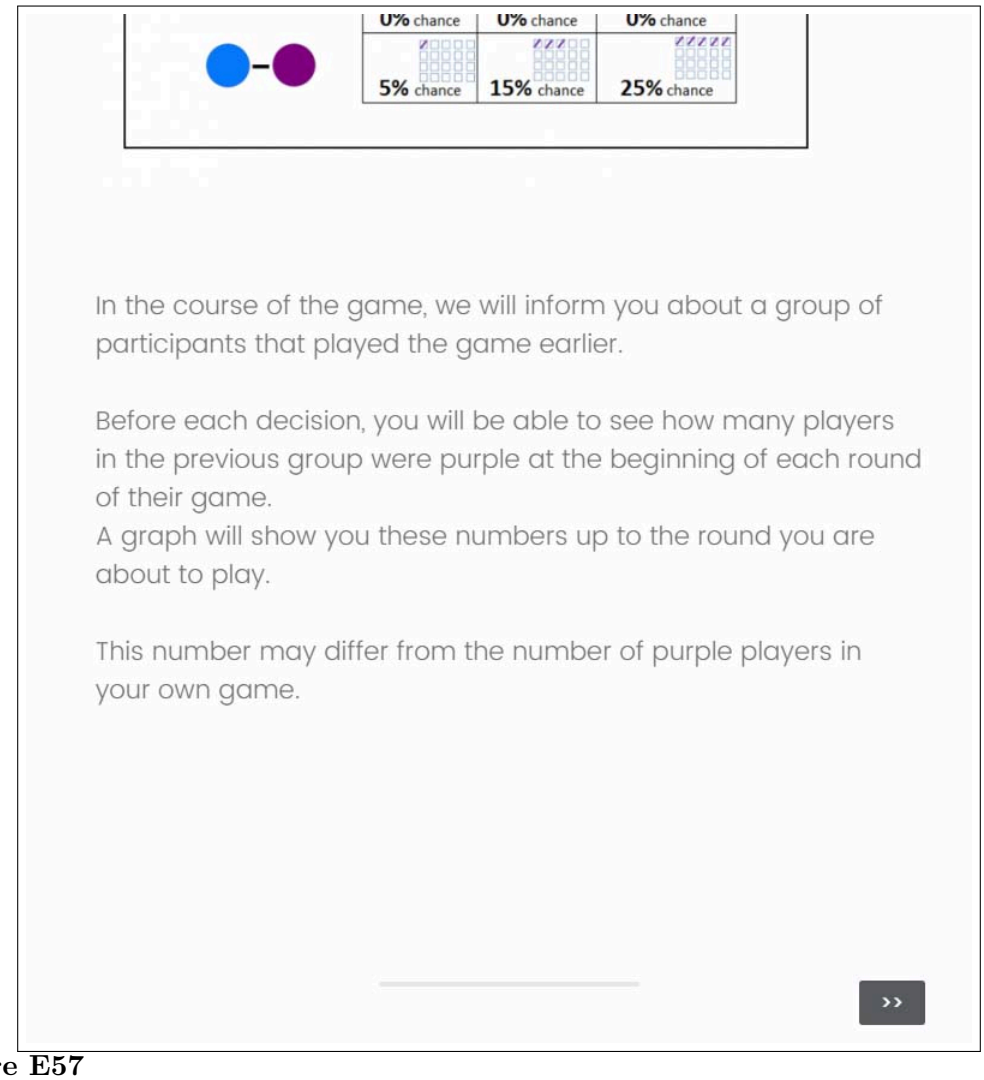


**Figure E57**

*Instructions (vicarious learning condition, the text was shown below the standard instructions)*

#### *E10.2.2 Introduction: Addition in the descriptive norms condition*

[*The following text was shown below the standard introduction part (see Fig. E58).*]

In the course of the game, we will inform you about a group of participants that played the game earlier.

Before each decision, you will be able to see how many players in the previous group chose action G in the round you are about to play.

This number may differ from the number of players choosing G in your own game.

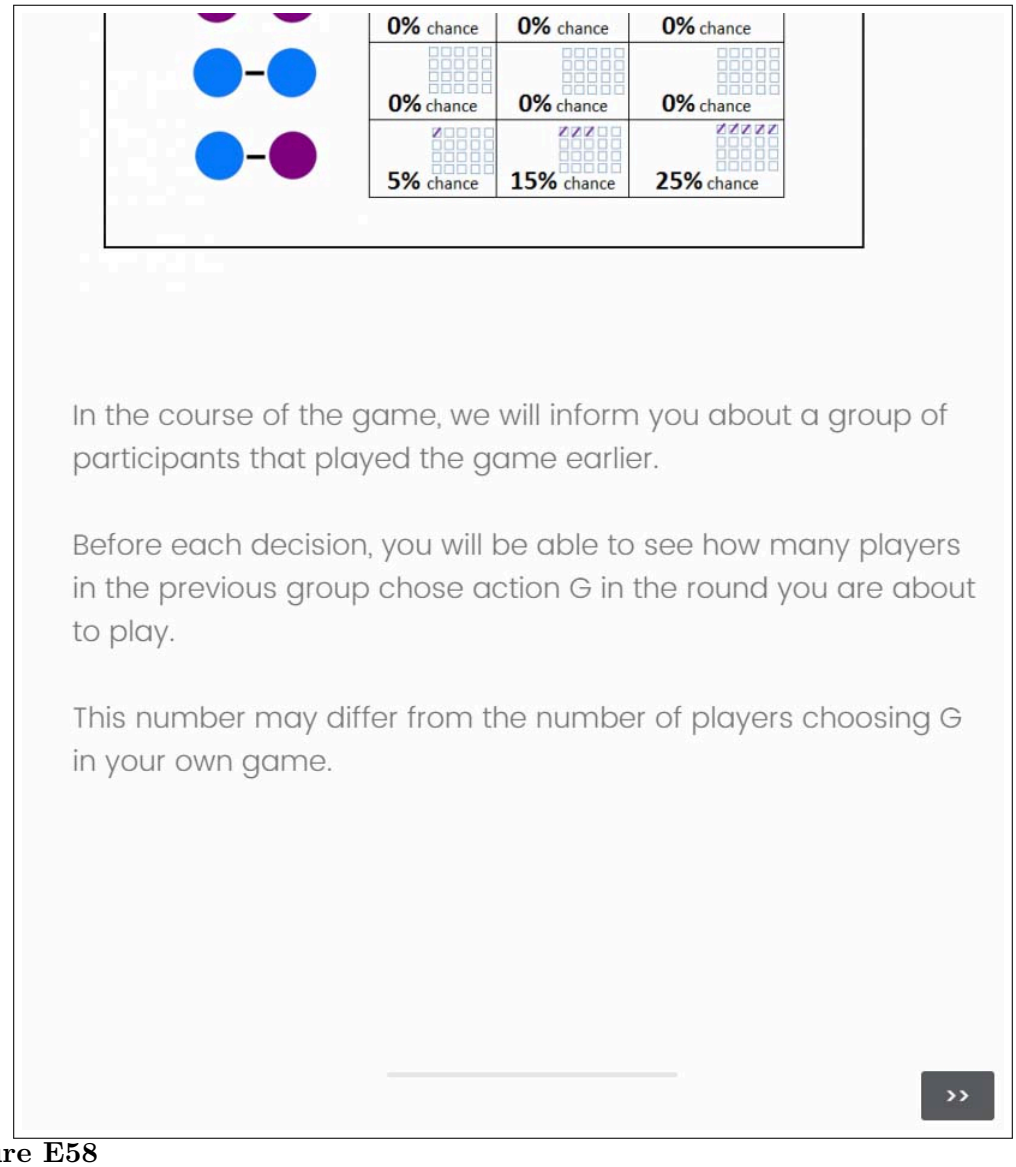


**Figure E58**

*Instructions (descriptive-norms condition, the text was shown below the standard instructions)*

#### *E10.2.3 Rounds (vicarious learning condition only)*

[*See Figure E59, Figure E60, and Figure E61.*]

Round [*current round*]/25
Points gained so far: [*current number of points*] (+ [*points gained in previous round*])
Possible bonus: £[*current bonus*] (+£[*bonus amount gained in previous round*])
Points will be converted into bonus money after the final round (for blue players only).

[*Graph of the number of purple participants in the previous group up to before the current round.*]

Above you can see information about the previous group of participants at the same point in the game.

Which of the two actions do you choose?

- Action G: I receive 8 points. (8)
- Action H: I receive 40 points. (40)

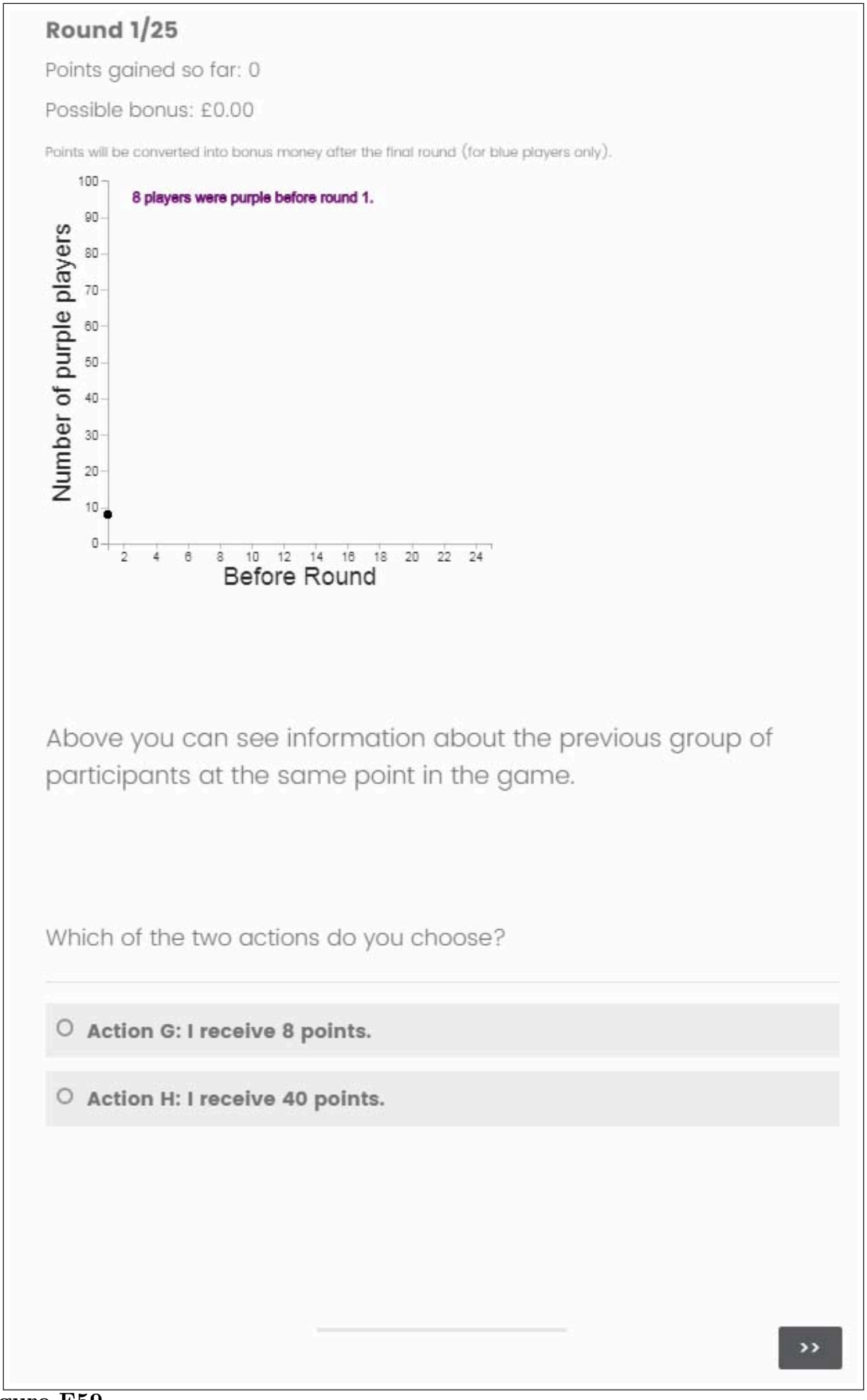


**Figure E59**
*Game round 1 (example in vicarious condition)*

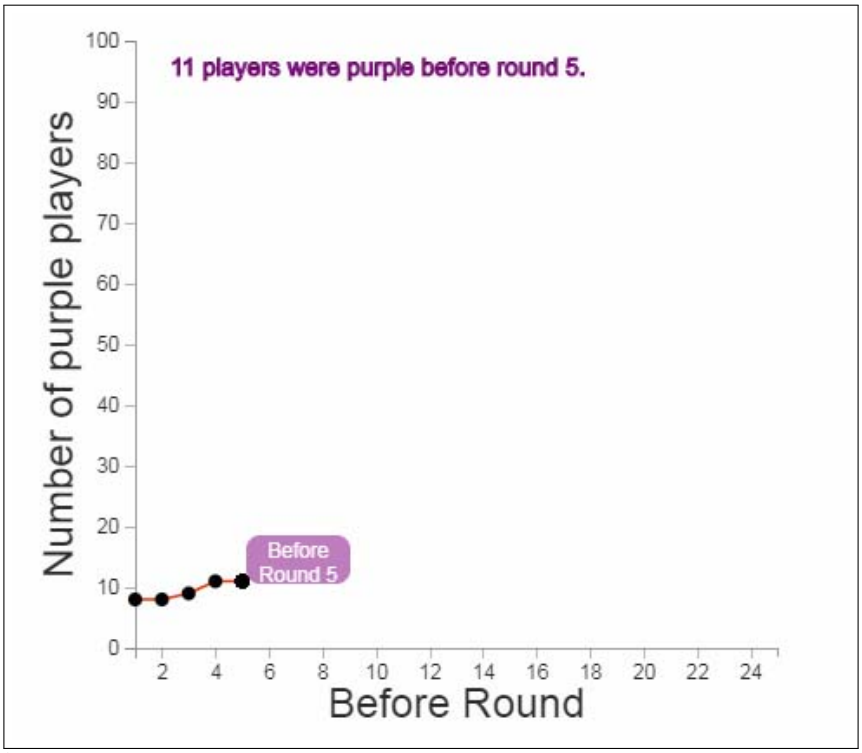


**Figure E60**

*Graph in game round 5 (example in vicarious condition with tooltip shown)*

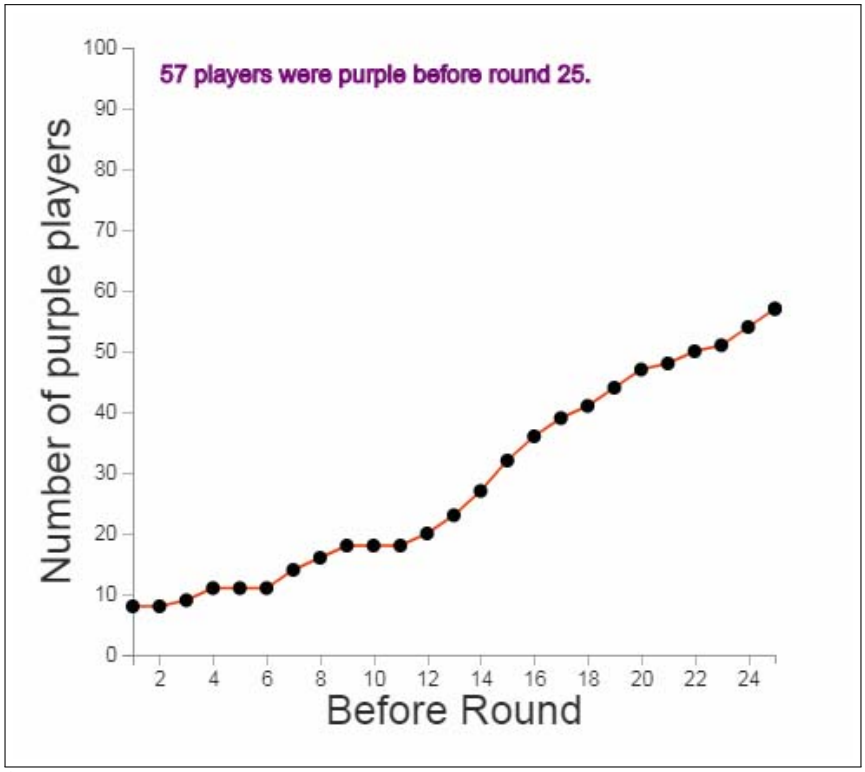


**Figure E61**

*Graph in game round 25 (example in vicarious condition)*

#### *E10.2.4 Rounds (descriptive norms condition only)*

[*All participants in this condition saw the same observed numbers, see Figure E62 and Figure E63.*]

Round [*current round*]/25
Points gained so far: [*current number of points*] (+ [*points gained in previous round*])
Possible bonus: £[*current bonus*] (+£[*bonus amount gained in previous round*])
Points will be converted into bonus money after the final round (for blue players only).

In the previous group of participants, [*number of participants in previous group who chose G*] of 100 participants chose action G in round [*current round number*].

Which of the two actions do you choose?

- Action G: I receive 8 points. (8)
- Action H: I receive 40 points. (40)

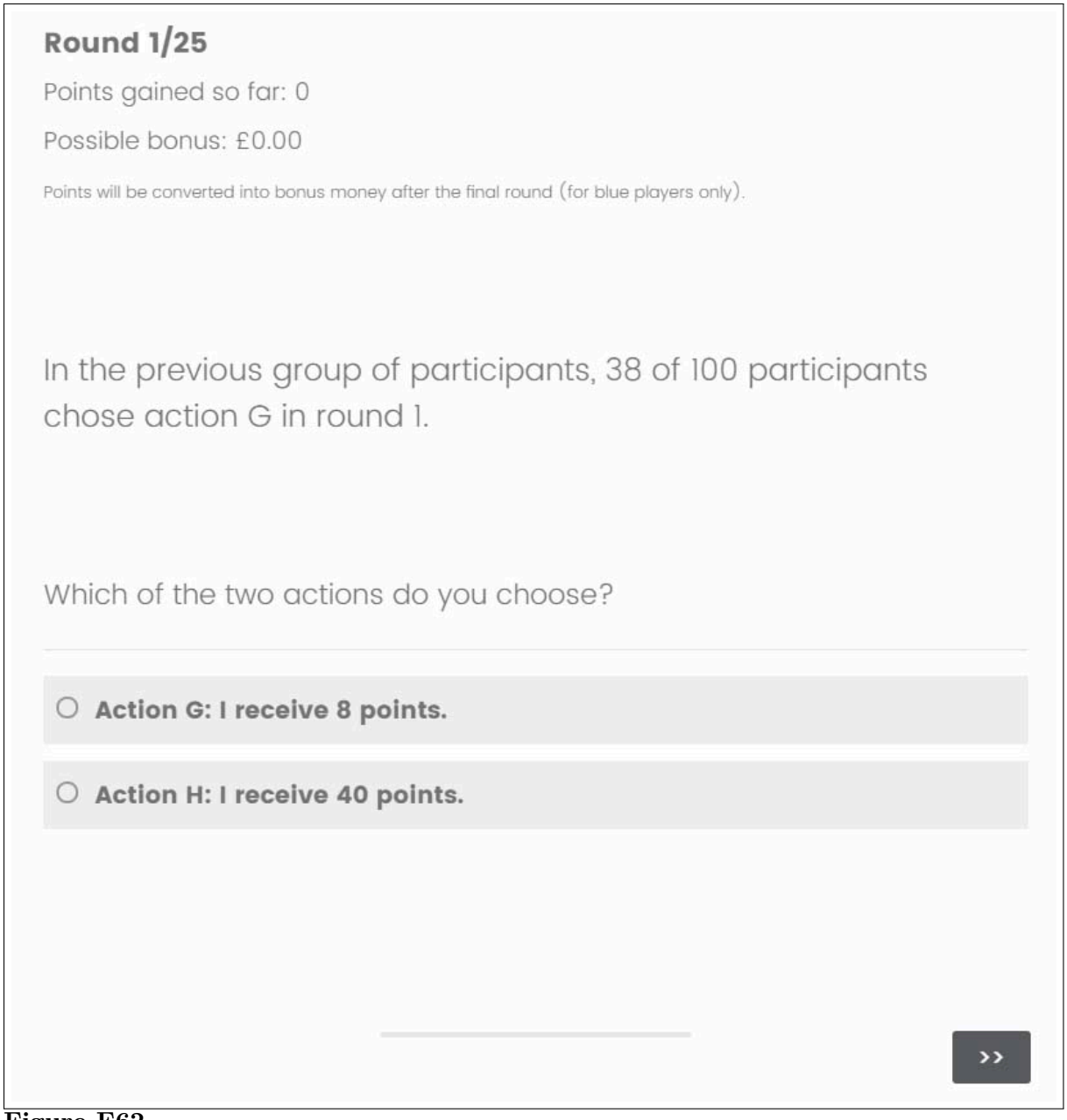


**Figure E62**

*Interface in game round 1 (example in descriptive-norms condition)*

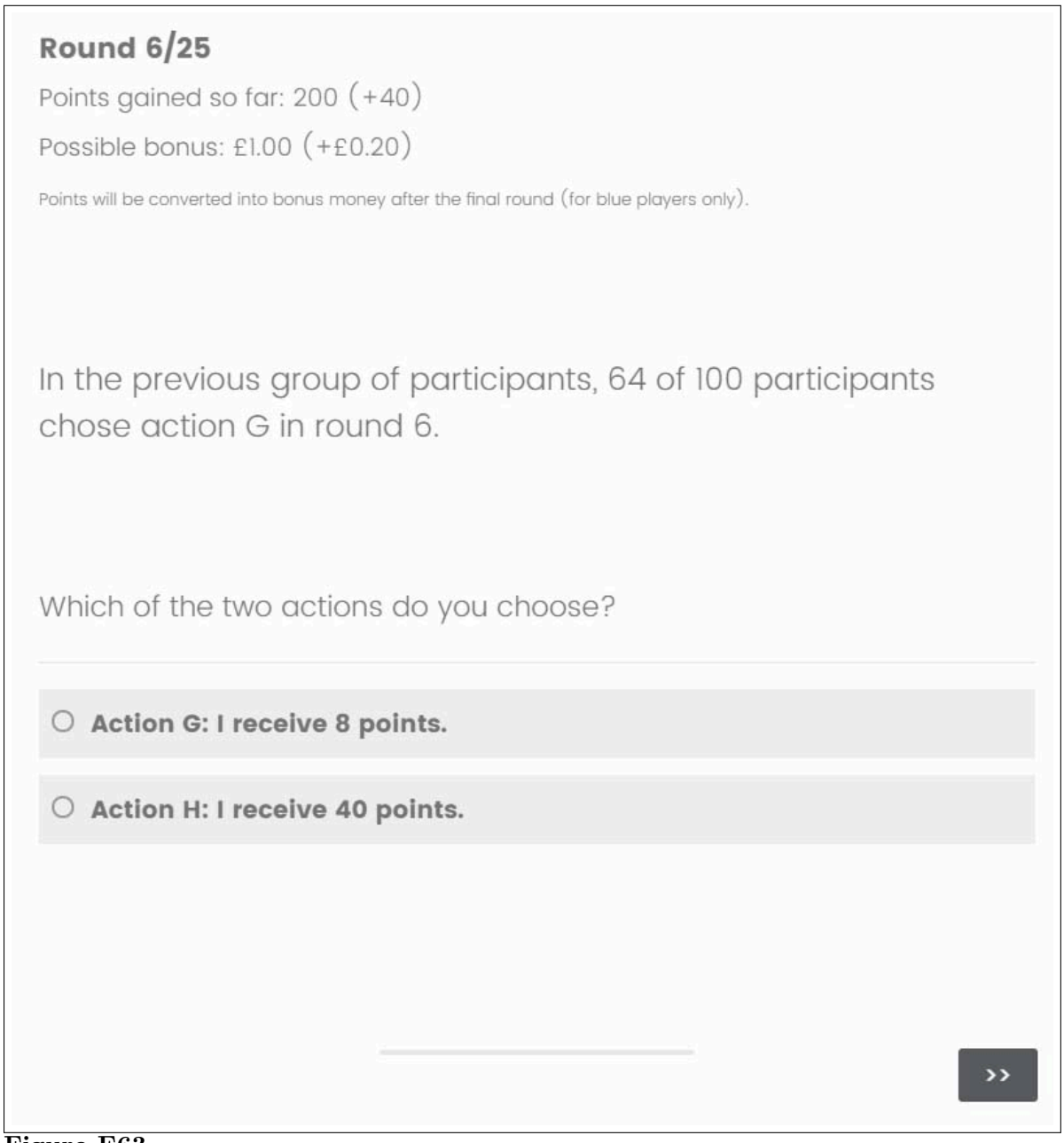


**Figure E63**

*Interface in game round 6 (example in descriptive-norms condition)*

#### *E10.2.5 Rounds (injunctive norms condition only)*

[*The same message was added in all rounds; see Figure E64.*]

Round [*current round*]/25
Points gained so far: [*current number of points*] (+ [*points gained in previous round*])
Possible bonus: £[*current bonus*] (+£[*bonus amount gained in previous round*])
Points will be converted into bonus money after the final round (for blue players only).

**Choose action G to protect your and other players' bonus money.**

Which of the two actions do you choose?

- Action G: I receive 8 points. (8)
- Action H: I receive 40 points. (40)

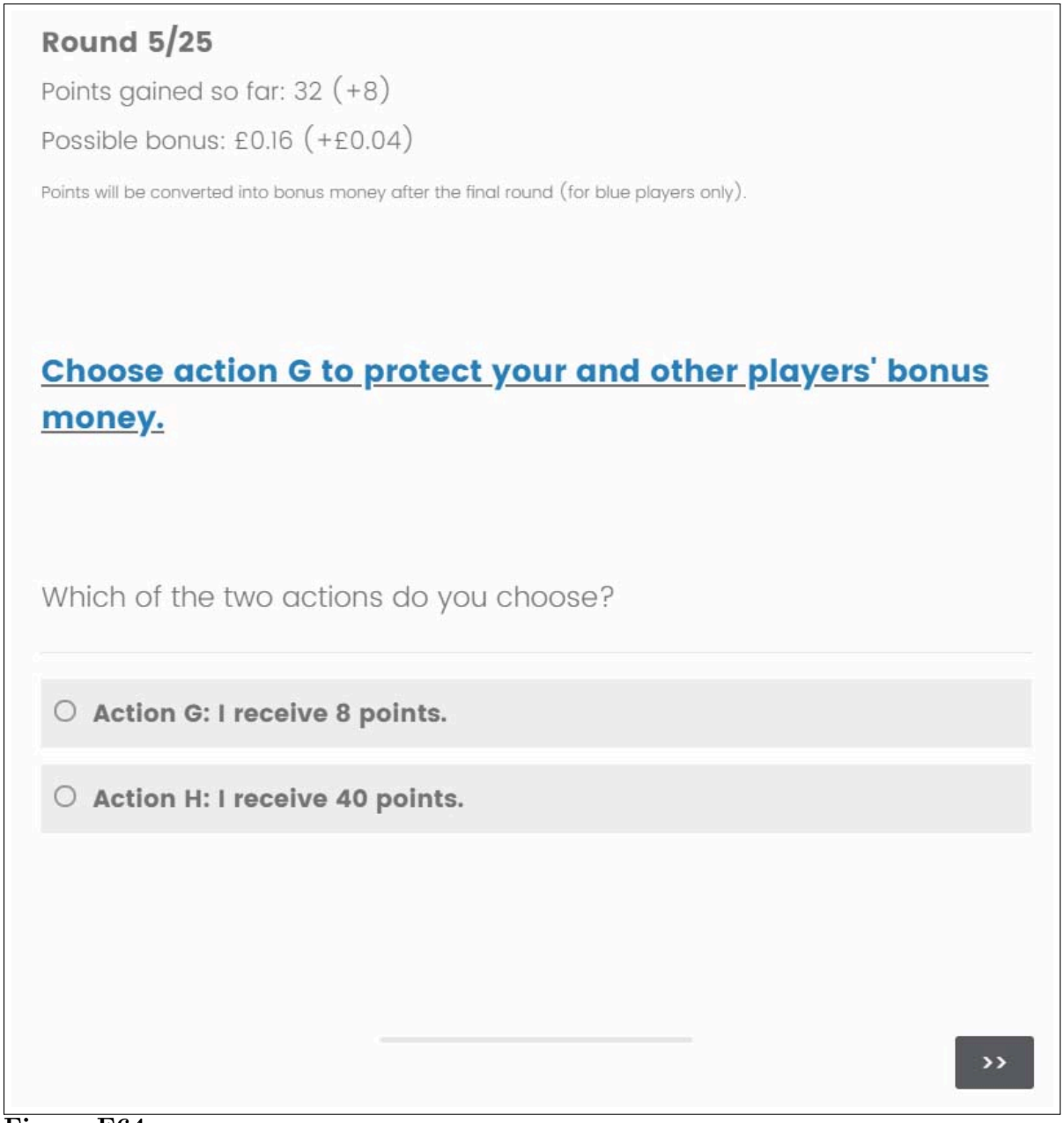


**Figure E64**

*Game round 5 (example in injunctive-norms condition)*

## E10.3 Measures and additional variables used in Study 2

A complete description of items and scale statistics can be found in **SM2** (all screen texts and screenshots) and **VM2** (distributions, scale statistics, scale formats, etc.). For the most part, the scales and items in Study 2 are a subset of those used in Study 1, with some small differences.

### *E10.3.1 Demographics*

Participants indicated their gender, age, education level, household income in 2019, current employment status, and activity on MTurk (if applicable).

### *E10.3.2 Cognitive measure*

**Cognitive Reflection Test (CRT).** Participants responded to four items measuring cognitive reflection (Frederick, 2005). Three items were taken from the CRTt (Woike, 2019), the final item from the CRT 2 (Thomson & Oppenheimer, 2016).

### *E10.3.3 Political preferences and attitudes*

**SECS.** The Social and Economic Conservatism Scale (Everett, 2013) consists of 12 items in 2 subscales: social conservatism and economic conservatism.

**Single-item measures.** Participants were asked about their voter registration status, their political position and political party affiliation, and the role of religion in their lives.

### *E10.3.4 Personality scales*

**Brief HEXCACO inventory (BHI).** In Study 2, we selected 16 of the 24-item Brief HEXACO Inventory (De Vries, 2013):

1. X: eXtraversion (4 items)
2. A: Agreeableness (4 items)
3. C: Conscientiousness (4 items)
4. O: Openness to Experience (4 items)

**HEXACO60-scales.** In addition, participants completed the full Honesty-Humility and the full Emotionality scale (not presented in Study 1) of the 60-item HEXACO-60 scales (based on the HEXACO-PI-R) in the self report form (Ashton & Lee, 2009).

**Psychological reactance scale.** We used a refined, 11-item version of the Hong psychological reactance scale (Hong & Faedda, 1996).

### *E10.3.5 Personality single-item measures*

**General Risk Taking.** Again, we used the general risk question has been used in the German Socio-Economic Panel (Dohmen et al., 2011).

**Intertemporal choice.** The time preference measure is used as a "qualitative measure of patience" in Falk et al. (2018, p. 1654).

**Trust.** The trust question was part of the global preference survey, as described in Falk et al. (2018).

#### *E10.3.6 Social preferences*

**Social Value Orientation.** Participants completed the six-item version of the ring measure of social value orientation (Murphy et al., 2011).

#### *E10.3.7 COVID-19 related measures*

**COVID-19 related worries.** This 5-item scale measured concerns about the consequences of COVID-19 for self, others and society.

**Compliance with counter-measures.** Five items measured compliance with measures against the spread of the virus. These now include items about social distancing and vaccinations.

**Willingness to accept tradeoffs.** Six items measured the willingness to accept tradeoffs between health outcomes for the vulnerable population and other outcomes.

**Further measures.** Participants also responded to questions about changes to their job situation and their willingness to donate to others suffering from the consequences of COVID-19. Again, participants were asked if they had ever been tested for COVID-19 and the result.

#### *E10.3.8 Postquestionnaires*

The postquestionnaire asked for participants' motivation in the game, and asked for estimates concerning the likelihood of ending the game as a purple player, the number of purple players in several selected game rounds, and the number of H-actions in the entire group in selected rounds. This was complemented by several (but fewer, compared to Study 1) open-format questions, asking for comments on the game, hypothetical messages to other players choosing cooperative or competitive actions, and for general comments. Participants who had experienced one of the five intervention conditions were asked further questions regarding their opinion of the intervention and its effectiveness.

## E11 Study 2: Selected Results

*Note: The full set of results is found in document **SA2** (Supporting Analyses Study 2, see Section E1). Variables and measures are described in document **VM2**.*

You can find the full set of analyses and the R code for calculations and figure creation in **SA2**. Here, we show the main results relating to the game, and some results linking political attitudes and COVID-19 related scales.

### E11.1 Behavior in the transmission game

### E11.2 Risky choices by round

Figure E65 allows for a comparison of game behavior across rounds between conditions. One condition is set apart from the others: players in the injunctive norms condition chose safer actions more frequently than in any other condition. In the control condition, the majority initially chose the risky option and, similar to Study 1, the frequency of risky

choices decreased over the course of the game with a slight increase during the final rounds. Both in the simulator and the chain reaction condition, H-choices tended to be less frequent than in the control condition. H-choices in the descriptive norms condition, on the other hand, tended to be more frequent than in the control condition. In the vicarious learning condition, risky choices started at a higher frequency than in the control condition but ended at a lower frequency than in the control condition in the final rounds.

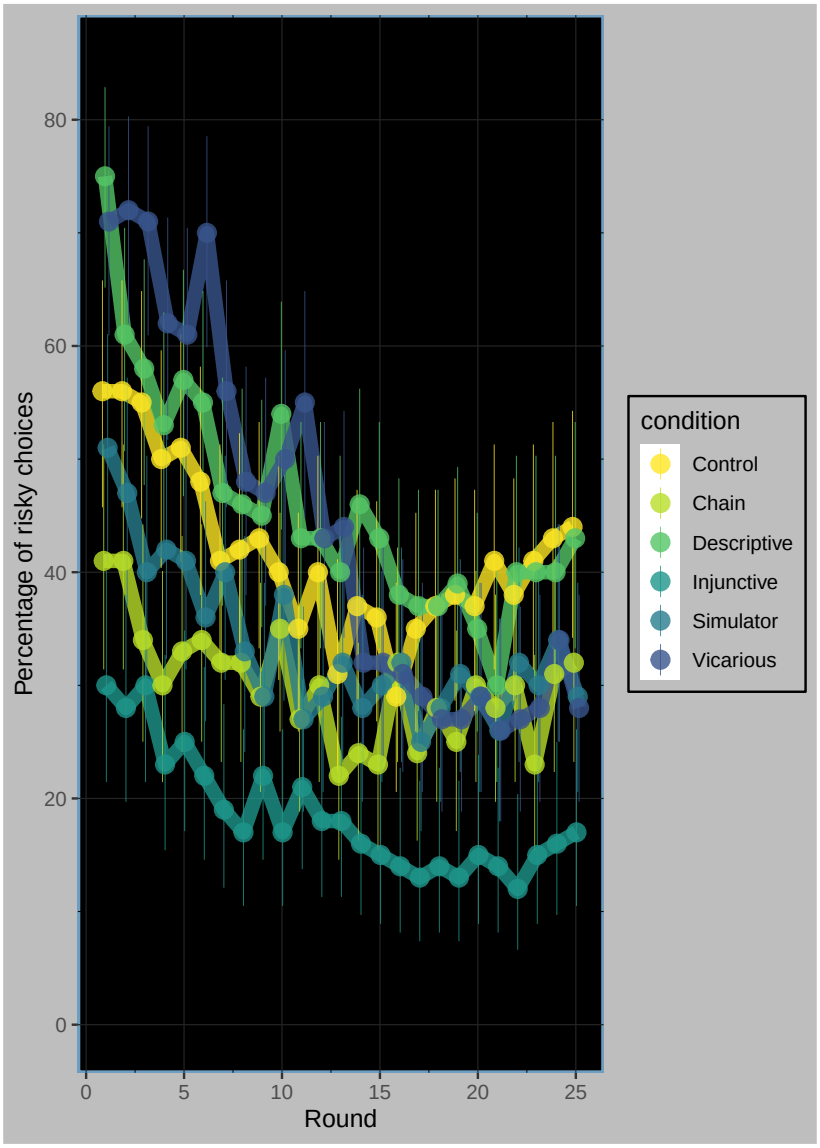


**Figure E65**

*Transmission Game round-wise results in Study 2: Points mark the percentage of H-choices in each round for the six groups, and error bars mark the 95-percent CI for percentages.*

### E11.3 Distribution of risky choices across participants

Looking at the distribution of the percentage of risky choices across players (see Fig. E66), these observations can be fleshed out.

The injunctive norms condition had the largest number of players with a very low percentage of risky actions, and a smaller number of risky players than the control condition. In the chain reaction and simulator conditions, the distribution was shifted to the left compared to the control condition, and it was shifted to the right in the descriptive norms condition. The vicarious learning condition had the largest number of players with medium levels of risk-taking and fewer players with extreme strategies.

A statistical analysis of differences (see Fig. E67) confirms these impressions. Compared to the control condition the final score after the game (before considering color switches) was significantly lower in the injunctive norms condition compared to the control condition. The same is true for the chain reaction condition.

Results in the vicarious learning, descriptive norms, and the simulator condition were not significantly different from the control condition. Results in the simulator condition were significantly lower than in both the vicarious learning and descriptive norms condition.

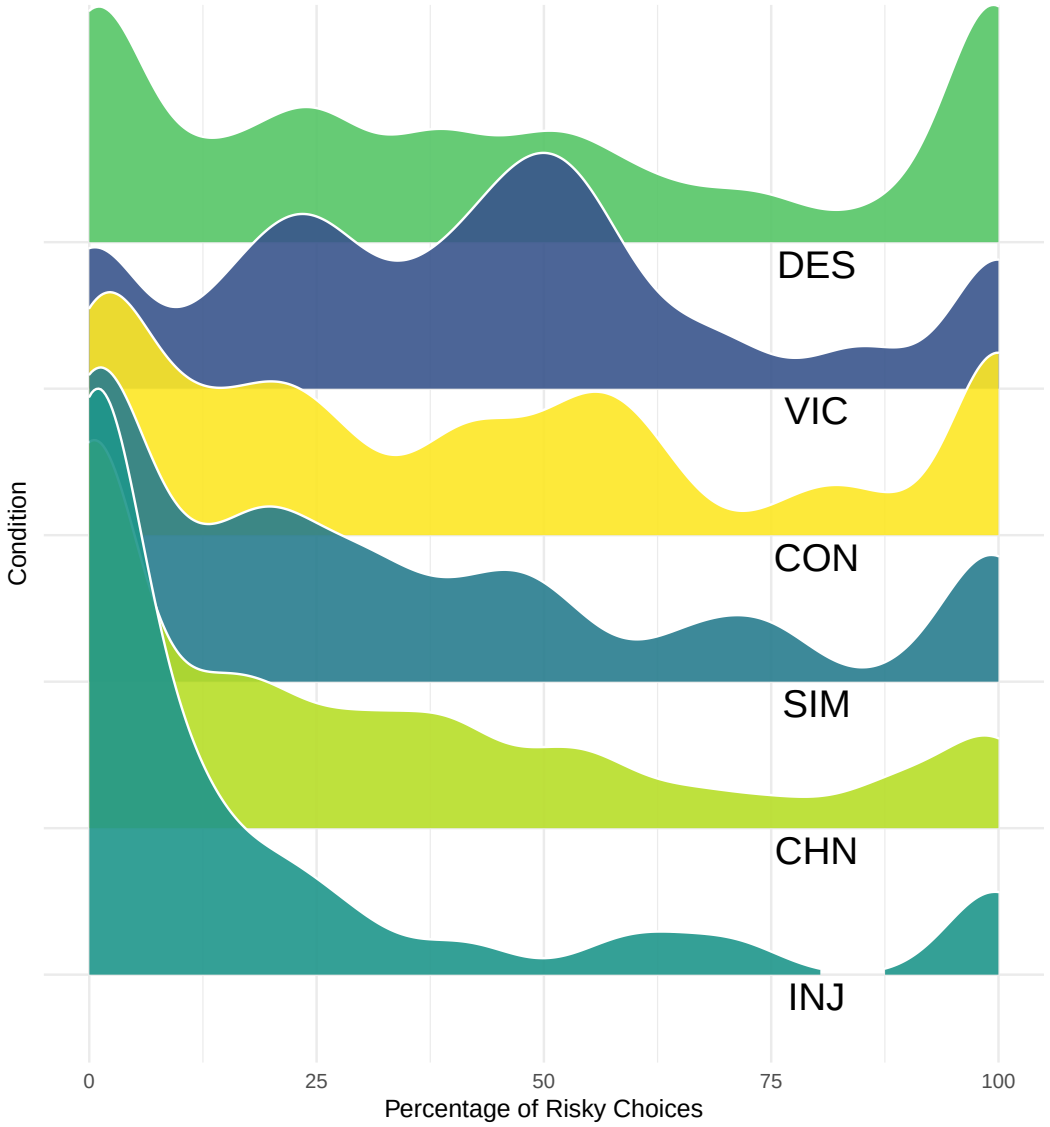


**Figure E66**

*Distribution of percentage of risky choices for all experimental conditions (DES: descriptive-norms, VIC: vicarious learning, CON: control condition, SIM: simulator, CHN: chain condition, INJ: injunctive-norms).*

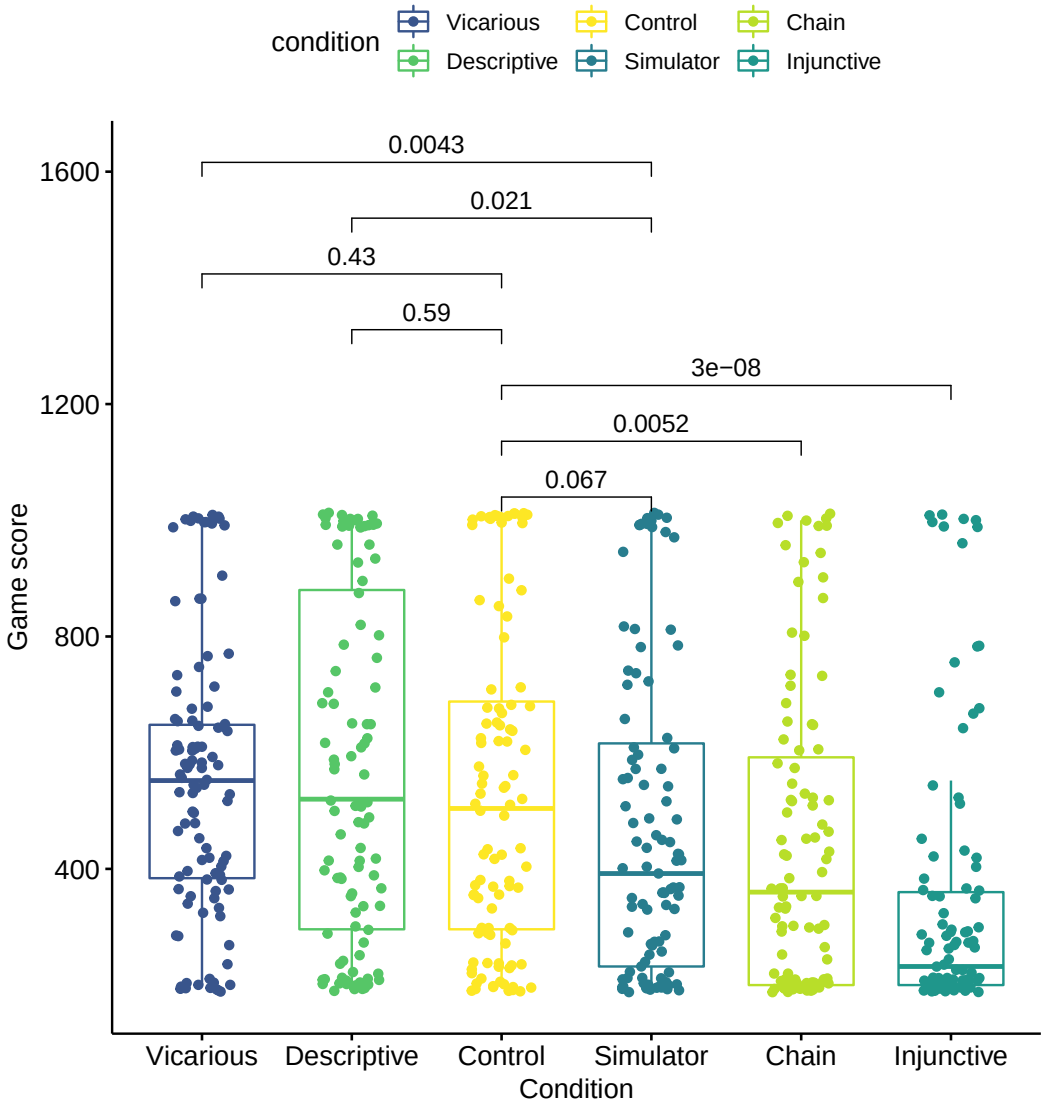


**Figure E67**

*Game results across conditions in Study 2: Jittered points within the box-plots represent individual results, p-values of selected comparisons are shown.*

## E11.4 Behavioral categories and predictor variables

### *E11.4.1 Behavioral categories across conditions*

We categorized behavior into four distinct types. Figure E68 shows the distribution of four behavioral strategies across conditions, defined as:

1. AllG: Chooses action G in every round.

2. H-G: Starts with action H and switches once in the course of the game.

3. Switching: Switches at least twice in the course of the game or starts with action G and switches once.

4. AllH: Chooses action H in every round.

Type "H-G" is distinguished from type "Switching", as only the behavior shown by the former type can be justified by our theoretical simulation results: Taking H actions before G actions has been shown to produce favorable results, switching back to H has been shown to produce expected losses.

Comparing the distribution across conditions, one can observe that the highest proportion of H-G-strategies occurred in the vicarious learning condition. This corresponds to a property of the presented graph based on the previous game: The graph exhibits an increasing slope over time with the first rounds showing low proportions of purple participants and later rounds showing much higher values. Some participants judging the current-round risk based on the other group's experience, might have decided to reverse course at some point during the game. Switching-strategies were prevalent across all conditions, the smallest proportion occurred in the injunctive norms condition that simultaneously produced the largest proportion of AllG-patterns.

In the following sections, we investigate the relationship between strategy type and a selection of additional scales and variables, focusing on those that successfully predicted the number of risky choices in both studies.

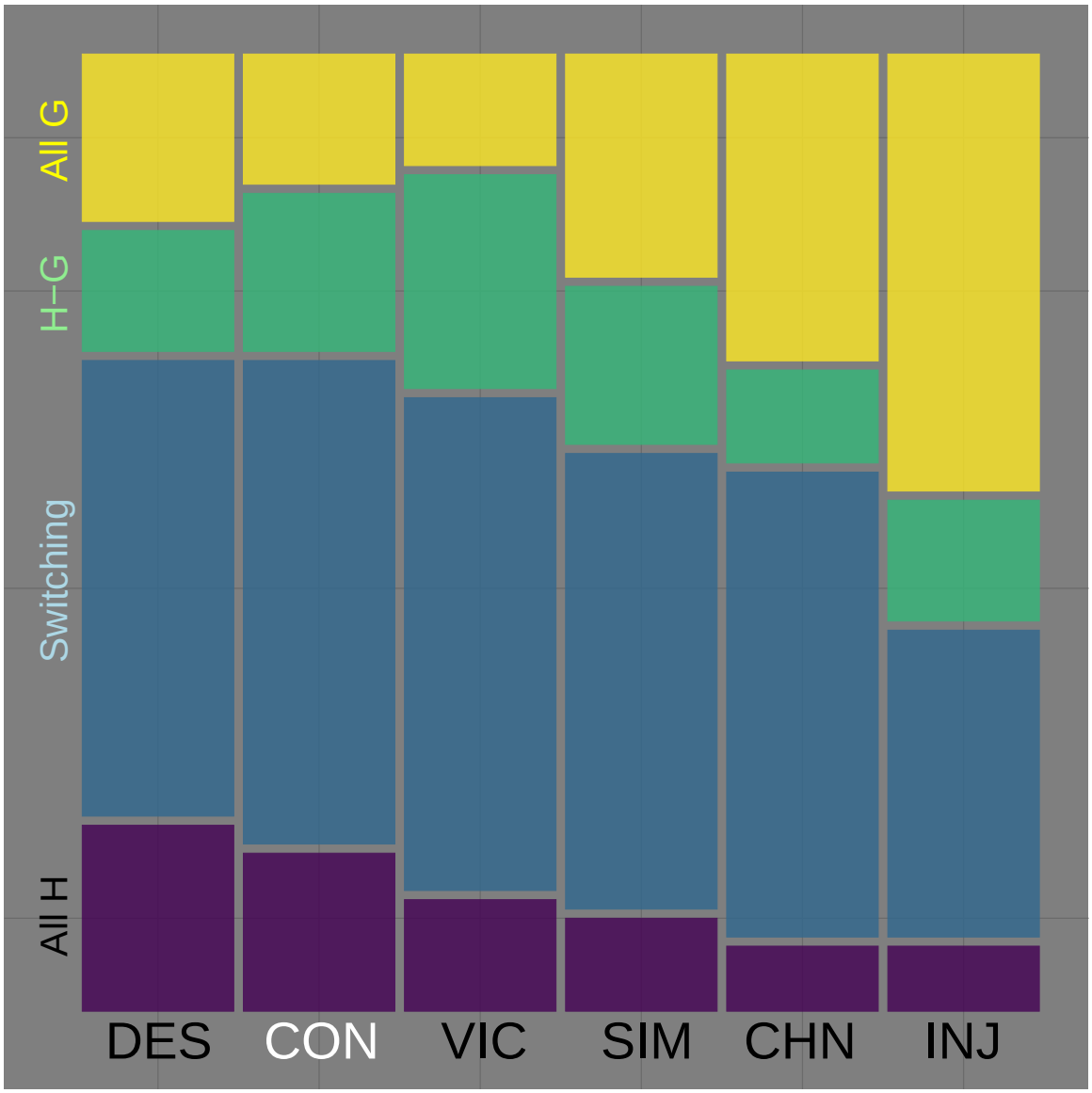


**Figure E68**
*Mosaic plot of the proportion of behavioral patterns observed in the six experimental conditions (DES: descriptive-norms, VIC: vicarious learning, CON: control condition, SIM: simulator, CHN: chain condition, INJ: injunctive-norms).*

### *E11.4.2 Behavioral categories and social value orientation*

We noted above that strategy choice in the transmission game should be influenced by social preferences, as higher individual scores imply lower scores for other group members, and there is a social dilemma at the heart of the transmission game.

Consistent with this interpretation, we observed clear differences between behavioral types in terms of social value orientation (see Fig. E69). The SVO angle of the SVO slider measure (Murphy et al., 2011) represents the relative weights put on the outcomes of the decision maker and another person, with higher values corresponding to prosocial preferences, and lower values to individualistic or even competitive choices (placing a negative value on the other person's outcome).

The highest proportion of prosocial values was found in the AllG-group, the lowest in the AllH-group. Competitive values were only seen in the AllH-group together with the highest proportion of individualistic values. The H-G group fell in between the extreme groups regarding both proportions. The Switch-group had a higher proportion of untypical values, but resembled the AllG-group more than the AllH-group.

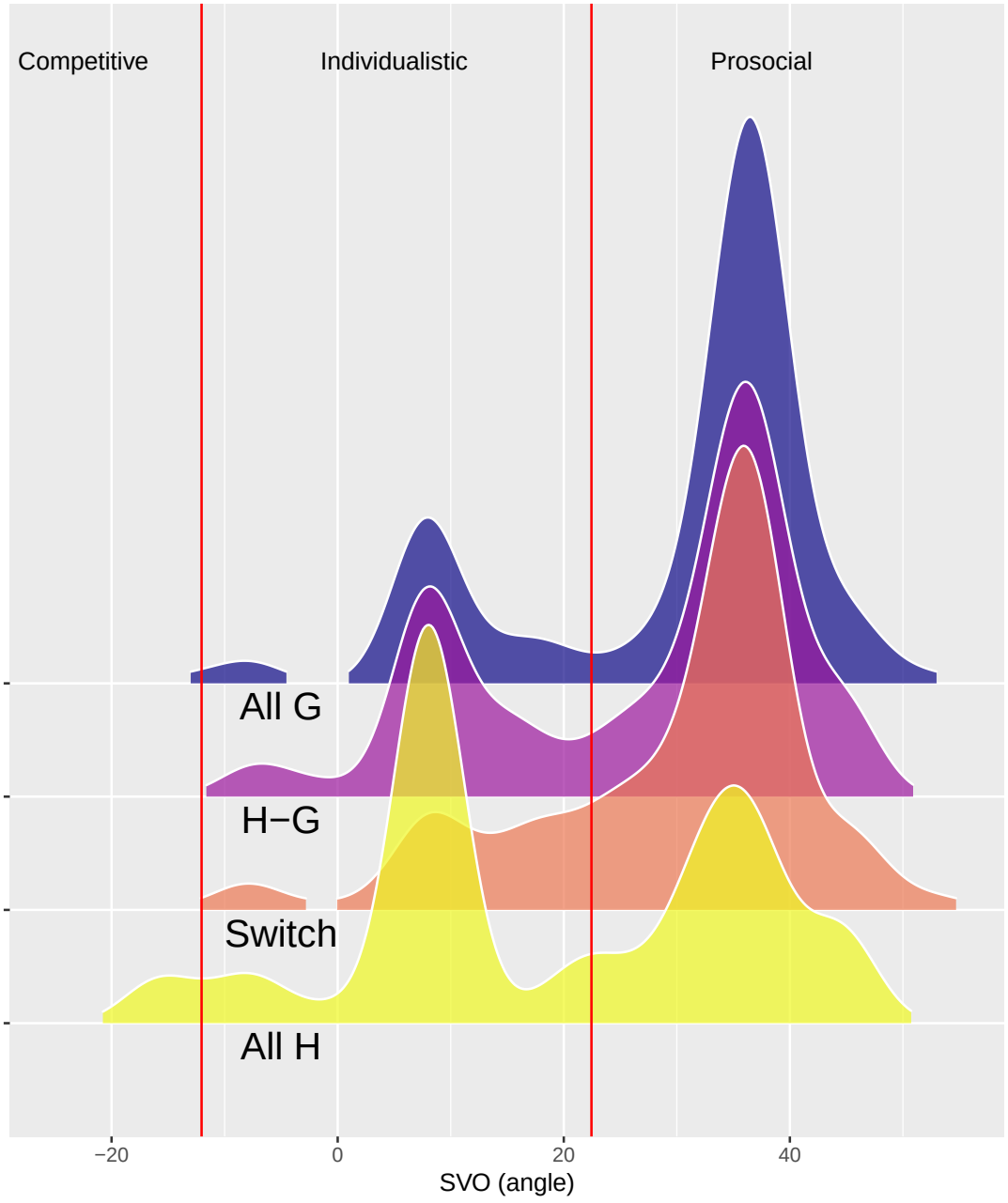


**Figure E69**
*Distribution of SVO angles for the four behavioral types: Red lines mark thresholds for regions considered competitive, individualistic, and prosocial preferences.*

As reported in **SA2**, SVO was positively correlated with worries about the coronavirus ($r = .11$), compliance with mitigation measures ($r = .11$) and negatively with the acceptance of tradeoffs ($r = -.19$).

### *E11.4.3 Behavioral categories and risk preference*

In previous sections, we referred to H-actions as risky actions. While scoring more points and increasing a player's expected outcome (if every other action were kept constant), they increase interaction risks and the chance to end the game with zero bonus money.

If this structure is transparent for participants we would thus expect a relationship with risk preference. Consistent with this assumption we see differences in risk preference—measured with the general risk item (Dohmen et al., 2011)—between the four behavioral types (see Fig. E70).

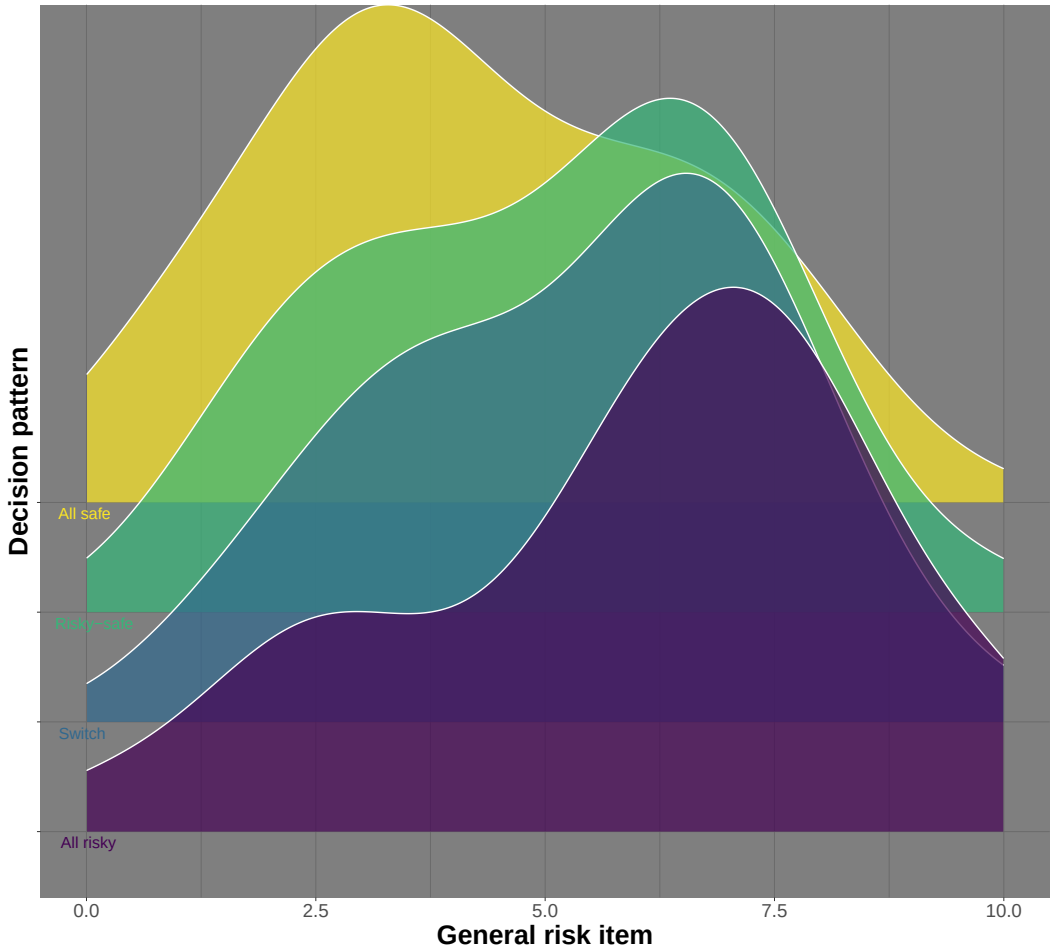


**Figure E70**
*Behavioral pattern and risk preference: Distribution of values across the four behavioral types.*

The AllG-type was the only group with a modal response below the mean scale point. The AllH-type had the smallest number of members that indicated a low degree

of risk preference. The two remaining groups were closer to the AllH-type than to the AllG-type, but their distribution was less shifted to higher values than for AllH.

Consistent with the relation to game results, we observed a positive correlation of the risk preference measure with the willingness to accept tradeoffs ($r = .13$), and a negative correlation with compliance with social distancing measures ($r = . - 13$; see **SA2**).

In summary, the distribution of risk values conforms to the idea that players choosing riskier actions had a higher appetite for risk and see themselves as active risk-takers in other situations, as well.

#### *E11.4.4 Behavioral categories and cognitive reflection*

Cognitive reflection was measured with four items, resulting in scores between 0 and 4. The mosaic plot in Figure E71 shows the distribution of scores in the four behavioral groups. It is easy to see that higher scores were more prevalent in the AllG and H-G groups, and the highest proportion of top scores (and the lowest proportion of 0-scores) was found in the H-G group.

When discussing the simulation results, we noted that identifying the optimal temporal placement of H-actions might require cognitive sophistication and at least the willingness to reflect on the advantages and disadvantages of different patterns. The comparison of the H-G and Switch groups shows higher cognitive reflection scores in the group making advantageous choices.

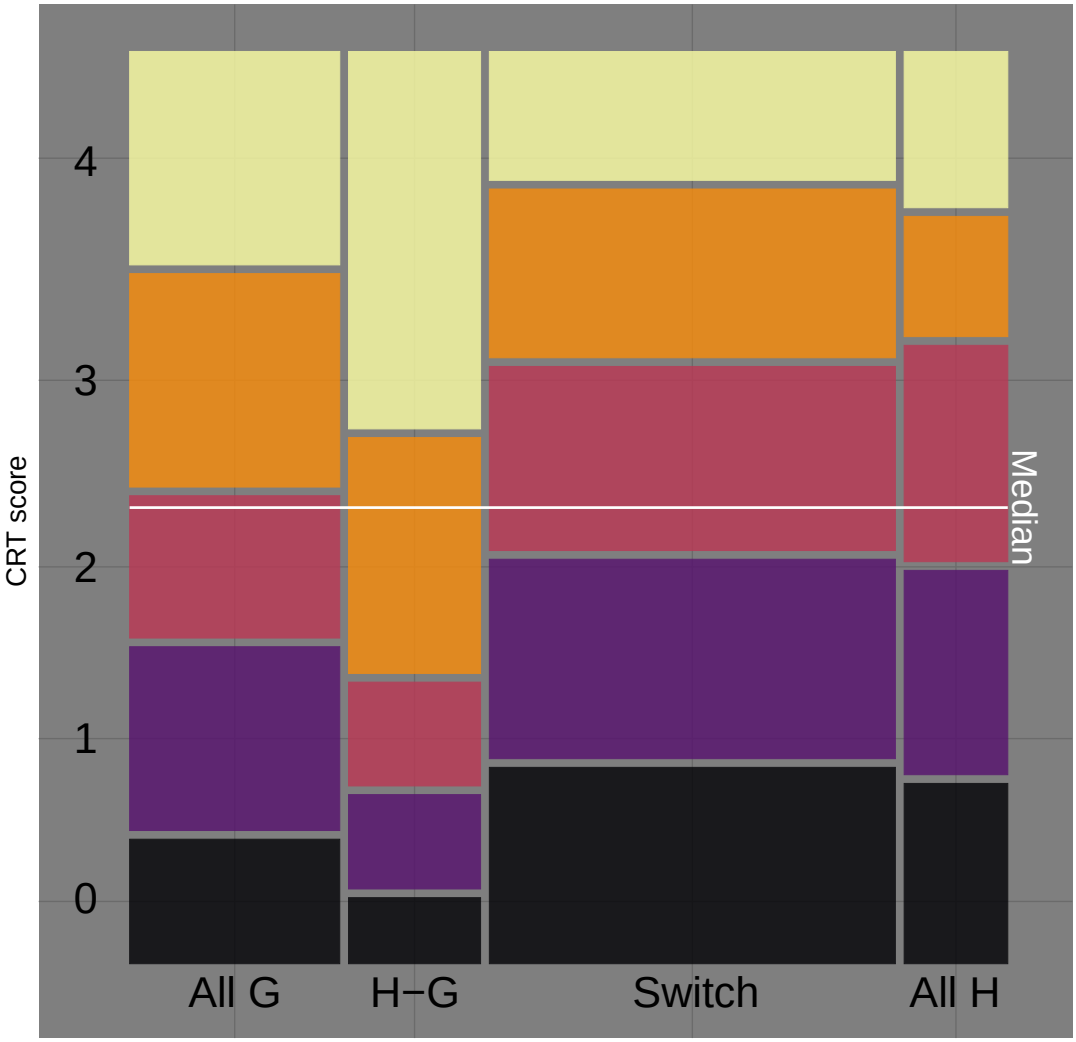


**Figure E71**
*Behavioral pattern and CRT score: Mosaic plot.*

## E11.5 Postquestionnaire-results

The postquestionnaires consisted of several parts. All participants responded to the first questionnaire, measuring expectations and general game motivation. Participants in the intervention conditions also responded to a second postquestionnaire with additional, intervention-specific questions.

### *E11.5.1 Expectations*

Participants were asked to estimate the proportion of purple participants after several rounds and the proportion of H-actions chosen in several rounds. Figure E72 plots individual estimates and fit lines for each condition.

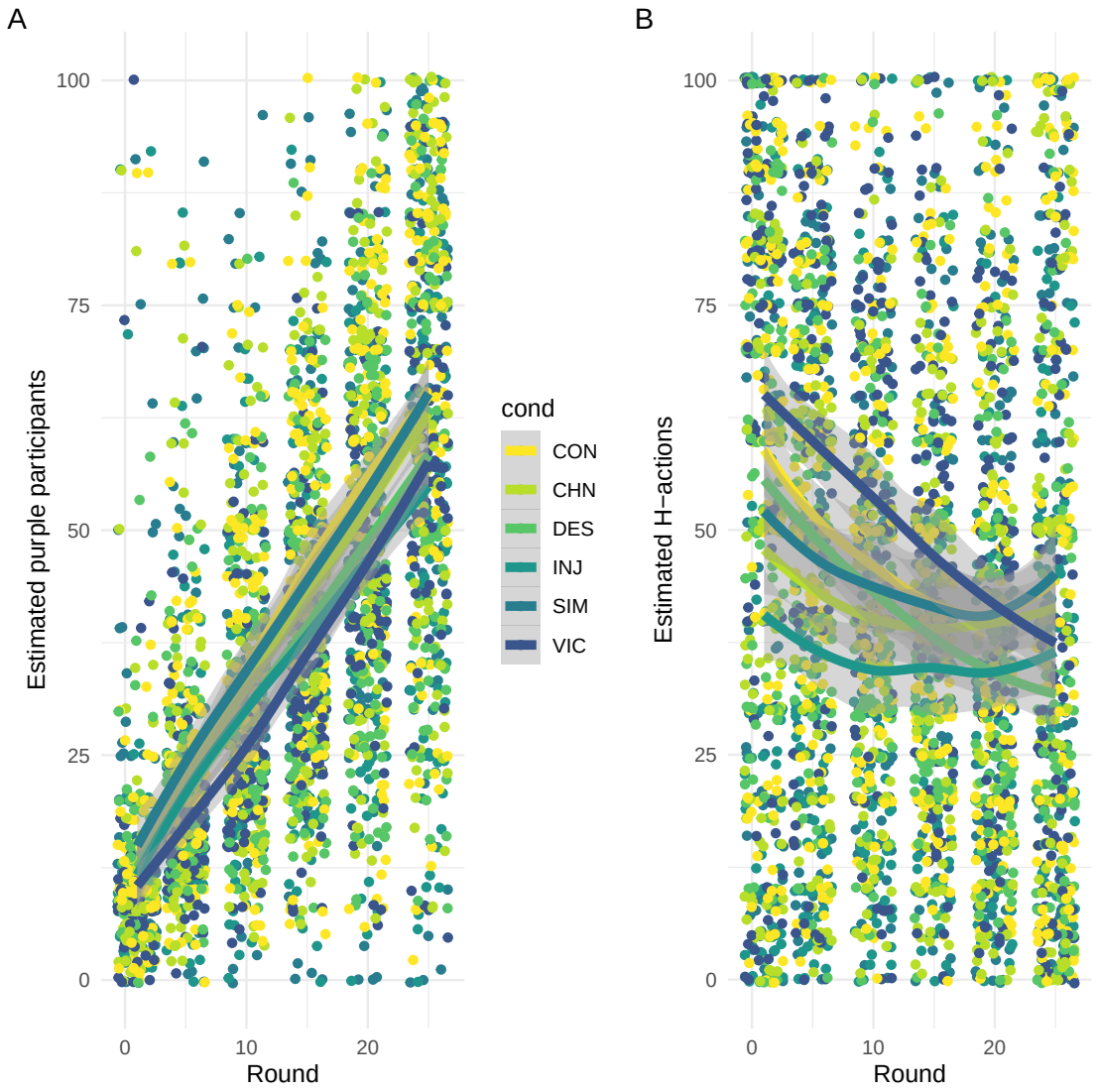


**Figure E72**

*Postquestionnaire expectations: Plots show individual estimates (points) and trend lines split by condition (DES: descriptive-norms, VIC: vicarious learning, CON: control condition, SIM: simulator, CHN: chain condition, INJ: injunctive-norms) with condition indicated by color for (A) estimates of the number of purple participants after rounds 1, 5, 10, 15, 20, 25, and (B) estimates of the proportion of H-actions chosen in rounds 1, 5, 10, 20, and 25.*

Changes in the number of purple participants are estimated to be linearly increasing over time (at least based on the average estimates). Points in the upper left and lower right of the diagram correspond to participants who might have estimated the number of blue participants, instead. The differences between conditions were relatively small (and underestimated the expected differences, see below). Slightly lower rates were estimated for the vicarious learning, injunctive norms, and descriptive norms conditions. These differences do not in general correspond to observed differences in behavior. Both the simulator and the chain reaction intervention increased the estimates for purple participants relative to the other interventions.

In contrast, the highest proportion of risky actions was expected in the vicarious learning conditions, but with a near-linear reduction and lower levels in the final rounds. Consistent with the actual behavior, the lowest proportion of H-choices was estimated in the chain reaction and injunctive norms conditions. Differences in estimated actions are larger than differences in estimated color switches. A comparison with actual expected color switches will be made in the simulation section, below.

#### *E11.5.2 Expectation for self versus others*

We directly asked for an estimate for the probability of switching to purple in the game. We also asked for the number of purple players after different rounds in the game including after the final round. Comparing this final estimate for all players (which—converted into the relative frequency—we take as a probability estimate for the average player) with the estimated probabilities for the player themselves, allows us to judge the degree of relative optimism or pessimism. The comparison is plotted in Figure E73.

We computed a mixed ANOVA with self/average as a within-factor and condition as a between-factor. We found a main effect for self/average ($F(1, 594) = 116.57$, $p < .001$, $\eta^2 = .16$), a weaker main effect of condition ($F(5, 594) = 2.46$, $p = .03$, $\eta^2 = .02$), and a non-significant interaction ($F(5, 594) = 1.76$, $p = .12$, $\eta^2 = .01$). Clearly, participants in all conditions described themselves as more likely to stay blue than the average player, which constitutes relative optimism.

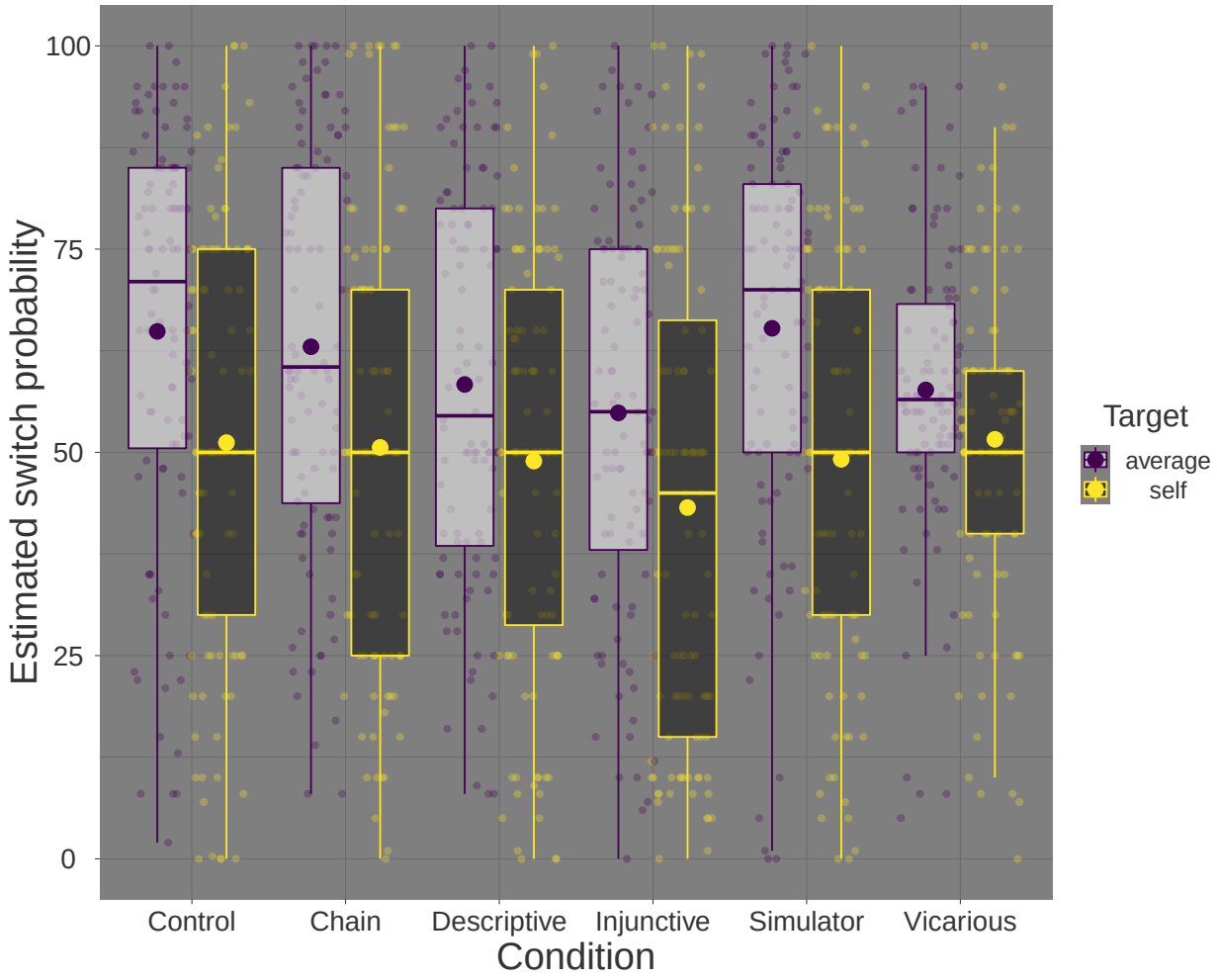


**Figure E73**

*Expected probability of switching to purple for oneself and for the average participant: Boxplots for self-estimates and average estimates split by condition. Larger dots represent means, small dots individual estimates.*

### *E11.5.3 Motivations (general postquestionnaire)*

The general postquestionnaire asked for participants' motivation in the game with eleven questions (on a scale from 1–5, the name at the start of each line is only used in the figures below):

- MaxBonus: I wanted to make as much bonus money as possible.
- MoreBonus: I wanted to make more bonus money than other players.
- Responsible: I felt responsible for other players.
- MakeSwitch: I wanted to make other players switch color.
- AfraidSwitch: I was afraid to switch color in this game.

- Anticipate: I tried to anticipate what others were doing in this game.
- FollowGut: I followed my gut in this game.
- TakeRisk: I wanted to take some risk in this game.
- ProtectOthers: I wanted to protect others in this game.
- DidNotCare: I did not care at all what happened in this game.
- AllReceive: I wanted that the entire group receives as much money as possible.

Relationships between these questions and game behavior are split up into two Figures (see Fig. E74 and Fig. E75).

As in Study 1, most participants indicated that they wanted to have the maximum bonus possible; fewer participants wanted to have more bonus money than others. Both answers to these questions were significantly correlated with the number of H-choices in all groups.

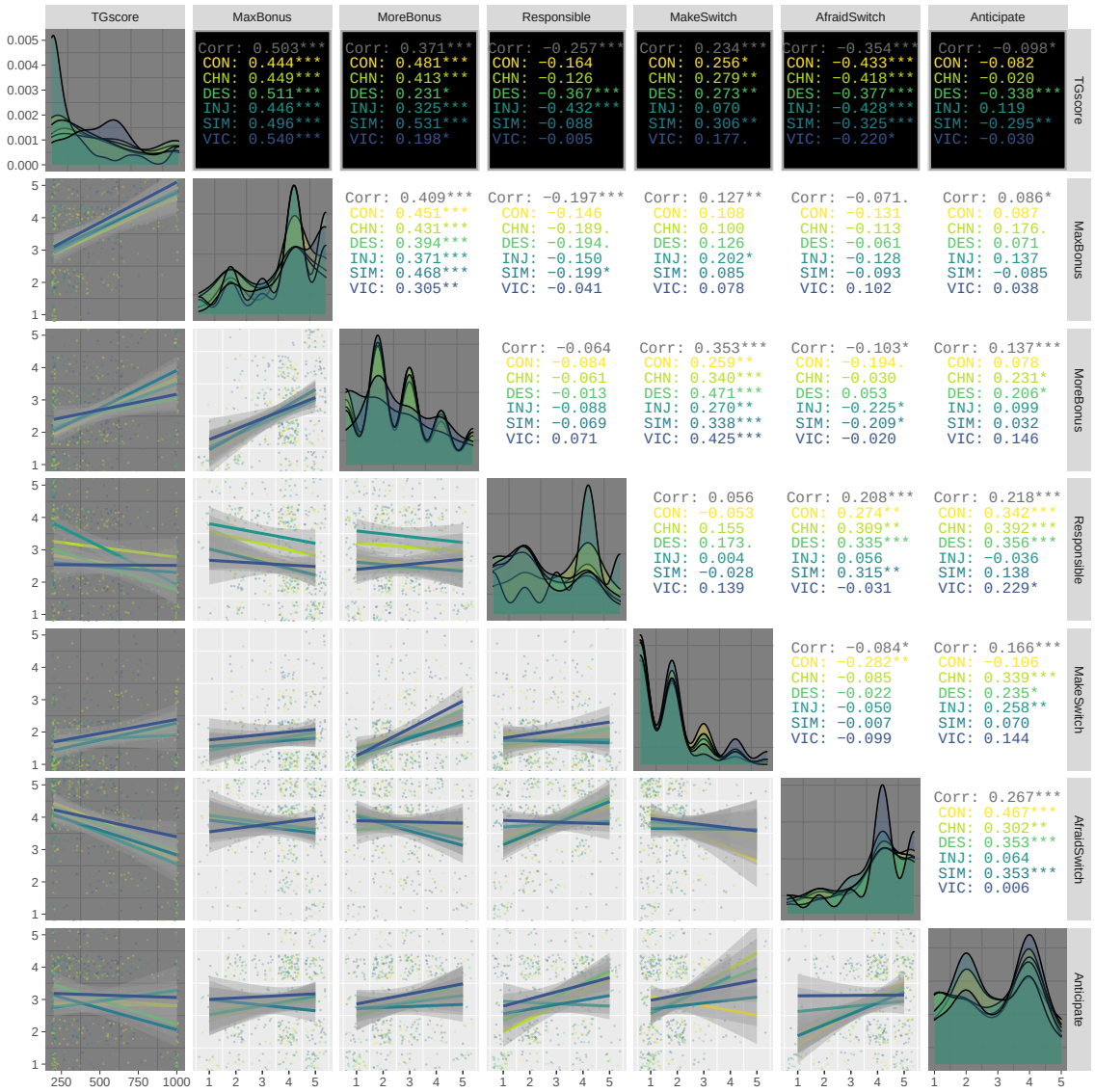


**Figure E74**

*Relationship between game results and postquestionnaire responses (part 1): Distribution of values (diagonal), general and group-specific bivariate correlations (above diagonal) and scatter plots with linear fit line (below diagonal).*

Feeling responsible and being afraid to switch were correlated negatively with the number of H-choices. Most participants were afraid to switch colors; this was more frequent in the vicarious learning condition. The negative correlation for this variable was found in each of the six groups. Feelings of responsibility for others were most pronounced in the injunctive norms condition (in which the highest negative correlation with risk-taking was observed). The second highest (and still significant) correlation was observed in the descriptive norms group. There was virtually no relationship in the vicarious learning condition, and reduced negative relationships in the rest of the groups.

Few participants responded that they wanted to make others switch color, and we observed positive correlations between this intention and the number of H-choices in all groups, most pronounced in the simulator condition and least pronounced in the injunctive norms condition. Participants were split in most conditions between answering that they had anticipated others' actions or not. There was a weak negative correlation with H-choices across all groups, with higher values in the descriptive norms and simulator condition (both of which explicitly focused on the behavior of others, either in the form of direct information or in the form of simulation assumptions).

The strongest predictor[12] of H-choices in the game was the response to the risk-taking question. In all groups high correlations between admitted risk-taking and the number of H-choices were found. Participants who responded affirmatively that they were following their gut tended to choose H more often, those who were willing to protect others tended to choose H less often (similar to responsibility). Few participants admitted that they did not care about their result and those tended to choose H more often, whereas those who wanted everyone to receive a similar outcome chose H less often. In general, the pattern of postquestionnaire responses aligned with behavior, with some notable differences between conditions.

[12]This is meant in the statistical sense, the postquestionnaire was obviously presented after the game.

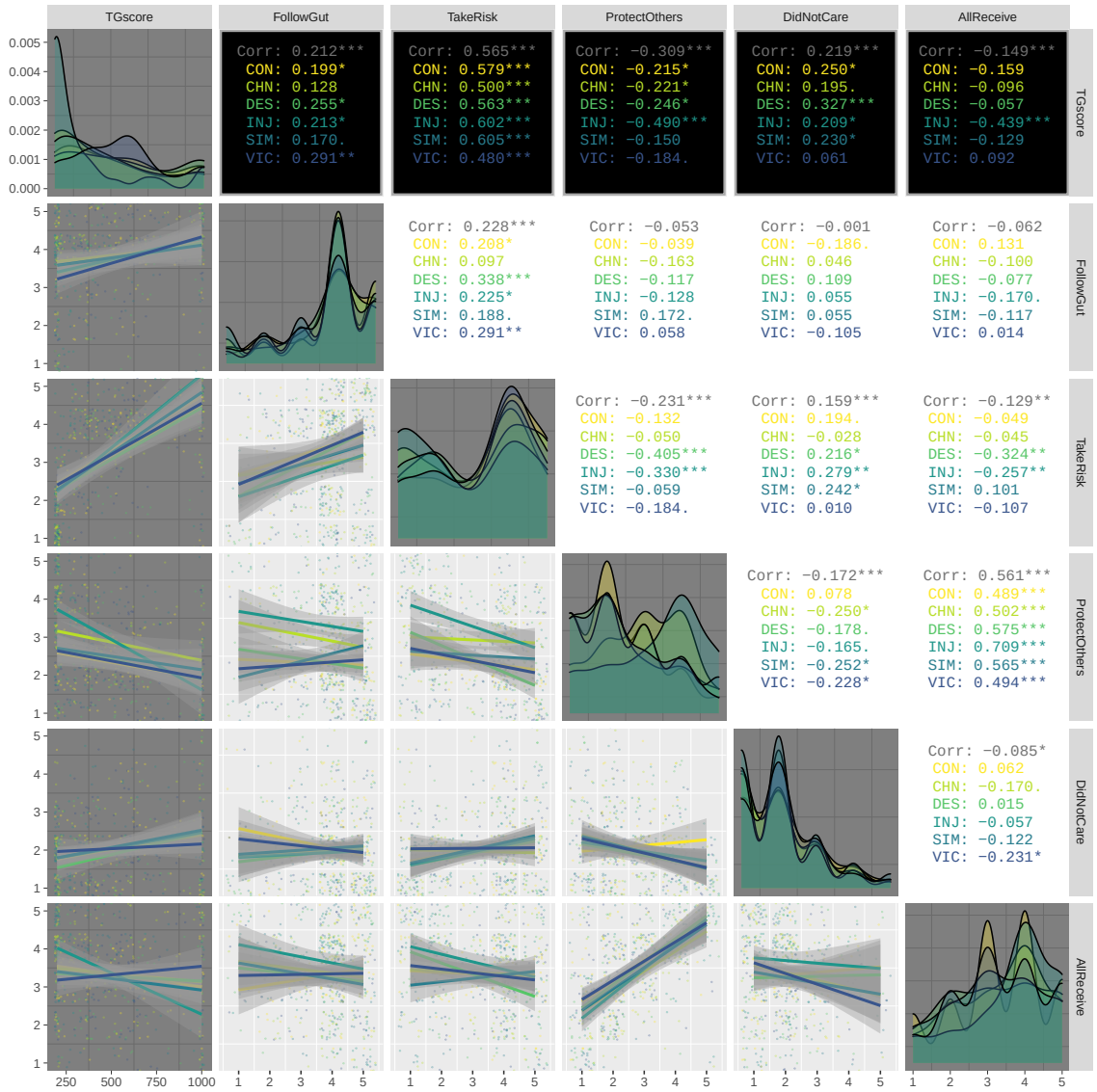


**Figure E75**

*Relationship between game results and postquestionnaire responses (part 2): Distribution of values (diagonal), general and group-specific bivariate correlations (above diagonal) and scatterplots with linear fit line (below diagonal).*

### *E11.5.4 Motivations (specific postquestionnaires)*

To analyze the specific postquestionnaires across conditions we grouped similar questions together (see Fig. E76).

Participants in general indicated to like and understand the five interventions (vicarious learning, descriptive norm, injunctive norm, chain reaction, and simulator); they felt influenced by the experienced intervention and did not consider it a waste of their time. Few participants believed that they had reduced the number of G-choices in response to the intervention, with some of them simultaneously indicating that they had also reduced the number of H-actions[13].

There was a small negative correlation between liking the intervention and H-choices, mainly driven by the injunctive norms and simulator condition. Similarly, there was a negative correlation between participants' indicating that they had understood the intervention and H-choices. The strongest relationship was found in the simulator condition, which might have required the most cognitive effort to understand, followed by the vicarious learning condition (which also featured a graphical presentation), and the chain reaction condition (also a relatively complex intervention).

There was a negative correlation between indications that the interventions had influenced decisions in the game and the number of H-choices. A strong relationship was only observed in the injunctive norms and the simulator condition. Participants who felt that the intervention was a waste of their time tended to choose a larger number of H-actions. This was especially true for the injunctive norms condition.

The two questions asking participants to imagine that the intervention had not taken place seemed to have caused interpretation problems for some participants. The general direction of the relationship was as expected: Participants who affirm that they would have chosen more H-actions without the intervention selected fewer H-actions in the game. A strong relationship was only found in the simulator condition, though[14].

Finally, participants who indicated that they had been influenced by the intervention also tended to indicate that they had reduced the number of H-choices. These correlations were strong for all conditions. A positive correlation with being influenced and reducing the number of G-choices was only present in the vicarious learning and descriptive norms condition. This may indicate that these two interventions may have had a somewhat ambiguous effect. A subgroup of participants might have increased the number of H-choices in reaction to these interventions.

[13] A charitable interpretation would be that the influence was different in different rounds.

[14] There was also an unexpected positive correlation for agreement with the suggestion that without the intervention G-actions would have been chosen less frequently and H-choices in the chain reaction condition. Likewise, there was no strong correlation between the two counterfactual questions, which seems counterintuitive.

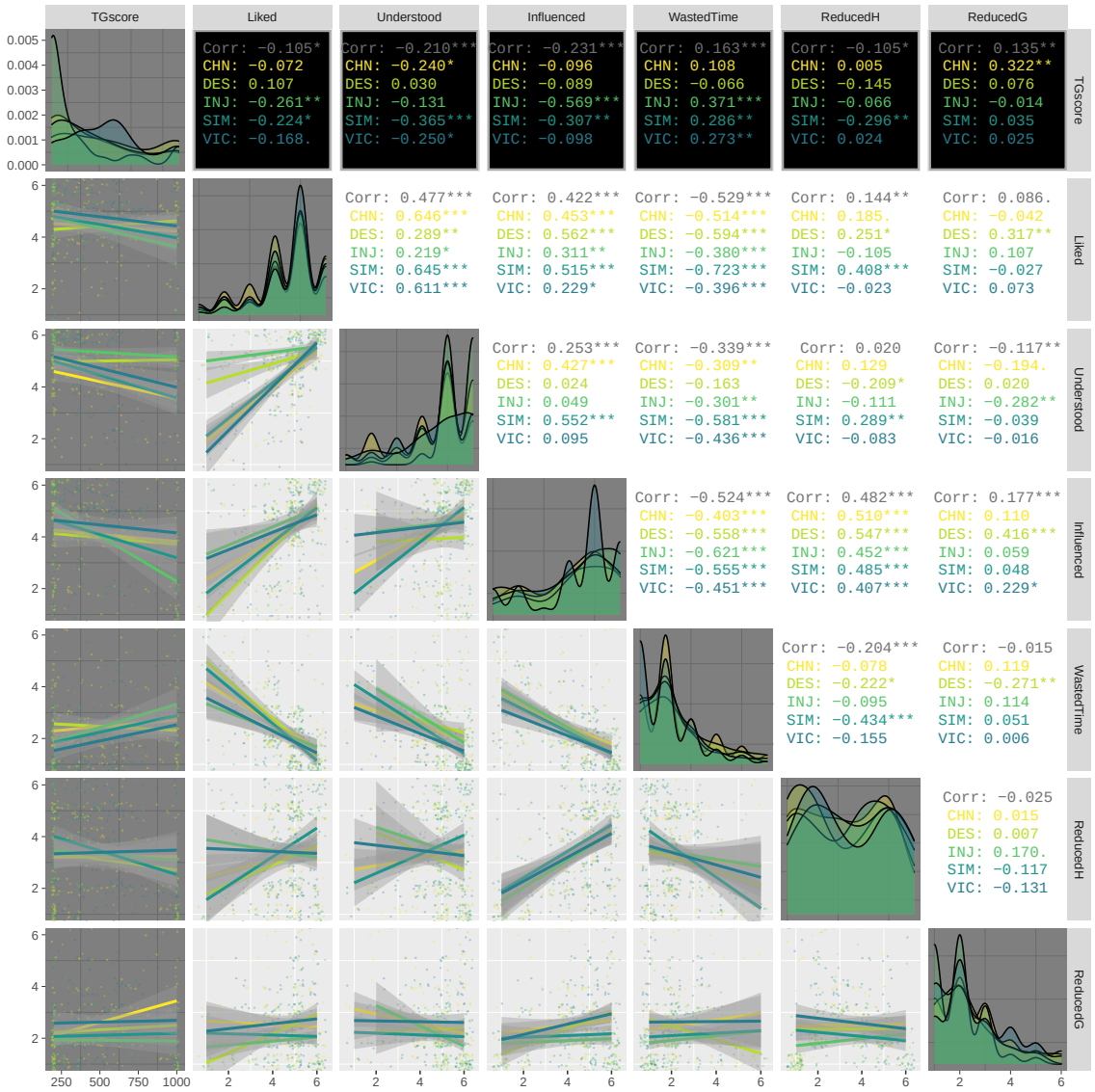


**Figure E76**

*Relationship between game results and condition-specific postquestionnaire responses: Distribution of values (diagonal), general and group-specific bivariate correlations (above diagonal) and scatter plots with linear fit line (below diagonal).*

Participants' understanding of an intervention was negatively correlated with considering it a waste of their time and highly positively correlated with liking it. This might help to explain the differences between interventions. Both the simulator and the chain reaction intervention were cognitively demanding, and not every participant might have positively engaged with them. These interventions seemed to have worked for some, but not all of the participants. In contrast, understanding the injunctive norms intervention is simpler and straight-forward.

It is possible that the vicarious learning condition signalled to some participants in the beginning that the risk was low and that H-actions could be relatively safely chosen. This corresponds to the high initial number of H-actions seen in this condition. The descriptive norms condition, in contrast, might have signalled both that a substantial number of other participants had chosen H-actions in the past and that a similarly substantial number had chosen G-actions. This intervention might therefore have had an ambiguous effect with some participants finding reasons and justifications for choosing H and others for choosing G.

### E11.6 Predictor variables and game results

Figure E77 presents correlations between predictors and game results. There was a significant positive correlation between general risk-taking and the number of H-choices in every condition (it is slightly weaker in the chain reaction condition). The CRT scores were consistently and negatively correlated with the number of H-choices. This correlation was strongest in the cognitively most demanding interventions: chain reaction and simulator. There was a negative correlation between SVO angles (prosocial preferences) and the number of H-choices. This relationship was non-existent in the chain reaction and descriptive norms condition, and strongest in the injunctive norms and vicarious learning condition.

The relationships between personality scales and game behavior that we had observed in Study 1 did not materialize in Study 2. There was no systematic pattern for the honesty and the emotionality scale. We observed a negative relationship between emotionality and the risk item, a positive relationship between honesty/humility and SVO, in basically all conditions.

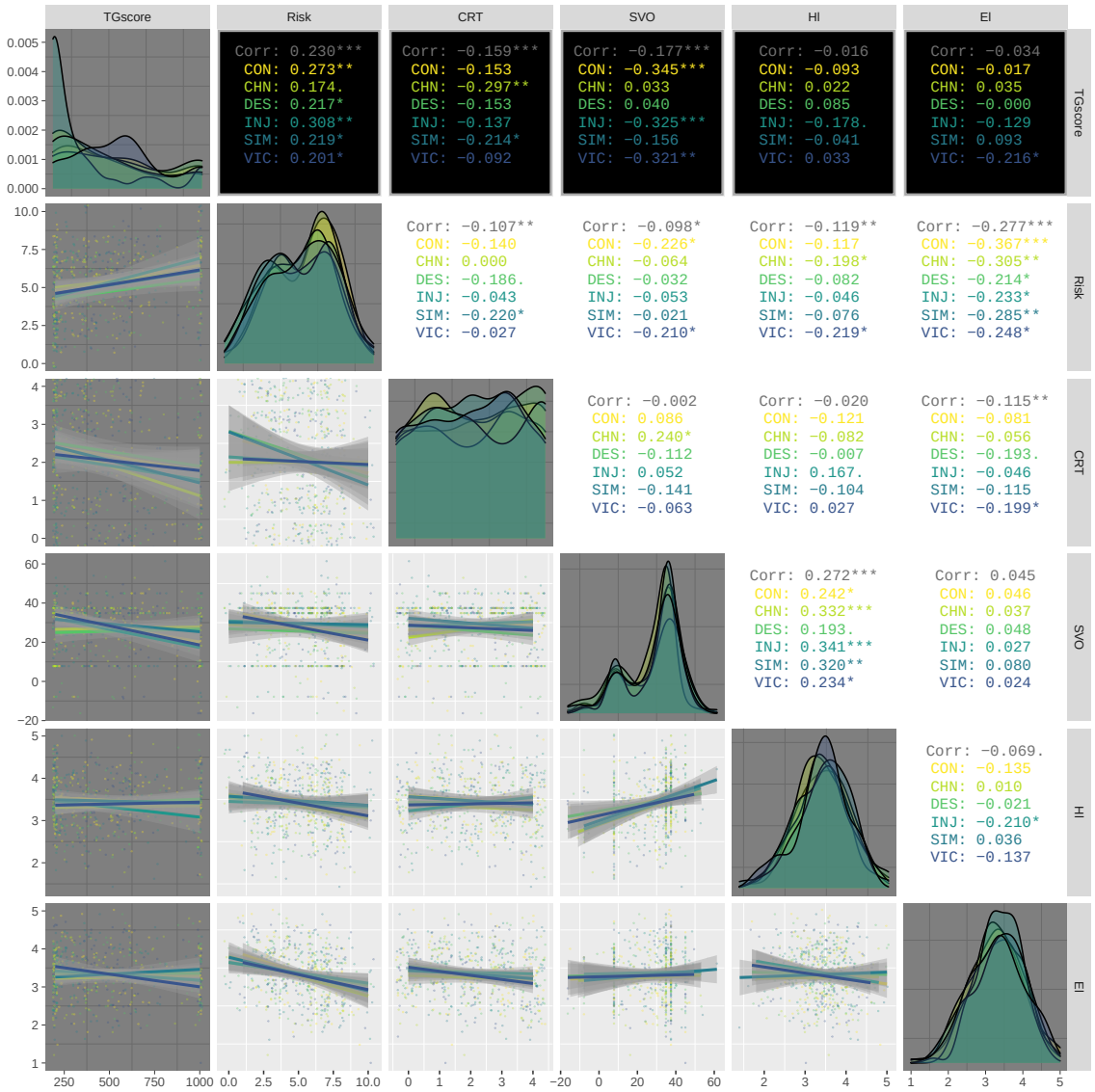


**Figure E77**

*Predictor variables and game results: Distribution of values (diagonal), general and group-specific bivariate correlations (above diagonal) and scatterplots with linear fit line (below diagonal).*

### E11.7 Game results, political attitudes and COVID-19 related scales

In Study 1, we observed strong relationships between political preferences and COVID-related scales, but no substantial relationship of either variable with game behavior.

Replicating this analysis for Study 2, we found significant correlations between conservatism and preference for Donald Trump and game behavior. A comparison between conditions demonstrates that this relationship was not robust across intervention conditions and mostly restricted to the control condition. Likewise, we found no consistent relationship between COVID-19 related scales and game behavior.

In contrast, we replicated the strong relationship between political attitudes and scales. Conservatism and preference for an electoral win for Donald Trump was correlated with lower compliance with virus counter-measures, reduced worry about the consequences of COVID-19 and a stronger endorsement of tradeoffs.

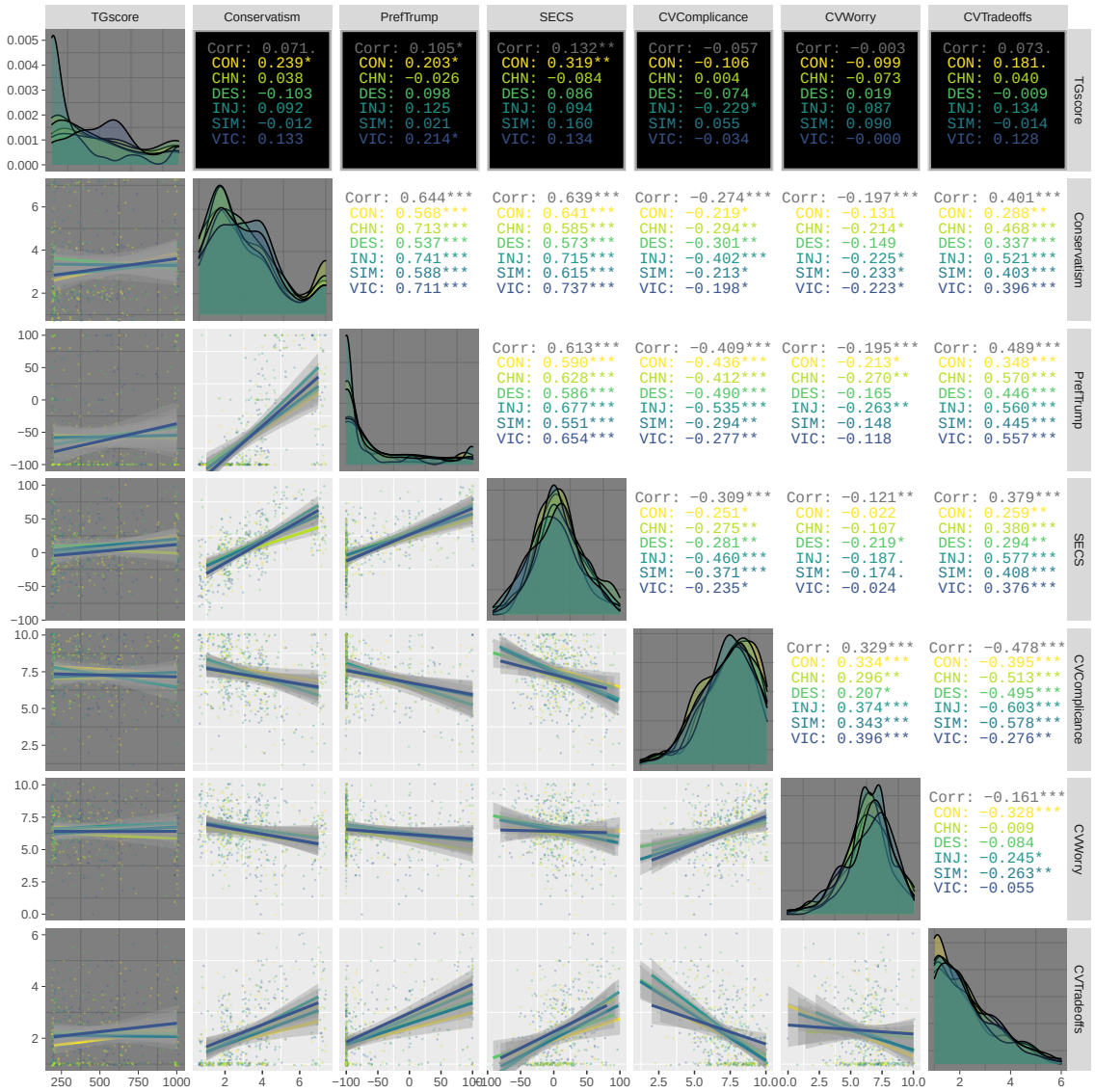


**Figure E78**

*Game results, political attitudes and COVID-19 related scales: Distribution of values (diagonal), general and group-specific bivariate correlations (above diagonal) and scatter plots with linear fit line (below diagonal).*

## E11.8 Qualitative Data

An independent coder categorized comments about the game and hypothetical messages to players choosing G-actions and H-actions, respectively. The coder had no information about other variables, neither game behavior nor condition.

### *E11.8.1 Transmission game associations*

We asked participants whether the game reminded them of anything they had encountered before. Table E6 lists the non-exclusive categories[15] that we used to code responses and the frequency with which each category was coded. We combined some similar categories with low frequency counts after the coding was completed. For example, all but one participant who described a social dilemma, described the prisoner's dilemma.

Only a minority of participants (14.3% with the most permissive criterion) connected the game to the novel coronavirus or epidemiology in general. This is encouraging, as it means that participants did not perceive the paradigm as an obvious disease scenario—even at a point in time when they were experiencing a large number of virus-related surveys. At the same time, the connection between game and pandemic was not perceived as arbitrary. More than half (26 out of 51) of the participants who mentioned COVID-19 or the coronavirus by name—and some participants talking in more general terms about diseases—translated game elements back into virus-related entities or behaviors. The following are

[15] "No comment" is, of course, an exclusive category.

**Table E6**

*Categories for associations with the transmission game in open-format asnwers.*

| Main category | Subcategory | n | Percent |
|---|---|---|---|
| Virus | COVID-19 | 51 | 8.5 |
| | Virus (unspecific) | 4 | 0.7 |
| | General epidemiology | 31 | 5.2 |
| | Total | 86 | 14.3 |
| Economic Game | Social dilemma | 19 | 3.2 |
| | Other game | 16 | 2.7 |
| | Total | 35 | 5.8 |
| Game of chance | Lottery/gamble | 15 | 2.5 |
| | Unspecific | 5 | 0.8 |
| | Total | 22 | 3.7 |
| Novel | Not seen before | 94 | 15.7 |
| | Total | 94 | 15.7 |
| Other | Biology | 4 | 0.7 |
| | Finance/Business | 11 | 1.8 |
| | Unspecific | 130 | 21.7 |
| | Total | 145 | 24.2 |
| No comment | No comment | 246 | 41.0 |
| | Total | 246 | 41.0 |

examples of such responses:

- It reminded me of the choice between gain and risk many are facing now during the pandemic. Expose yourself at work or isolate and make no money.
- This seems to replicate the dilemma currently occurring with shelter-in-place orders and the individual decision for people to either stay at home or go out. You can play it safe and help everyone else out, or you can be risky to have the better potential bonus (participation in regular activities) but your risk of turning purple or not getting a bonus is increased, and also raising the odds other people turn purple as well.
- Reminds me of "Flattening the curve" with COVID.
- This game reminds me of mathematical epidemiological models in which a virus spreads through a group of people. I would guess that the purpose of this study is to gauge to what degree people making self servicing and dangerous decisions affects the population at large.

Comparing the number of virus-mentions across conditions, there was a slightly higher number in the chain reaction condition ($n = 18$) than in other conditions ($5 < n < 10$). Other associations included economic games (social dilemmas for the most part), games of chance, or a diverse range of experienced activities, exercises or demonstrations. A substantial number of participants (15.7%) explicitly stressed the novel character of the transmission game.

#### *E11.8.2 Messages to other players*

Our coding scheme for messages was inspired by the "might over morality" hypothesis Liebrand et al. (1986), which suggests that cooperators perceive behavior in a social dilemma as morally good vs. bad, whereas defectors primarily perceive behavior as strong vs. weak.

We coded the messages for G and H separately on four binary dimensions: strong (e.g., smart or powerful), weak, moral/altruistic, and immoral/selfish. In addition, we coded whether the message implied that the behavior was safe or risky (in two binary ratings). We also evaluated the message as either positive (1), neutral (0) or negative (-1). An example of a moral vs. immoral contrast is the message pair: "(G) Good job - that will make it more likely we'll all earn money, including you" vs. "(H) your greed is going to backfire for not only yourself, but also for the rest of us." In contrast, the weak vs. strong contrast is exemplified by "(H): They are brave and are risk takers, people like this can be good leaders." vs. "(G): Sometimes you have to take a risk in life to get something, you cannot be scared all the time in your life".

Some codes were used very infrequently. Only 15 participants (2.5%) considered G to be immoral and only 2 participants (0.3%) described H as moral or altruistic. Similarly, only 25 participants (4.2%) described action H as weak. Figure E79 and Figure E80 show how often messages to G-players and H-players were categorized into the more frequent categories, split by game behavior (Fig. E79) and game condition (Fig. E80).

Messages to H-players often described the behavior as selfish and immoral, but this did not depend on game behavior which is inconsistent might-over-morality hypothesis. Describing G as a moral or altrusitic option is more common in AllG-players than in

AllH-players (leftmost and rightmost points), but less than a quarter of AllG-players are categorized as using this characterization. It is more common to classify H as immoral than to classify G as moral.

A characterization as strong or weak is relatively rare, and characterizing G as a strong action is even more common than characterizing H in that way. Again, this is inconsistent with the idea that AllH-players use a distinction of strong vs. weak.

The general evaluation of these actons as observed in the written messages shows the expected result that AllG-players evaluate G-actions positively and H actions negatively. In contrast, AllH-players characterized H-actions as positive, and G-actions as (slightly negative). Players with a number of H-actions between 0 and 25 show intermediate positions with roughly equal and relatively neutral mean evaluations for 12–15 H-choices. This result is not trivial, as the coding of messages was done without access to game behavior information.

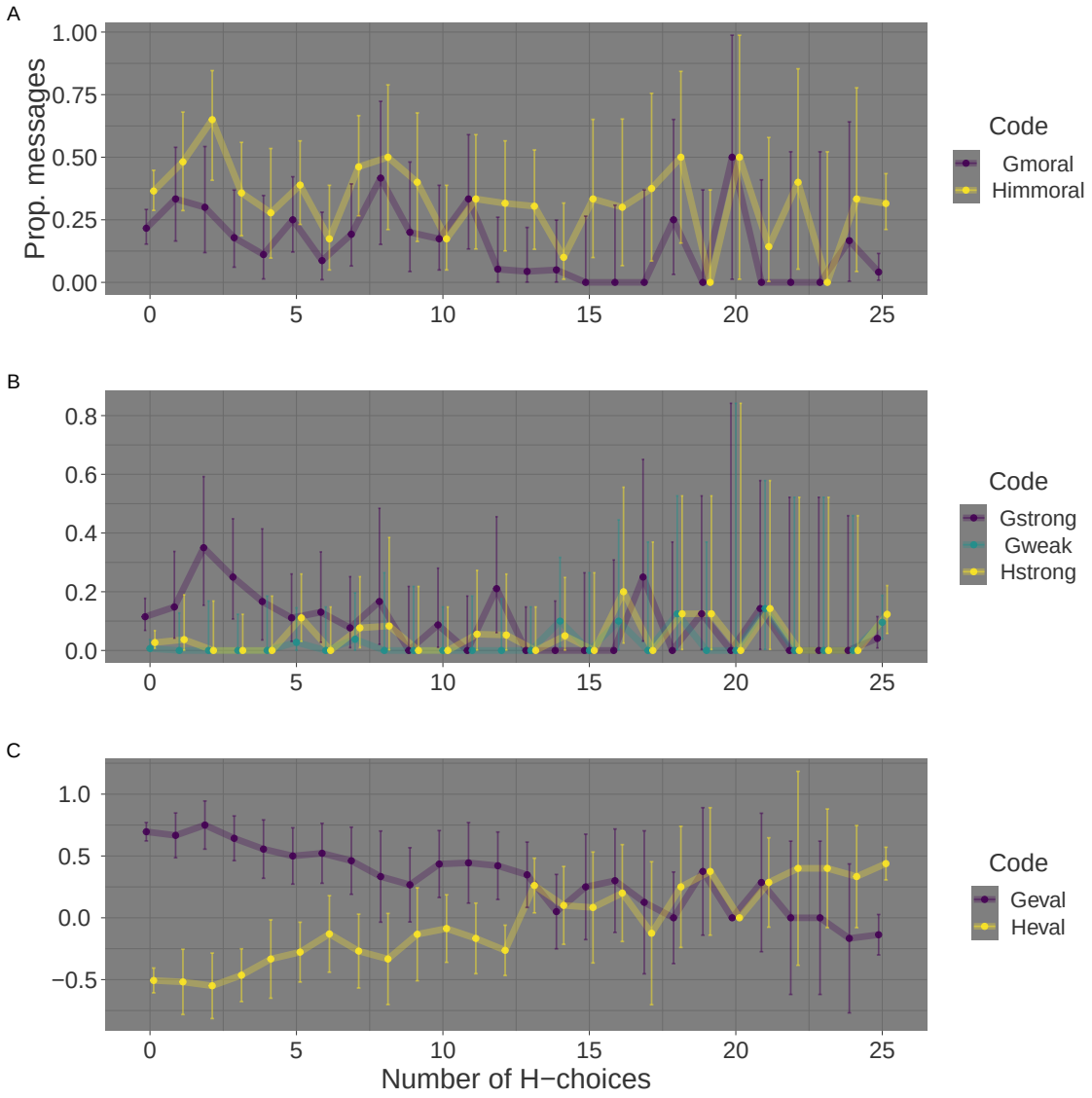


**Figure E79**

*Categorization of messages split by the number of H-choices in the game: Dots represent the proportions of messages categorized into the respective classes, error bars represent the 95% CI of the proportions. The three panels show proportions—for each number of H-choices—for (A) messages to G characterizing G-actions as moral/altruistic and messages to H characterizing H-actions as selfish/immoral; (B) messages to G coded as strong, messages to G coded as weak, messages to H coded as strong. (C) Mean evaluation of messages to G-players and H-players with 95% CI for the mean.*

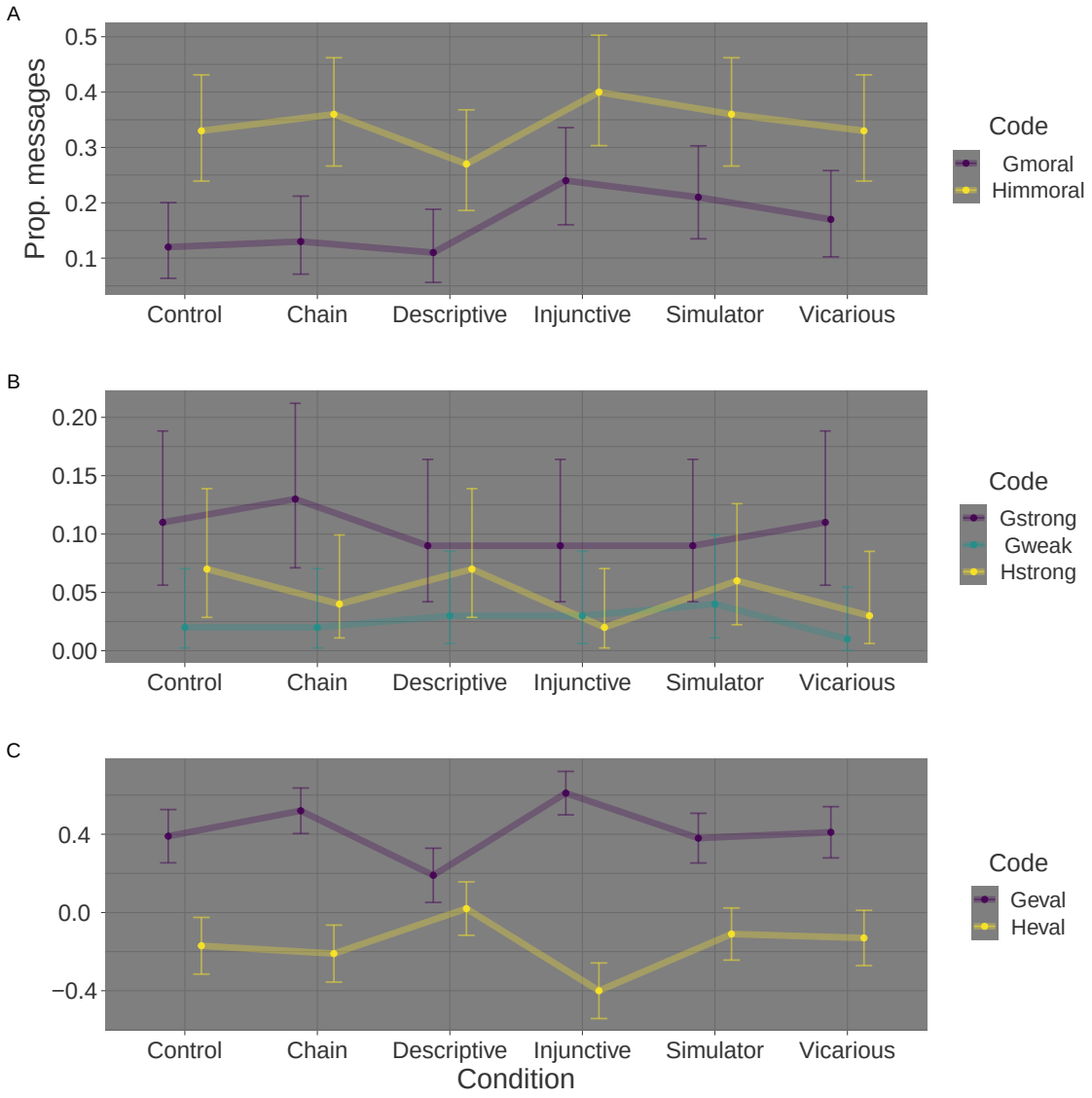


**Figure E80**

*Categorization of messages split by game condition: Dots represent the proportions of messages categorized into the respective classes, error bars represent the 95% CI of the proportions. The three panels show proportions—for each number of H-choices—for (A) messages to G characterizing G-actions as moral/altruistic and messages to H characterizing H-actions as selfish/immoral; (B) messages to G coded as strong, messages to G coded as weak, messages to H coded as strong. (C) Mean evaluation of messages to G-players and H-players with 95% CI for the mean.*

The results split by condition show a similar general pattern. Participants viewed action H as immoral slightly more frequently in the injunctive condition and slightly less frequently in the descriptive norms condition. It is again difficult to discern a pattern for strong and weak categories. The difference in average evaluation was most pronounced in the injunctive norms condition and smallest in the descriptive norms condition (consistent with behavior in these conditions). Compared to the other four conditions, the difference in rating resulted from differences in both G-ratings and H-ratings: The evaluation of H was highest in the descriptive norms condition and lowest in the injunctive norms condition and vice versa for the evaluation of G.

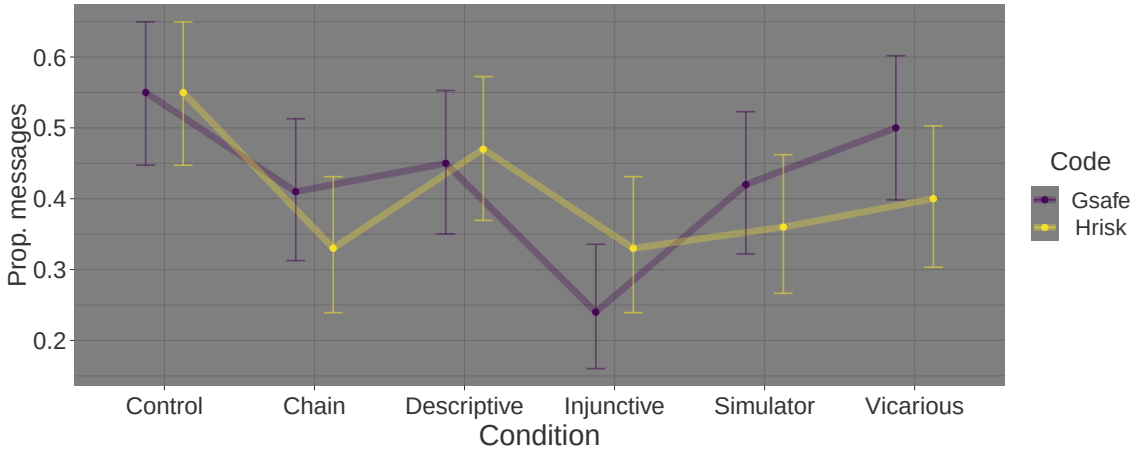


**Figure E81**

*Categorization of messages split by game condition: Dots represent the proportions of messages categorized into the respective classes (messages to H-players characterizing H-choices as risky and messages to G-players characterizing G-actions as safe), error bars represent the 95% CI of the proportions.*

Beyond the categories of might and morality, the most frequent categorization related to relative safety and risk. An example is the message pair "(G) playing it safe, no risk" vs. "(H) aiming for the most money, playing a riskier game". Figure E81 shows the proportions of messages referring to risk and safety. The majority of participants in the control condition made references to risk in their message to H-participants and to the lack of risk or safety in their message to G-participants. This decreased to about a third of participants in the injunctive norms condition. **VM1** and **VM2** contain more examples of answers to the open-format questions.

## E12 Study 2: Selected Simulation Results

We simulated 1,000,000 games for each group of 100 participants, using the 2,500 decisions made by each group. Figure E82 summarizes the results for these simulated games. The pattern of these results roughly corresponds to the comparison of H-choices across conditions, with some notable distinctions. Conditions with the fewest number of H-choices resulted in distinctly lower rates of color switching and subsequently in distinctly higher expected outcomes for group members. Among conditions with high numbers of H-choices,

the curves illustrate the importance of the temporal spacing of H-choices: Results for the vicarious learning condition were much better than for the descriptive norms condition. Participants in the vicarious learning condition seemed to react to the increase in purple participants by switching to G-choices later in the game, which more than compensates their slightly higher rates of H-choices in early rounds. Their curve shows some evidence of flattening and they have a higher expected outcome than participants in the control and descriptive norms condition. These outcomes already take zero-scores for purple participants into account.

To test for the robustness of our results, we repeated the full set of simulations while using a resampling method: For each trial we constructed a new group by randomly sampling participants from the original group with replacement. This resulted in a slight increase in variance, but mean results were close to indistinguishable from the original results (see Fig. E83 for a side-by-side comparison).

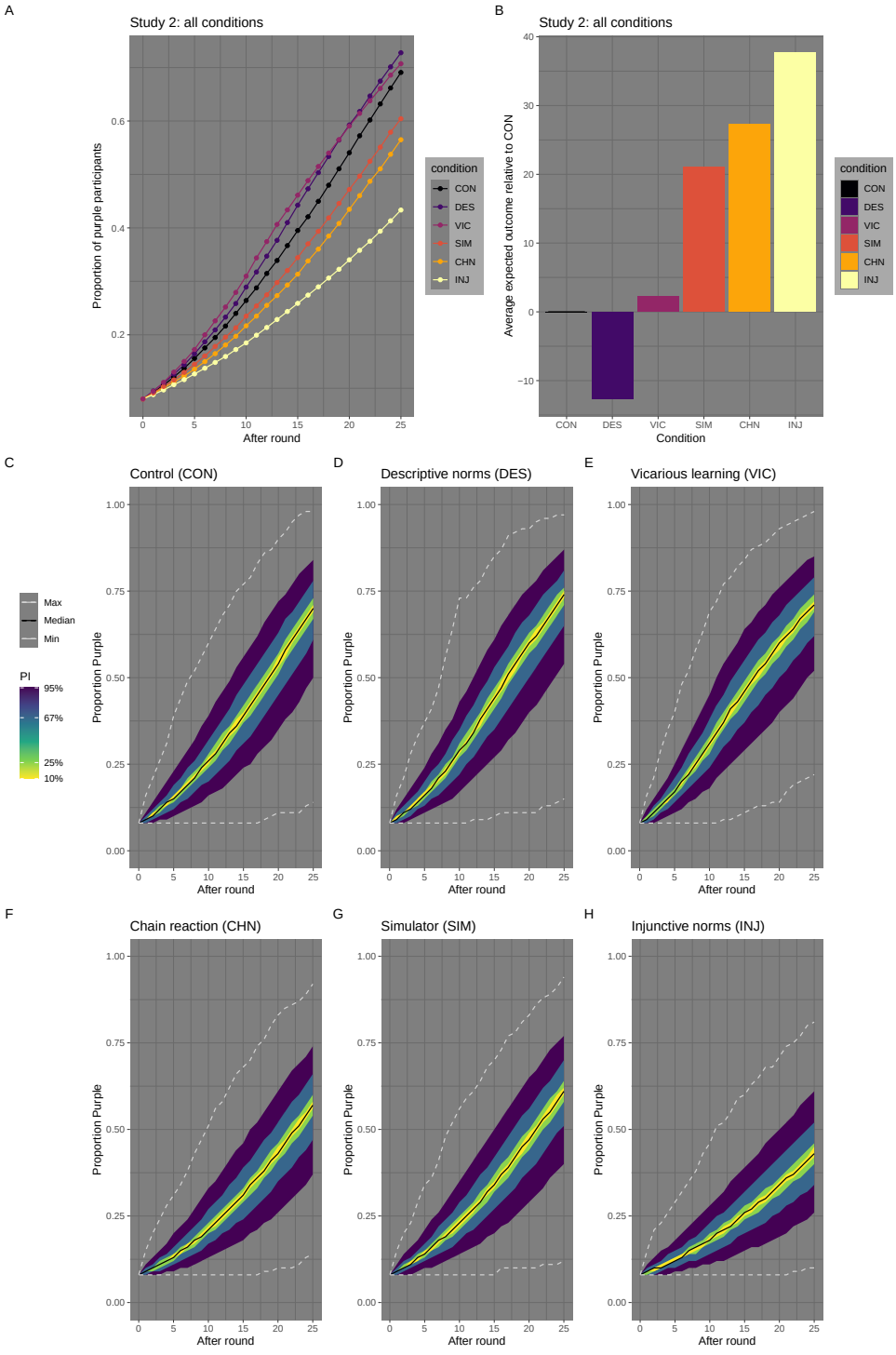


**Figure E82**

*Results for 6,000,000 simulated games based on the observed decisions in six conditions in Study 2: Two panels show results summaries across conditions: (A) Average proportion of purple participants for each round within each condition, and (B) difference in average outcomes from the control condition. Six panels show results within conditions: Graphs show distributions of observed proportions of purple participants by round, indicating the median, minimum, maximum, and the inner 10%, 67% and 95% of results, for the (C) control (D) descriptive norms, (E) vicarious learning, (F) chain reaction, (G) simulator, and (H) injunctive norms condition.*

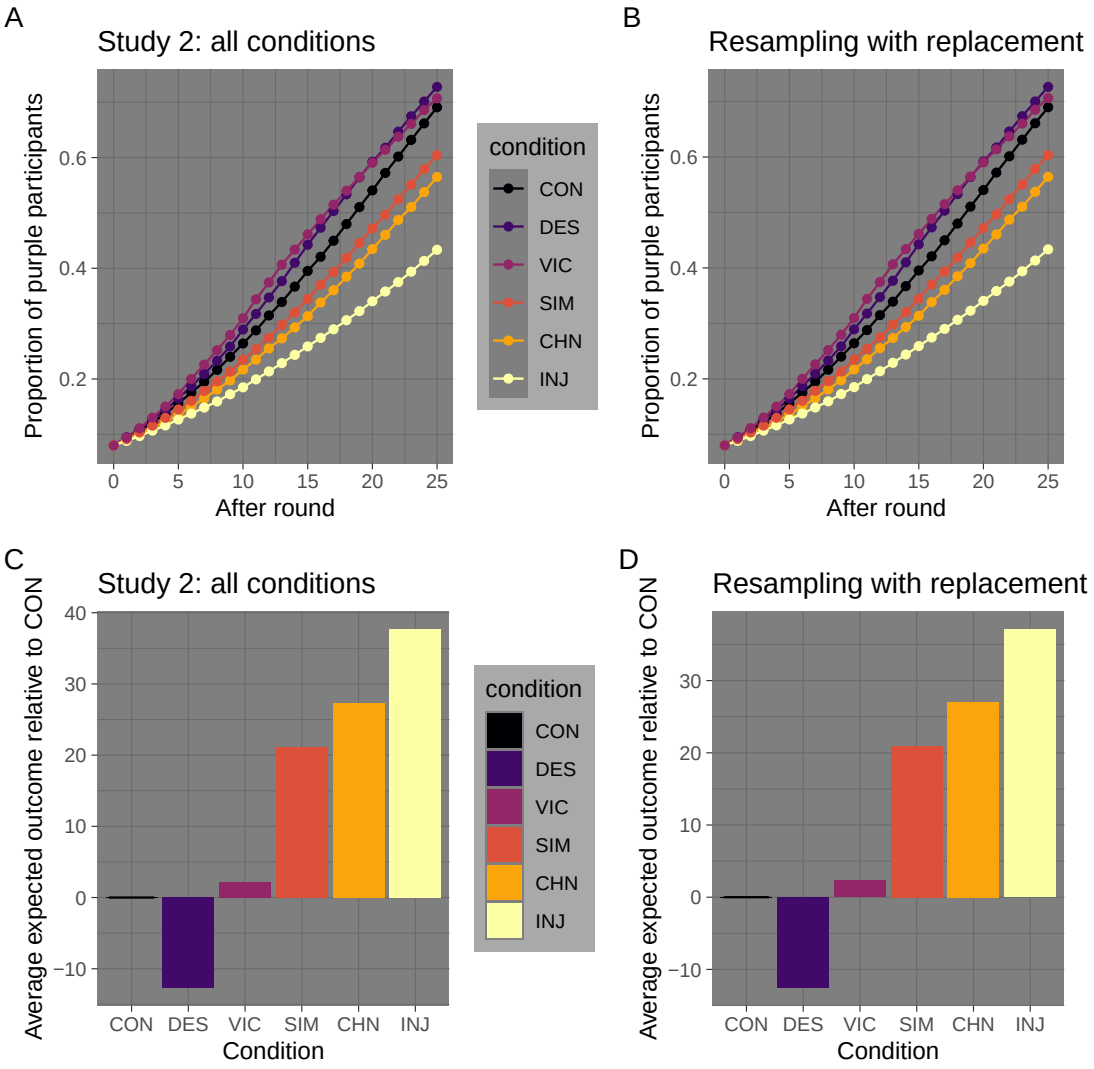


**Figure E83**

*Comparison of original simulation results and simulation results after resampling: Upper panels show round-wise average proportions of purple participants for each of the six conditions (A) for the original, and (B) for the resampled groups. Lower panels show average outcomes for each of the six conditions (C) for the original, and (D) for the resampled groups.*

Finally, Figure E84 demonstrates that the basic social dilemma structure of the game is reflected in outcomes for each of the six conditions. Here, results are shown split by strategy used in each of the six conditions. We still observed higher rates of receiving a bonus for players using the AllG strategy (see **SCR** for additional results), but given the large score differences, AllH players had a higher expected outcome than AllG players in each of the six conditions, with mixed strategies falling in between in most conditions. HG-players on average outperformed Switch-players[16].

Replicating the observations for theoretical populations, the difference in payoffs for competitive and cooperative strategies was higher in conditions with a larger proportion of cooperative choices. Free-riders thus profit from other participants' efforts to reduce the amount of color-switching. It is important to consider these results jointly with the difference in payoff rates. While risk-taking players can expect higher average outcomes, they are faced with a lower probability of receiving a bonus at all. The observed correlation between general risk-taking and competitive choices might indicate that participants were at least to some degree aware of this tradeoff. The observed correlation with social preferences might also indicate that participants were at least to some degree aware of the social dilemma structure of the game.

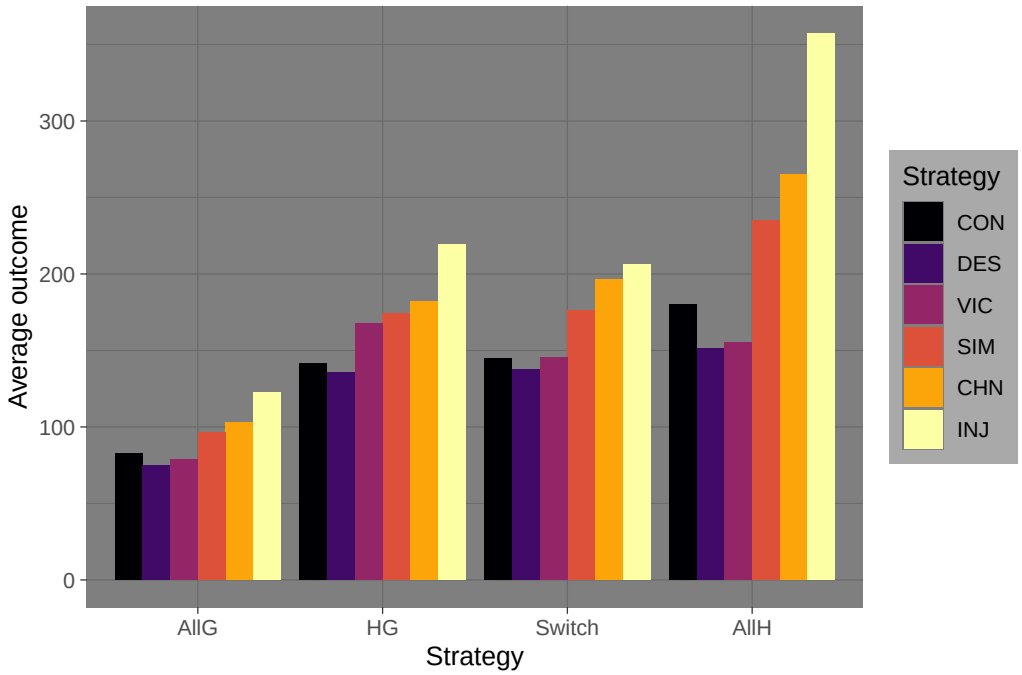


**Figure E84**

*Average outcomes across simulated games for participants categorized as using each of the four possible strategies for each of the six conditions.*

[16]Results are not aligned by the number of H-choices, this explains the variation between conditions.

# Part V
# Discussion and Extensions

## E13 Discussion of Selected Results

### E13.1 Results for predictor variables

Here, we present results for the relation between game behavior and the predictor variables identified in Study 1.

#### *E13.1.1 Social value orientation*

We observed a negative relationship between the social value orientation (SVO) measure and risk-taking in the transmission game. Participants who were classified as altruistic chose safer actions than participants on the individualistic and competitive side of the spectrum. As reported in **SA2**, SVO was positively correlated with worries about the coronavirus, compliance with mitigation measures and negatively with the acceptance of tradeoffs.

**SVO and social distancing.** These findings are consistent with previously reported patterns: Oosterhoff and Palmer (2020) found more hoarding behavior in response to the virus for teenagers who put a larger value on their self-interest. Jordan et al. (2020) observed that a perceived public threat showed a stronger association with prevention intentions than a perceived personal threat and that prosocial behavior in economic games predicted these intentions. Similarly, prosociality predicted social distancing, buying cloth masks, and donating to COVID-19 aid funds Campos-Mercade et al. (2020). Relatedly, Pfattheicher et al. (2020) found that empathy drives social distancing. In contrast, (de Neys et al., 2020) argued that respect for social distancing was driven by a lower assessment of risks, not a difference in care for others.

**SVO and messaging.** One's social value orientation might impact the appeal of received messages, and messages appealing to prosociality might be more effective (see, e.g., Heffner et al., 2020). Blagov (2020) concluded, for example, that appeals to altruism might be successful with the majority of recipients, but create the opposite effect for antagonistic individuals. (Pfattheicher et al., 2020) found that induced empathy for the vulnerable increased the motivation to socially distance. Some of these effects may be fragile (see, e.g., Utych & Fowler, 2020): Jordan et al. (2020) saw the effectiveness of prosocial appeals decrease over the course of the pandemic, while Cao et al. (2020) explained a reduction in prosocial willingness over time with compassion fatigue. Leder et al. (2020) even found, in contrast, that protective measures were mainly used for self-protection and not the protection of others, even in a predominantly prosocial sample of participants. Further, SVO did not predict compliance.

#### *E13.1.2 Risk-taking*

General risk preference predicted risky decision making in the transmission game, and at the same time non-compliance with social distancing measures and the willingness to accept tradeoffs between human lives and continuity of public life. It should be noted that objective risk needs to be perceived before it can exert an influence on behavior.

Abdelrahman (2020) and Plohl and Musil (2021) found risk perception to predict social distancing and compliance. However, risk perception might not always be grounded in objective risk, as Jarynowski et al. (2020) observed in a Polish sample. In contrast, (Du et al., 2020) found prevalence rates of COVID-19 to increase web searches for protective behaviors. The effect of perceived risk might also vary between domains (Galandra et al., 2020). For example, Bostrom et al. (2020) suggested that pandemic risks are perceived as more natural than risks posed by climate change, and thus might evoke less moral concerns.

#### *E13.1.3 Cognitive reflection*

We found a negative relationship between CRT scores and risk-taking (r=-.16), but the analysis of behavioral pattern showed that the group moving from risky to safe strategies (H–G) was the only group with higher average CRT scores. CRT scores further showed a weak positive correlation with compliance (r=.09) and a weak negative correlation with worries (r=-.09). The CRT is a popular measure and has been included in many COVID-related current studies. For example, Teovanovic et al. (2020) observed that CRT scores predicted a greater adherence to guidelines. Pennycook, McPhetres, Zhang, et al. (2020) found a better performance in judging the accuracy of information in participants with higher CRT scores, Stanley et al. (2020) a positive correlation between CRT and social distancing (r=.14), and a negative correlation with believing the pandemic to be a hoax (r=-.46; see Alper et al., 2020; Cavojova et al., 2020; Erceg et al., 2020, and the section on conspirational thinking below).

#### *E13.1.4 Personality*

In Study 2, we were not able to replicate the pattern of results in Study 1 for the HEXACO-scales, namely that honesty/humility and emotionality negatively predicted risk-taking. Inspecting correlation patterns with COVID-related scales, we found that honesty/humility showed a small negative correlation with accepting tradeoffs (r=-.15), while emotionality predicted worries (r=.28), compliance (r=.15), and tradeoffs (r=-.15). In addition, openness followed the same pattern as emotionality, but to a lower degree. In a large study (total N $>$ 10k), (Zettler et al., 2020) reported large variations between subsamples regarding correlations between personality scales and adherence with health recommendations, with within-study replication failures for almost all found effects (and the general size of effects relatively small; supported by Martin, 2020). In their meta-analysis, they reported a positive correlation between behavioral adjustment and honesty/humility, and a positive correlation between perceived risk and emotionality. In contrast to our findings in Study 2, they also reported that conscientiousness predicted adjustment. Muto et al. (2020) reported a negative correlation between extraversion and adherence that we likewise did not observe. Folk et al. (2020) found no evidence for the hypothesis that extroverts fared worse than introverts during the crisis (and neither did we). Among the predictors for social distancing in Abdelrahman (2020) were conscientiousness and neuroticism (related to emotionality in the HEXACO scales). Stockpiling was predicted by lower values of honesty/humility (Columbus, 2020)

Some studies investigated links between general susceptibility to anxiety, as measured in emotionality and neuroticism scales, and specific fears about the coronavirus (see

also Hongbo et al., 2020; McElroy et al., 2020). Consistent with the definition of emotional stability, Flesia et al. (2020) found a negative association with distress levels, and Kroencke et al. (2020) found more negative affect in participants with higher neuroticism. Zettler et al. (2020) linked emotionality to willingness to accept restrictions. Other studies supported a link between worry about COVID-19 and more responsible behavior (Erceg et al., 2020), similar to the moderate correlation between worries and compliance in our study (r=.33). Further scales found to predict behavior in response to the coronavirus were the dark factor of personality (reducing the acceptance of restrictions or health behavior Blagov, 2020; Zettler et al., 2020), openness for experience (Götz et al., 2020; Martin, 2020), and agreeableness (Blagov, 2020).

Given this general pattern of effect sizes and variation between studies, large samples and meta-analyses across national contexts might be the best way to address the question of how personality traits affect reactions to the novel virus. Some connections seem uncontroversial (emotionality, humility) for some target variables, but the lack of consistently strong effects should be considered as a warning against assuming simple, unconditional relationships.

## E13.2 Postquestionnaire results

### *E13.2.1 Linear transmission curves*

Consistent with the findings reported by Lammers et al. (2020), participants' average estimates of the number of purple participants across rounds were close to linear. In the vicarious learning condition, there was a slight curvature corresponding to the previous group's data presented to participants. At the same time, participants assumed in all conditions that there was a reduction of risky behavior over time. However, there could be an alternative interpretation of the estimated linear relationship: Participants might have believed that mitigation measures were sufficient to keep the growth at a constant rate. This, of course, is speculation given the lack of pertinent measures. Another interesting finding was the relative proximity of estimated end points across all conditions. This could be plausibly interpreted as limits in understanding the precise dynamics unfolding in each condition based on the assumed behavior. Instead, participants might use the risk-reward heuristic (Leuker et al., 2018) to generate plausible final values. Based on the maximum bonus of 5 GBP that can be calculated based on the number of rounds and the communicated point conversion rate, participants might assume that less than half of the participants will receive a bonus in this game.

### *E13.2.2 Optimism*

As seen in Figure E73, participants tended to estimate their chances of receiving a bonus as higher than those of other participants. This relative optimism had a correspondence in participants' estimates of how likely they were to become infected during the pandemic. This type of optimism had been observed for the H1N1 flu virus (Cho et al., 2013; Rudisill, 2013). Raude (2020) found in a large survey in European countries that coronavirus risk estimates were optimistically skewed in the early stage of the epidemic. Similar optimism was witnessed by Gerhold (2020) and Kuper-Smith et al. (2020). Further, the perceived risk of infection was linked to the willingness to practice social distancing and

other counter-measures (de Neys et al., 2020; Wise et al., 2020). At the same time, relative optimism might not reflect absolute optimism: for example, Glöckner et al. (2020) found in a German sample that participants exhibited underconfidence regarding their chances to remain unimpacted by the virus. Globig et al. (2020) saw private optimism and public pessimism, with protective behaviors therefore mainly driven by concern for others.

#### *E13.2.3 Survey responses vs. behavior*

Surveying the postquestionnaire answers regarding participants' reactions to the five interventions, it is striking to see that the stark differences in behavior across conditions were not necessarily reflected in similar differences regarding their stated evaluation of the interventions (see also **SA2**). Here, we would argue that the observed behavior is the more relevant target variable. This might also explain the pattern of findings in Pink et al. (2020), who conducted five experiments with persuasive short messages in the US over a period of five months. The authors selected the most promising messages based on scale ratings by online participants, and measured behavioral compliance intentions in the remainder of the studies. They did not find consistent effects of any message across all studies relative to their control condition. One should mention first, that the control condition already had moral implications for action. The control message explained the way the virus is transmitted[17] and told participants that by staying home and avoiding contact they could help to prevent the spread of the disease. This statement alone was not vastly different from our injunctive message. One way to interpret the lack of consistent benefits of other messages in Pink et al. (2020) is as a robustness check for the strength of our manipulation. Along these lines, the authors also reported a negative impact of a descriptive norms intervention.

A second way to interpret the findings of Pink et al. (2020) is to point out that all dependent variables in their study were ratings on scales indicating intentions, not incentivized behavior. An interesting way to approach these questions would be to translate all messages that do not use deception into the language of our paradigm and test their effectiveness in this setting. Pink et al. (2020) further suggested that the effectiveness of messages might change over time in a pandemic, for example, because a lot of information is shared about the virus over time or the perceived nature of the pandemic calls for different approaches at different times. The study also showed a potential ethical problem associated with testing a diverse set of messages targeting behavior in a pandemic: Negative effects on intentions to adhere to guidelines and mitigation behavior might put participants at risk.

### E13.3 Attitudes towards COVID-19

#### *E13.3.1 Political ideology*

As reported above, we found that political ideology, party affiliation, and preference for presidential candidates were strong predictors of self-reported compliance with anti-measures against COVID-19, but also of worries about the virus and the willingness to accept tradeoffs. Conservatives reported lower compliance, less worries and were more willing to accept tradeoffs. The tendency to reject counter-measures and accept tradeoffs was even more strongly predicted by a stated preference for Donald Trump to be re-elected.

[17] They call the virus a "respiratory illness" though, which is technically incorrect

We found relatively smaller correlations between game results and conservatism (r=.07 and r=.13, for the two measures) and the preference for Trump and game results (r=.11).

At first glance, this pattern does not seem to be self-evident, especially since risks posed by viruses tend to be viewed as natural risks (Bostrom et al., 2020). Yet, there are two components that might help to understand this relationship. In many countries, and especially in the United States, conservative positions reject regulations and perceived restrictions of freedom, while valorizing individual choice. Kahn et al. (2020) argued that conservatives are more concerned than liberals about commission threats (such as imposing restriction measures), but not about omission threats (such as a failure to protect a vulnerable population). Second, political messaging differed substantially along partisan divisions, also across individual states (Rosenfeld, 2020) and counties (Kavanagh et al., 2020) in the US. Thus, political predictors featured in many surveys studying the response to COVID-19.

The inverse relationship between conservatism and adherence to guidelines replicates previous findings (Briki & Dagot, 2020; Raihani & de-Wit, 2020; Rothgerber et al., 2020). The inverse relationship between conservatism and concern was likewise observed in multiple studies (Briki & Dagot, 2020; Conway et al., 2020; Pennycook, McPhetres, Bago, et al., 2020a; Raihani & de-Wit, 2020; Rothgerber et al., 2020). Conway et al. (2020) linked these relationships by suggesting that ideological differences motivate differential perceptions of the crisis by liberals and conservatives, leading to less concern and less compliance. Plohl and Musil (2021) argued for differences in trust in science as playing this mediating role. Differential consumption of conservative media, especially in the US, might amplify existing differences (Pennycook, McPhetres, Bago, et al., 2020a; Rothgerber et al., 2020).

Based on the patterns of international findings, the close relationship between conservatism and corona-related attitudes might not be universal and vary substantially between countries. Harper et al. (2020) found no effect of politically relevant variables for predicting mitigating behavior, Pennycook, McPhetres, Bago, et al. (2020a) found a weaker relationship in the UK than in the US. On the other hand, Farias and Pilati (2020) found similar relationships for a sample in Brazil, where reactions to the coronavirus are similarly politicized as in the US.

#### *E13.3.2 Conspiratorial thinking*

One of the strongest predictors of attitudes towards social distancing in Study 1 ($n_1$ = 100; misinformation was not measured in Study 2) was the endorsement of misinformation (r=-.70). Misinformation endorsement also predicted acceptance of tradeoffs (r=.49) and compliance with counter-measures (r=-.27). The endorsement of misinformation was heterogeneous across the political spectrum and strongly correlated with conservatism (r=.48), preference for Donald Trump (r=.58) and negatively with preference for Joe Biden (r=-.43). We also observed a very high correlation (r=.63) between misinformation about the coronavirus and climate skepticism (all results shown in **SA1**). We did not find a substantial relationship between misinformation and our cognitive measures (numeracy, CRT). Note that conspirational thinking was also produced in open-format comments, such as "Its a weaponized hoax meant to take down our current president and reverse all his gains. China is responsible due to their need to preserve their position globally." See **VM1** and **VM2** for further examples.

Our sample size in Study 1 was relatively small, thus the lack of a negative correlation between misinformation and CRT scores (r=.01) should not be taken as strong evidence for the null hypothesis. In previous studies, small to moderate negative correlations were observed between CRT scores (with some variation in actual measures) and belief in COVID-19 conspiracies (Alper et al., 2020; Cavojova et al., 2020). National narcissism predicted health-related behaviors in both the UK and the US in Sternisko et al. (2020) with more variance explained in the US (compare also Alper et al., 2020). Velásquez et al. (2020) described the organization of newly forming clusters in the online community around COVID-19 conspiracies spanning multiple platforms in a decentralized pattern (Velásquez et al., 2020).

Marinthe et al. (2020) and Imhoff and Lamberty (2020) see conspirational thinking as linked to a general "conspiracy mentality". High values on this dimension may both increase the perception of risk and reduce the willingness to engage in standard preventive behaviors. Thus, even when the risks involved in climate change and viruses might be categorized as different in nature (Bostrom et al., 2020), conspirational thinking might lead to affirmation of both. Brzezinski et al. (2020) discussed that the concentration of climate change skeptics predicted lags in adopting lockdown policies in counties. Bertin et al. (2020) found that conspiracy mentality also predicted the lack of willingness to be vaccinated without hesitation, which might well become a singular focus in 2021 (Fisher et al., 2020; Korn et al., 2020). The correlation between conservatism and unscientific beliefs has been documented often, e.g., in van der Linden et al. (2018) and Lewandowsky et al. (2020).

## E13.4 Intervention results

### *E13.4.1 Vicarious learning*

Participants in the vicarious learning condition chose relatively more risky actions than participants in the control condition, but those risky actions were mainly concentrated in earlier rounds. While some participants might have chosen this pattern for strategic reasons, it is also consistent with an underestimation of the growth rate based on a linear projection into the future (Fansher et al., 2020; Lammers et al., 2020; Trueblood et al., 2020). Only when participants actually witnessed the higher than estimated rise in numbers did their behavior change.

Moreover, many studies report that risk perception during the pandemic did not necessarily track changes in objective risk (Jarynowski et al., 2020; Wise et al., 2020). Our intervention confronted participants directly with objective information and there was no reason to assume that participants did not trust the veracity of this information. We stressed at the start of the study that the studies published under the first author's Prolific account do not deploy deception in any study and that information can be relied on to be truthful. In this sense, the situation in the game might have been more favorable for the intervention to work than a real-life situation that are more fraught with problems of accessing relevant and trustworthy data.

Information presented in graphs might be more difficult to accurately understand than information presented in tables (Fansher et al., 2020). To alleviate problems in graph comprehension, we always included a statement describing the round state in addition to

the graph. We also presented data on a linear and not a logarithmic scale that has been shown to cause further interpretation problems (Rentsch et al., 2020).

Boin et al. (2020) argued that governments might be ill-prepared for a crisis that does not materialize in large, and immediate events, but begins (and ends) slowly. They wonder whether current structures of crisis management are up to the task of mitigating a pandemic. Our results suggest that public risk perception might not help in addressing these deficits. Vicarious learning might have a stronger impact, if the observed behavior allows for projections into one's personal future. This happened at the beginning of the pandemic, when other nations observed China and Italy dealing with their first wave. Notably though, this benefit can only be reaped if people's vicarious experience is translated into accurate risk projections. The simulator condition might be an optimistic test case for this potential.

#### *E13.4.2 Descriptive norms intervention*

The descriptive norms condition exhibited the most risk-taking and the worst average outcomes across all conditions. There are positive examples of normative feedback interventions resulting in a reduction of the target behavior (Burger & Shelton, 2011; Halpern, 2015; Lewis & Neighbors, 2006a). Gouin (2020) reported that a perception that most others follow directives made it more likely that participants followed these directives themselves.

At the same time, interventions using descriptive norms have an uneven track record with many examples of ineffective interventions (Huh et al., 2015; Lees et al., 2020). Moreover, normative interventions have been observed to backfire and move behavioral tendencies in the undesired direction (Cialdini, 2003; Richter et al., 2018; Schultz et al., 2007; Whitt et al., 2019). Reasons why interventions backfired in the past included that the targeted behavior was not shown by an overwhelming majority of others or that the source of information was not trusted (Bicchieri & Dimant, 2019; Cialdini, 2003; Schultz et al., 2007, 2018). In our case, though, the normative information was given directly by us. A majority of participants chose safe actions from the second round forward, but not a large majority. Thus, this might have given license for participants to show risky behavior, as it may have appeared to them as widespread. Potentially unwilling to forego befits reaped by others (Woike & Hafenbrädl, 2020) this could have led to an increase in risky choices.

Fortunately, we implemented this intervention within an abstract infection simulation. The increase in risky behavior thus has no impact on responses to the pandemic. Some attempts at testing descriptive norms interventions appeared to be less fortunate. Leary and Dvorak (2020) reported that their normative intervention led to less social distancing by participants subsequently, Kitamura and Yamada (2020) observed less adherence to social distancing measures in one of their normative feedback conditions. This poses an ethical dilemma for studying interventions in the wild without estimating the directionality of their effect, first.

Recent and not so recent studies show some evidence that the weight of social information is higher when referring to sources close to or in the in-group of the receiver (Schulze & Pachur, 2019; Tunçgenç et al., 2020; Van Bavel et al., 2020). For example, one participant wrote in an open-format comment: "My father passed away and if social distance was enacted sooner it could be prevented". A manipulation of social distance is difficult in our anonymous paradigm, and may yet backfire when facing clusters of divergent behavior, as political polarization might create.

### *E13.4.3 Injunctive norms intervention*

Our injunctive norms intervention–the simplest intervention in the set—turned out to be the most effective intervention by far. There are multiple, non-exclusive candidate suggestions found in the literature to explain this success.

**Appropriateness.** March and Olsen (2013) argued that behavior followed a "logic of appropriateness". The game situation is novel and players may therefore face ambiguity as to which action is appropriate. The injunctive norms intervention resolves this ambiguity by relating the actions to general principles (Nozick, 1993) that many participants have already accepted for a range of situations. These principles include doing no harm, protecting others, not taking unnecessary risks, etc. Cialdini and Goldstein (2004) likewise focused on appropriate responses to situations informed by norms and stressed the importance of highlighting normative information to make it effective.

**Effective message components.** Grant and Hofmann (2011) argued that a focus on consequences for others should be more effective than a focus on personal consequences for increasing beneficial health behavior (see also Globig et al., 2020; Pink et al., 2020). Luttrell et al. (2016) argued for the importance of a perceived moral basis for one's attitudes for producing corresponding behavior. If a moral basis depends on contested scientific results, the perceived scientific consensus can play a central role in disambiguating the situation (Lewandowsky, Gignac, & Vaughan, 2013). In a similar vein, (Yoeli & Rand, 2020) stress the importance of a message being short and unambiguous, combined with the implication that others expect the receiver to comply. Sources of ambiguity that might render messages ineffective are hedging expressions that allow for interpretation, such as "wear mask whenever possible" or "as much as practicable" (Michie et al., 2020). At the same time, the communicator of messages needs to be perceived as honest and credible for the message to have an effect (Bostrom et al., 2020).

**Response to COVID-19.** Seale et al. (2020) found that "being the right thing to do" was the strongest motivation for adhering to social distancing guidelines, and Kuiper et al. (2020) observed that intrinsic moral motivation was a predictor of compliance. Some studies explored the effectiveness of giving potential reasons for these measures to be the "right thing to do". Similar to Grant and Hofmann (2011), Pfattheicher et al. (2020) identified empathy for others as an essential motivation for distancing, and Moore et al. (2020) argued that the link to population-level outcomes might be an essential argument for convincing younger people to reduce individual risk-taking. Everett et al. (2020) found a deontological intervention stressing one's duty to be more effective than a focus on consequences. Consistent with these findings, Gouin (2020) found that a belief in duty, protection for others, and perceived consensus drove adherence to social distancing; Capraro and Barcelo (2020) stressed the importance of the level of one's community. The effectiveness of these elements might undergo change over time during the pandemic: Jordan et al. (2020) found a prosocial framing to be effective in the beginning of the pandemic in the US in March, but not a month later and Favero and Pedersen (2020) did not observe substantial variation between framings at the beginning of April.

#### *E13.4.4 Boosting interventions*

Both boosting interventions showed some reduction in risky actions, with the simulator intervention performing worse than the chain reaction intervention. On the positive side, both interventions seem to have succeeded in establishing a better understanding of the risk involved. Both interventions underline that transmissions do not progress in a linear fashion, as assumed by many observers (Lammers et al., 2020). If violators were primarily driven by a different risk assessment of the virus (de Neys et al., 2020), a simulator intervention would be able to put this assessment into question.

On the negative side, both interventions demand more attention than the other interventions we tested. The chain reaction story is less technical than the simulator condition, and it also offers a simple solution to the described problem of transmission chains: safe behavior may prevent the chains from happening. This is close to the effective means of coping with threats —an offer that Wood (2000) argued may increase the likelihood of following a recommendation. Finding this solution in the simulator condition requires a comparison of results between simulation conditions and translating gained insights into action.

The success of the chain reaction condition and the mixed results for the simulator condition might also be an indicator that interventions might need to be tailored to recipients. Some of the feedback in the simulator condition was very positive, whereas other participants did not appreciate this exercise at all.

## E14 Variants, extensions, and research ideas

There are many possible variations that could help to address specific aspects of pandemic behavior or behavioral interventions. In this section, we explore a number of these variants.

First, we present a paper-and-pencil presentation of the game and the possibility to simulate the game even without electricity. Second, we discuss possible variations of parameters to show how specific changes can be compensated by others to create feasible scenarios for a broad parameter range. Third, we give an example of a more substantial change to the game format. In this example, we introduce infectious periods and immunity and discuss possible applications.

This does not exhaust the range of possible variants. Using either the strategy method, small groups, or mixed groups of anonymous players and identified computer players, many more variants are feasible that remain deception-free, properly incentivized and flexible. One could incorporate mechanisms for messaging, punishment, point transfers, differential scoring rules, etc. Features that were cut to decrease complexity, could be reintroduced into the game as a focus for experiments. While we would be careful not to overburden participants with rule complexity, further limits are imposed only by the researcher's creativity.

### E14.1 Paper and pencil version of the Transmission Game

The transmission game does not necessarily require access to computers and can be played, simulated and paid out in paper-and-pencil format. We believe that this option may be attractive when conducting studies in the field. In contrast, a simultaneous variant of the game would require simultaneous access to many computers, which would severely limits its appropriateness for many many settings. Here, we demonstrate how experimenters can prepare the game material and analyze the results.

#### *E14.1.1 Game sheets*

A paper-and-pencil version of the game requires instruction material in paper format. Experimenters should allow for control questions to be checked before the game and offer corrections and answers to questions before the game.

Here, we focus on game sheets that could be used to carry out a full, incentivized experiment. The proposed game sheets are shown in triple version in Figure E85.

The game slips have 25 pairs of check boxes to mark decisions for each round. They have three more elements. The element in the lower left can be used to note down a participant code that allows to identify participants in a separate list. The other two elements are used for playing the game and determining outcomes.

Figure E86 shows two filled-in sheets handed in by hypothetical participants. If 100 sheets are collected the game is ready to be played out.

#### *E14.1.2 Determining results*

Playing out the game requires the sheets, a pen and a twenty-sided die (D20, see Fig. E87). To keep track of game rules and the number of rounds completed in the game, the material shown in Figure E88 can be useful.

**Color game**

Round **01**: Ⓖ Ⓗ Round **02**: Ⓖ Ⓗ Round **03**: Ⓖ Ⓗ Round **04**: Ⓖ Ⓗ
Round **05**: Ⓖ Ⓗ Round **06**: Ⓖ Ⓗ Round **07**: Ⓖ Ⓗ Round **08**: Ⓖ Ⓗ
Round **09**: Ⓖ Ⓗ Round **10**: Ⓖ Ⓗ Round **11**: Ⓖ Ⓗ Round **12**: Ⓖ Ⓗ
Round **13**: Ⓖ Ⓗ Round **14**: Ⓖ Ⓗ Round **15**: Ⓖ Ⓗ Round **16**: Ⓖ Ⓗ
Round **17**: Ⓖ Ⓗ Round **18**: Ⓖ Ⓗ Round **19**: Ⓖ Ⓗ Round **20**: Ⓖ Ⓗ
Round **21**: Ⓖ Ⓗ Round **22**: Ⓖ Ⓗ Round **23**: Ⓖ Ⓗ Round **24**: Ⓖ Ⓗ
Round **25**: Ⓖ Ⓗ

**Color game**

Round **01**: Ⓖ Ⓗ Round **02**: Ⓖ Ⓗ Round **03**: Ⓖ Ⓗ Round **04**: Ⓖ Ⓗ
Round **05**: Ⓖ Ⓗ Round **06**: Ⓖ Ⓗ Round **07**: Ⓖ Ⓗ Round **08**: Ⓖ Ⓗ
Round **09**: Ⓖ Ⓗ Round **10**: Ⓖ Ⓗ Round **11**: Ⓖ Ⓗ Round **12**: Ⓖ Ⓗ
Round **13**: Ⓖ Ⓗ Round **14**: Ⓖ Ⓗ Round **15**: Ⓖ Ⓗ Round **16**: Ⓖ Ⓗ
Round **17**: Ⓖ Ⓗ Round **18**: Ⓖ Ⓗ Round **19**: Ⓖ Ⓗ Round **20**: Ⓖ Ⓗ
Round **21**: Ⓖ Ⓗ Round **22**: Ⓖ Ⓗ Round **23**: Ⓖ Ⓗ Round **24**: Ⓖ Ⓗ
Round **25**: Ⓖ Ⓗ

**Color game**

Round **01**: Ⓖ Ⓗ Round **02**: Ⓖ Ⓗ Round **03**: Ⓖ Ⓗ Round **04**: Ⓖ Ⓗ
Round **05**: Ⓖ Ⓗ Round **06**: Ⓖ Ⓗ Round **07**: Ⓖ Ⓗ Round **08**: Ⓖ Ⓗ
Round **09**: Ⓖ Ⓗ Round **10**: Ⓖ Ⓗ Round **11**: Ⓖ Ⓗ Round **12**: Ⓖ Ⓗ
Round **13**: Ⓖ Ⓗ Round **14**: Ⓖ Ⓗ Round **15**: Ⓖ Ⓗ Round **16**: Ⓖ Ⓗ
Round **17**: Ⓖ Ⓗ Round **18**: Ⓖ Ⓗ Round **19**: Ⓖ Ⓗ Round **20**: Ⓖ Ⓗ
Round **21**: Ⓖ Ⓗ Round **22**: Ⓖ Ⓗ Round **23**: Ⓖ Ⓗ Round **24**: Ⓖ Ⓗ
Round **25**: Ⓖ Ⓗ

**Figure E85**
*Empty game slips for playing the paper-and-pencil version.*

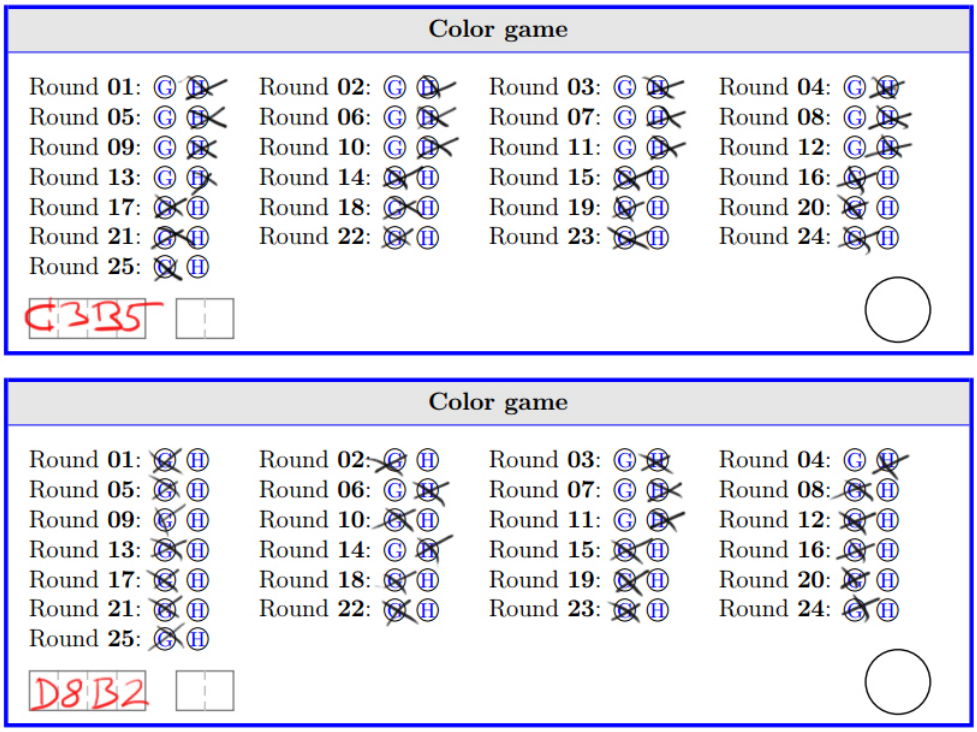


**Figure E86**
*Filled in game slips (paper-and-pencil version).*

Before the first round, the experimenter draws eight game sheets out of the stack and marks them as purple. The circle in the lower right on each sheet can be filled in to clearly demarcate purple players.

**Figure E87**
*Twenty-sided die (D20) for the paper-and-pencil version of the game.*
*Photo Credit: Lucas Santos, Unsplash.*

For playing out the game, the experimenter needs to shuffle the slips of paper at the beginning of each round. The experimenter takes sheets from the top of the stack in pairs

**Color Game: Round Tracker**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 21 | 22 | 23 | 24 | 25 | | | | | |

**Color Game: Dice rules (when blue is matched with purple)**

| Player 1 | Player 2 | Color switch on die result |
|---|---|---|
| H | H | 1–5 |
| G | H | 1–3 |
| H | G | 1–3 |
| G | G | 1 |

**Figure E88**
*Game round tracker and die results table for the paper-and-pencil version of the Transmission game.*

and checks for each pair, whether a blue and a purple player are matched (it is sufficient to check that exactly one of the two circles is filled in). For these mixed pairs, the current rounds actions are checked and a D20 is rolled to determine–according to the die result table—whether a color change takes place and the circle on the sheet is filled in. If a color change occurs, the current round on the sheet of the affected player can be circled to keep track of the development of infection rates.

The game progresses in this way over 25 rounds. At the end, all slips have to be checked one more time. Those slips with a filled in circle do not earn any bonus money and the remaining field at the bottom of the slip is crossed out. Those slips whose cricle is not filled out receive a check mark in the circle, and the number of H-choices is written into the remaining field (see Fig. E89 for one example of each case).

As a final step, the number of H-choices can be linked to the payoff distribution (see Figure E90) and participants may receive their bonus payment by linking payoff amounts, participant codes and names of participants.

### *E14.1.3 Possible applications of the paper-and-pencil variant*

An obvious benefit of the paper-and-pencil variant is the universal applicability in any field setting. In this form, the game can be administered to large groups of people without the requirement of electronic devices for anyone in the room. Even in industrialized countries, universal access to electronic input devices might not be safely assumed in

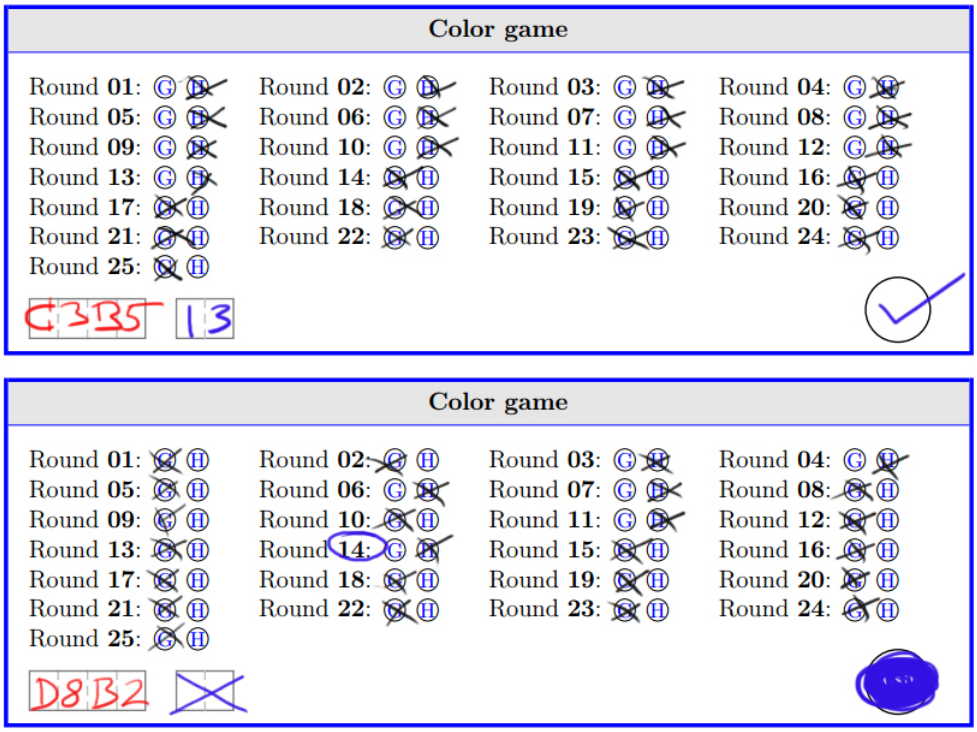


**Figure E89**

*Final state of game slips: The top slip belongs to a player who has not changed color during the game. The player made 13 H-choices and thus, 13 was entered into the result field. The bottom slip belongs to a player who switched color during round 14. The 14 is therefore circled and the result field is crossed out.*

all settings, but this concern is even more justified in more general experimental settings (Henrich et al., 2010a, 2010b).

A second benefit of using this variant would be its educational use. The paper-and-pencil version could be played and evaluated by a group of students. For example, while playing out the game students could determine roundwise population states and plot a graph of infection proportions across rounds, illustrating the concept of non-linear growth with a concrete example. The game might need to be scaled down to class size for this purpose (see some of the variants below).

Obviously, the paper-and-pencil variant could easily be combined with computerized analysis by transferring the entries as input for a simulation.

### E14.2 Varying parameters in the simulation

**Color Game: Payoff calculation**

| # of H-choices | Points | Payoff |
|---|---|---|
| 0 | 200 | £1.00 |
| 1 | 232 | £1.16 |
| 2 | 264 | £1.32 |
| 3 | 296 | £1.48 |
| 4 | 328 | £1.64 |
| 5 | 360 | £1.80 |
| 6 | 392 | £1.96 |
| 7 | 424 | £2.12 |
| 8 | 456 | £2.28 |
| 9 | 488 | £2.44 |
| 10 | 520 | £2.60 |
| 11 | 552 | £2.76 |
| 12 | 584 | £2.92 |
| 13 | 616 | £3.08 |
| 14 | 648 | £3.24 |
| 15 | 680 | £3.40 |
| 16 | 712 | £3.56 |
| 17 | 744 | £3.72 |
| 18 | 776 | £3.88 |
| 19 | 808 | £4.04 |
| 20 | 840 | £4.20 |
| 21 | 872 | £4.36 |
| 22 | 904 | £4.52 |
| 23 | 936 | £4.68 |
| 24 | 968 | £4.84 |
| 25 | 1,000 | £5.00 |

**Figure E90**
*Payoff calculation sheet: Table to translate the number of H-choices into monetary payoffs.*

We chose parameter values in the Transmission Game to capture several elements of preventive measured in a pandemic: The measures reduce risk, but not to zero. The chance of transmission even without preventive measures does not have a high probability. Measures protect the person choosing them, but also people that person comes into contact with. Finally, near-universal use of measures makes a real difference in expected transmission rates. We also created the scenario to capture non-linear case growth observed in the early days of the pandemic, as discussed in our presentation of the framework above.

One may ask whether the parameterization we have chosen constitutes a rare, atypical case. This might be linked to a concern that highly specific assumptions might limit the applicability of the paradigm when there is no room for modifications of its elements. Here, we show some examples of different parameter sets and variants that create a game structure fulfilling the requirements.

To demonstrate the consequences of changes, we simulate games with three different populations: The first population has 80% H-players (choosing H-actions in every round of the game) and 20% G-players (choosing G-actions in each round). The second population is equally split with 50% H-players, and the third population has 20% H-players and 80% G-players. Each game was simulated 1,000,000 times and we report the averages for these three populations to give an idea about a realistic range of possible outcomes.

### *E14.2.1 Reducing the number of players*

We discussed the question of desired scalability and it should be clear from the setup that increasing the number of players poses no problem for the paradigm. Here, we show one possibility for reducing the number of players while keeping the parameters roughly the same.

Figure E91 presents the results for two variants based on a population of only 24 players (the number of H-players are rounded in this case). The number of initial cases is changed to 2 (which corresponds to $8.\bar{3}\%$).

Keeping all parameters constant, the game results vary across a similar range as for the 100-player game (more extreme results would be seen for more extreme populations, but we present a reasonable range of possible populations). This result for variant 1 reinforces the idea that the game is scalable to most (even) player counts. We chose 24 players, as this compares roughly to the maximum occupancy of many experimental laboratories and the number would allow for combining the game with two-player, three-player, four-player, six-player, eight-player and twelve-player group games[18].

[18]This would not allow for a perfect-stranger-matching with a single group in the transmission game (Both et al., 2016).

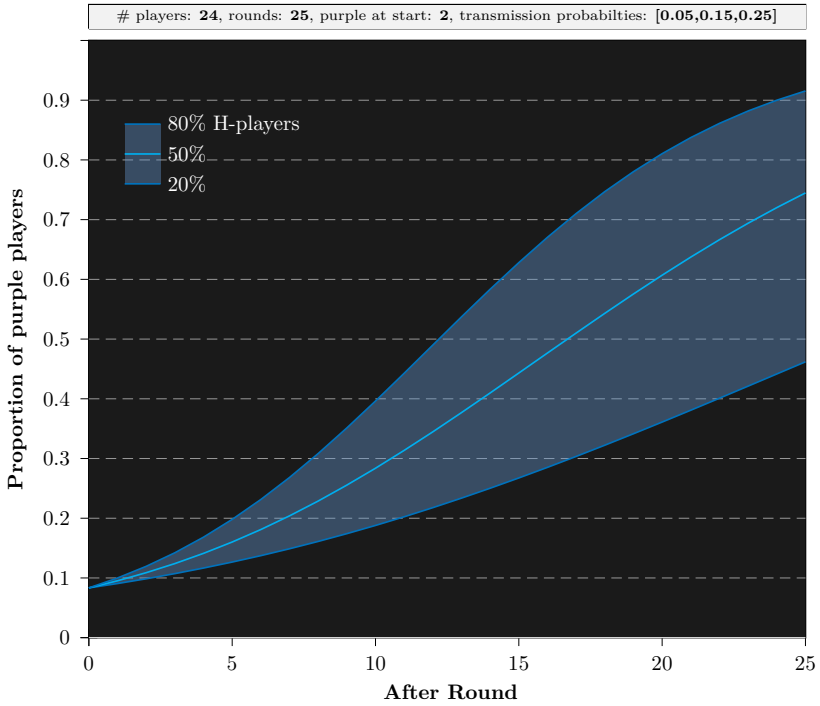


**Figure E91**
*Simulation results for two parameter variations with 24 players (variant 1): Curves in each graph mark results for 20%, 50%, and 80% H-players in the population (from bottom to top).*

### *E14.2.2 Extending the number of rounds to 40*

Another concern might be that the transmission probabilities of 5%, 15% and 25% do not match the risks for a considered class of situations. There are multiple ways to adjust parameters to allow for different probability choices while keeping the range of possible outcome varied enough.

When extending the number of rounds (and decisions) from 25 to 40, the range and size of transmission probabilities can be reduced to 3%, 6%, and 9%, while maintaining a reasonable range of outcomes (see variant 2 in Fig. E92).

Another lever is the number of initial purple participants. Increasing the number can shift the starting point to the right to require less rounds to reach the same state (and more rounds if the number is decreased). Variant 3 shows the consequences of changing the initial number from 8 to 12. This results in a shift of about ten rounds, resulting in higher proportions after round 40.

### *E14.2.3 Lowering the transmission probabilities further*

The number of purple participants in the first round corresponds to the stage of a pandemic that the time frame covers. One might argue that 8 participants out of 100 is already a considerable number of purple participants. Lowering this number can be done

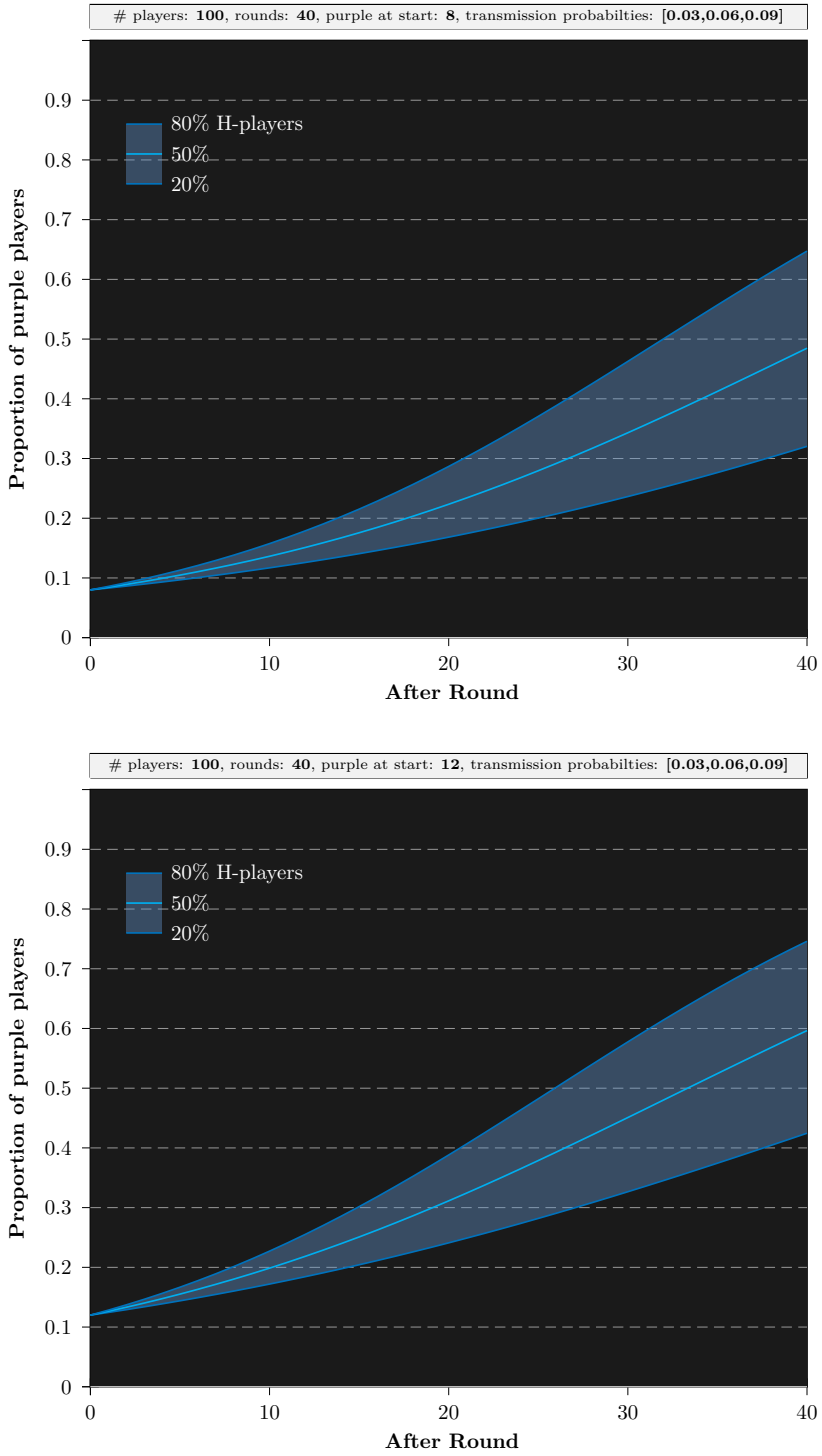


**Figure E92**

*Simulation results for two parameter variations with 40 rounds (variants 2 and 3, top and bottom figure): Curves in each graph mark results for 20%, 50%, and 80% H-players in the population (from bottom to top).*

by extending the length of the game (if one does not want to increase the transmission probability). We investigated the case of 2 initial participants with even lower transmission probabilities than before in variant 4 (see Fig. E93, top), setting them as 2%, 4%, and 6%. We extend the number of rounds to 100 to observe the evolution of infection rates over an extended period.

As expected, the initial infection rate evolves much more slowly than before and reaches 10% at about round 40. Nonetheless, extending the game to a longer number of rounds leads again to an interesting range of results.

Starting the game with 8 purple players as before, the game reaches this state earlier (variant 5 in Fig. E93). This range is smaller than in variant 4 as populations starting with 2 purple individuals will already vary in terms of the round in which they reach 8 purple individuals on average. An interesting range of results can be observed after about 80 rounds.

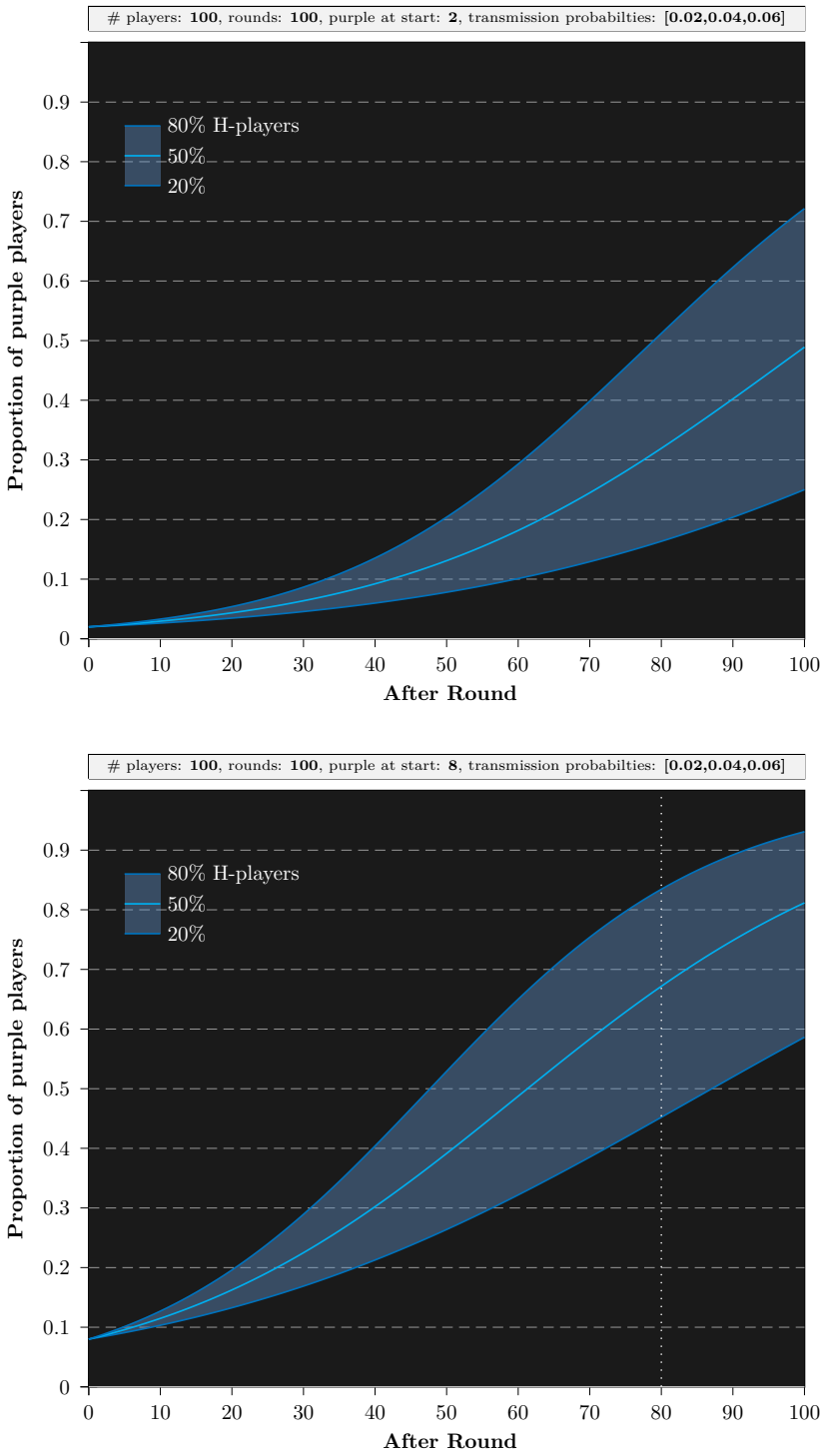


**Figure E93**

*Simulation results for two parameter variations with two purple participants at the start of the game (variants 4 and 5, top and bottom figure): Curves in each graph mark results for 20%, 50%, and 80% H-players in the population (from bottom to top). The dashed line suggests a possible number of rounds for a game based on the parameters in variant 5.*

### *E14.2.4 Further variations*

Finally, we compare different combinations of the variations we introduced. In variant 6, we repeat variant 5 after increasing the transmission probabilities to 3%, 6%, and 9% (see Fig. E94, top). The resulting increase is naturally steeper with an interesting divergence occurring after 60 rounds.

Increasing the transmission probabilities (to 5%, 10%, and 15%) while simultaneously lowering the number of initially infected creates a slower start but even steeper increase and interesting differences at about round 50 (variant 7, see Fig. E94, bottom).

In the final variant, lowering the transmission probabilities to 3%, 6%, and 9% while keeping the number of initially infected at 2 leads to the progression shown in Figure E95. After 75 rounds, there is again a wide range of possible outcomes.

### *E14.2.5 Notes on the choice of parameters*

These simulations demonstrate that none of the chosen parameter values is cast in stone in terms of enabling an experimenter to create relevant variants, in which differences in behavior are translated into a wide range of expected results. Our analytical framework is perfectly capable of dealing with all of the variations we discussed in this section. In our simulations (see **SCR**), we explored further ways of varying the transmission dynamics. For example, if there is even a small, exogenous risk of infection per round per player, transmission proportions can rise fast within fewer rounds. We explored this option but decided against implementing it in the standard version of the transmission game. We found that adding another probability might create confusion and threaten the proper comprehension of all relevant game rules. Nonetheless, for a researcher interested in the effect of risk that depends on one's actions versus the effect of risk that cannot be mitigated, this variant can offer a path.

The standard paradigm is not committed to a fixed interpretation of "round" or the temporal and spatial meaning of one interaction. In return, this offers the possibility to adapt parameters to a specific situation of interest.

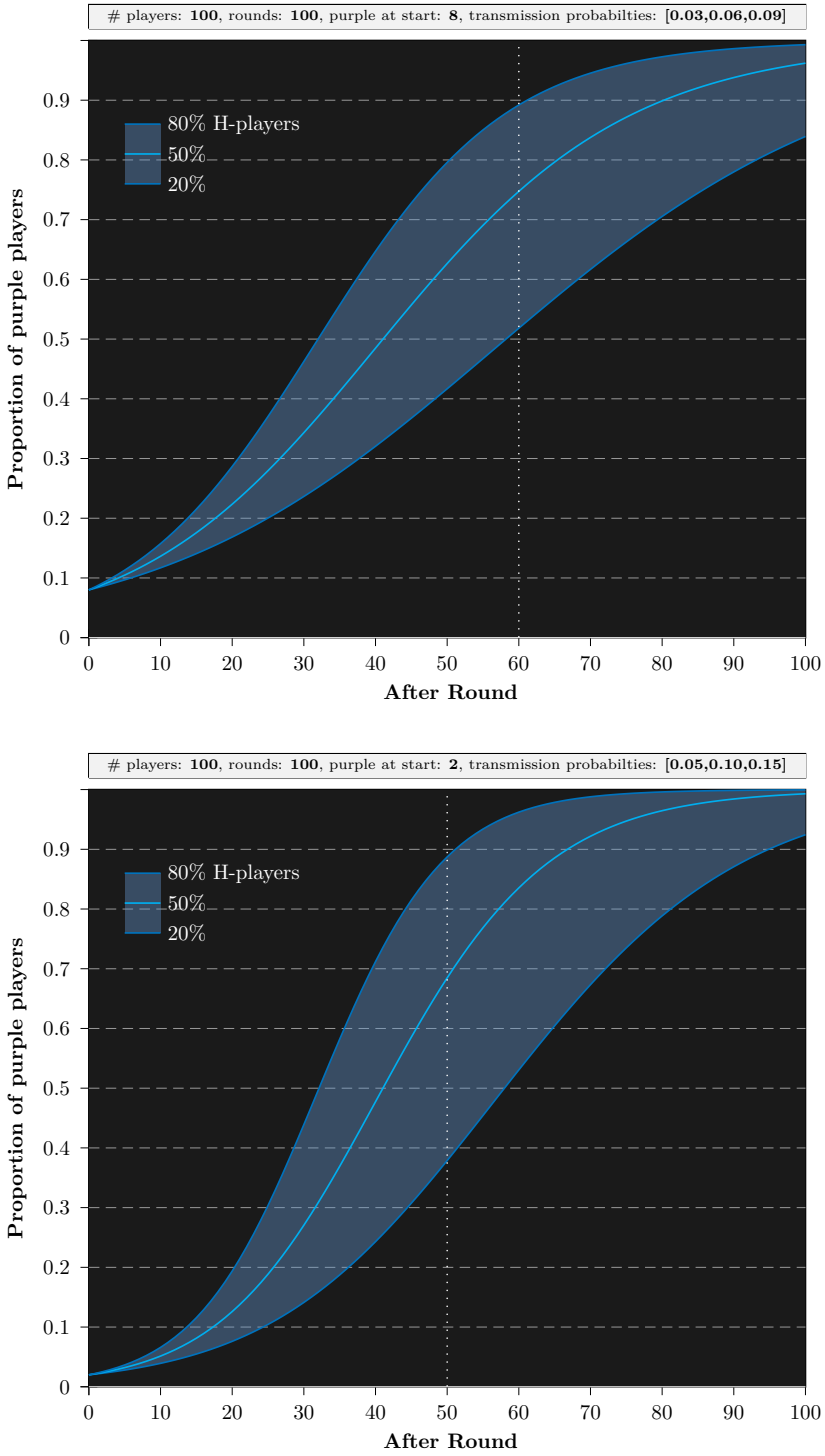


**Figure E94**

*Simulation results for two parameter variations (variants 6 and 7, top and bottom figure): Curves in each graph mark results for 20%, 50%, and 80% H-players in the population (from bottom to top). The dashed lines suggests a possible number of rounds for a game based on these parameters.*

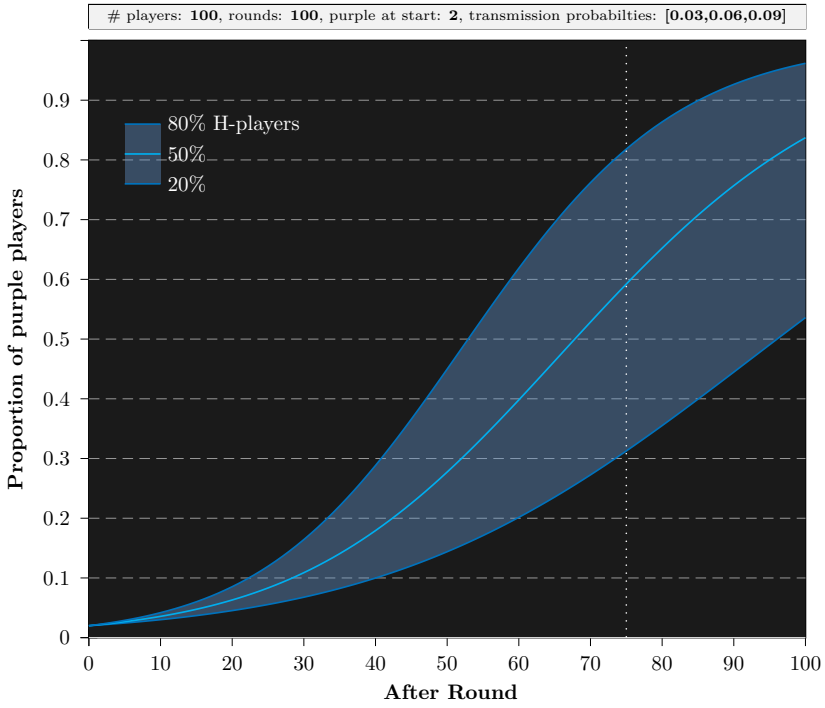


**Figure E95**
*Simulation results for a parameter variation (variant 8): Curves in each graph mark results for 20%, 50%, and 80% H-players in the population (from bottom to top). The dashed line suggests a possible number of rounds for a game based on these parameters.*

## E14.3 Adding elements to the simulation: Infectious periods and immunity

The game framework allows for straightforward additions of elements that may be of interest to different investigators. We advise to carefully consider whether these additions make a sufficient contribution, as any additional rule will add complexity to the game (see our discussion in Section E3.3.6). Here, we demonstrate the impact of adding a component impacting the dynamics of the transmission of infectious diseases.

Our game framework does not allow infected players to revert back to the original state. Moreover, infected players can affect others for the remainder of the game. Here, we offer a variant that sets a round limit for the period in which infected players can infect other players, before they enter a state of immunity. To keep the addition of rules to a minimum, we assume that immune players can neither be infected nor infect others.

### *E14.3.1 Simulation code*

Listing 19 presents the code required for adding this component and simulates a systematic array of conditions to test the consequences of this addition. The value of the variable *maxInfectionRounds* represents the number of rounds after infection, an infected individual is infectious to others. The initially infected individuals are considered to be infectious until the end of this time period.

To explore the dynamics, behavior is varied from 0% H-players to 100% players in steps of 25%, and the length of the infectious period is taking on the values 5, 10, 15, and 25. With 25 rounds, the variant becomes indistinguishable from the original game, as no player can lose the infectious state within the length of the game.

Listing 19: Simulation variant with infectious periods.

```
1
2  %SIMULATION META-PARAMTERS
3  tic;
4  numStrategies=5;
5  numInfectionTimes=4;
6  numTrials=1000000;
7  saveFolder="[Foldername]\res\";
8
9  %GAME PARAMETERS
10 popSize=100;
11 assert(mod(popSize,2)==0);
12
13
14 numRounds=25;
15 infectionRisk=[0.05; 0.15; 0.25];
16 initial=8;
17
18 savePopStateRound=zeros(numRounds+1,numStrategies,
       numInfectionTimes,numTrials);
19 saveInfectiousState=zeros(numRounds+1,numStrategies,
       numInfectionTimes,numTrials);
20
21 %SIMULATION RUN
22 for infectioni=1:numInfectionTimes
23     out=infectioni
24     toc
25 for strategyi= 1:numStrategies
26     out2=strategyi
27     toc
28 for triali=1:numTrials
29
30     pop=zeros(popSize,1);
31     infectionRound=-1000*ones(popSize,1);
32
33     switch strategyi
34         case 1
35             actions=zeros(popSize,numRounds);
36
37         case 2
```

```
38                 actions=zeros(popSize,numRounds);
39                 actionOrd=randperm(popSize)';
40                 actions(actionOrd(1:popSize*0.25,1),1:numRounds)=1;
41         case 3
42                 actions=zeros(popSize,numRounds);
43                 actionOrd=randperm(popSize)';
44                 actions(actionOrd(1:popSize*0.5,1),1:numRounds)=1;
45
46         case 4
47                 actions=zeros(popSize,numRounds);
48                 actionOrd=randperm(popSize)';
49                 actions(actionOrd(1:popSize*0.75,1),1:numRounds)=1;
50         case 5
51                    actions=ones(popSize,numRounds);
52     end
53
54     switch infectioni
55         case 1
56             maxInfectionRounds=5;
57         case 2
58             maxInfectionRounds=10;
59         case 3
60             maxInfectionRounds=15;
61         case 4
62             maxInfectionRounds=25;
63     end
64
65     %INIT POP
66 infectOrd=randperm(popSize)';
67 pop(infectOrd(1:initial, 1))=1;
68 infectionRound(infectOrd(1:initial, 1))=0;
69
70 savePopStateRound(1,strategyi,infectioni,triali)=sum(pop,1);
71 saveInfectiousState(1,strategyi,infectioni,triali)=sum(
      infectionRound(1:popSize,1)>0-maxInfectionRounds,1);
72
73 for ri=1:numRounds
74  matchOrd=randperm(popSize)';
75  for pi=1:popSize/2
76      %test whether individual is infectious
77      if ((pop(matchOrd(pi*2-1,1),1)==1) &&...
78              (ri-infectionRound(matchOrd(pi*2-1,1),1) <=
                   maxInfectionRounds ) &&...
79              (pop(matchOrd(pi*2,1),1)==0)) ||...
80              ((pop(matchOrd(pi*2,1),1)==1) &&...
```

```
81                (ri-infectionRound(matchOrd(pi*2,1),1) <=
                      maxInfectionRounds ) &&...
82                (pop(matchOrd(pi*2-1,1),1)==0))
83
84            countH=actions(pi*2-1,ri)+actions(pi*2,ri);
85            infectP=infectionRisk(countH+1,1);
86            drawR=rand(1,1);
87            if (drawR<infectP)
88                 if ( pop(matchOrd(pi*2-1,1),1)==0)
89                    pop(matchOrd(pi*2-1,1),1)=1;
90                    infectionRound(matchOrd(pi*2-1,1),1)=ri;
91
92                 else
93                    pop(matchOrd(pi*2,1),1)=1;
94                    infectionRound(matchOrd(pi*2,1),1)=ri;
95                 end
96                end
97            end
98       end
99      savePopStateRound(ri+1,strategyi,infectioni,triali)=sum(pop
            ,1);
100     saveInfectiousState(ri+1,strategyi,infectioni,triali)=sum(
            infectionRound(1:popSize,1)>=ri-maxInfectionRounds,1);
101
102  end
103
104 end%Trials
105 end %STRATEGY
106 end%INFECTION
107
108 meanPopStateRound=mean(savePopStateRound,4);
109 meanInfectiousStateRound=mean(saveInfectiousState,4);
110
111 for ci=1:numInfectionTimes
112    saveMatrix1= meanPopStateRound(:,:,ci);
113    saveMatrix2= meanInfectiousStateRound(:,:,ci);
114
115    dlmwrite(strcat(saveFolder,'ResPop--Infection',int2str(ci),'.
           txt'),saveMatrix1,'delimiter','\t','precision',12);
116    dlmwrite(strcat(saveFolder,'ResInfectious--Infection',int2str
           (ci),'.txt'),saveMatrix2,'delimiter','\t','precision',12);
117 end
118 toc
```

### *E14.3.2 Simulation results*

The results are distributed across five figures to allow to assess the impact of variation of the infectious time period for a given player behavior. For 0% H-players the results are shown in Figure E96, for 25% H-players in Figure E97, for 50% H-players in Figure E98, for 75% H-players in Figure E99, for 100% H-players in Figure E100.

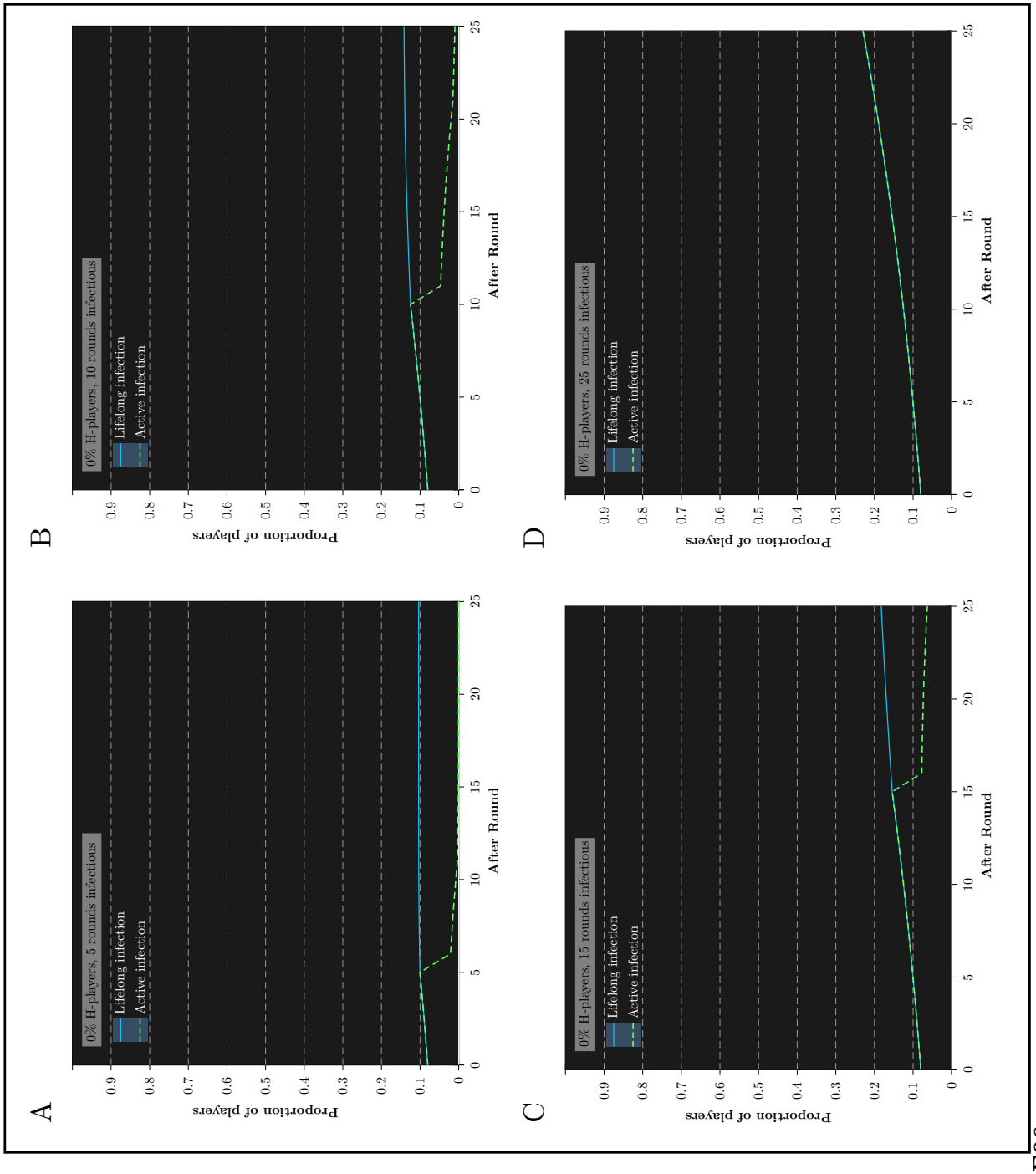


**Figure E96**

*Results for simulations with infectious periods and 0% of players choosing H-actions in each round and a maximum number of infectious rounds set at (A) 5, (B) 10, (C) 15, and (D) 25. Lines represent the total number of infectious and recovered players (blue solid line) and the number of players that can infect others (green dashed line).*

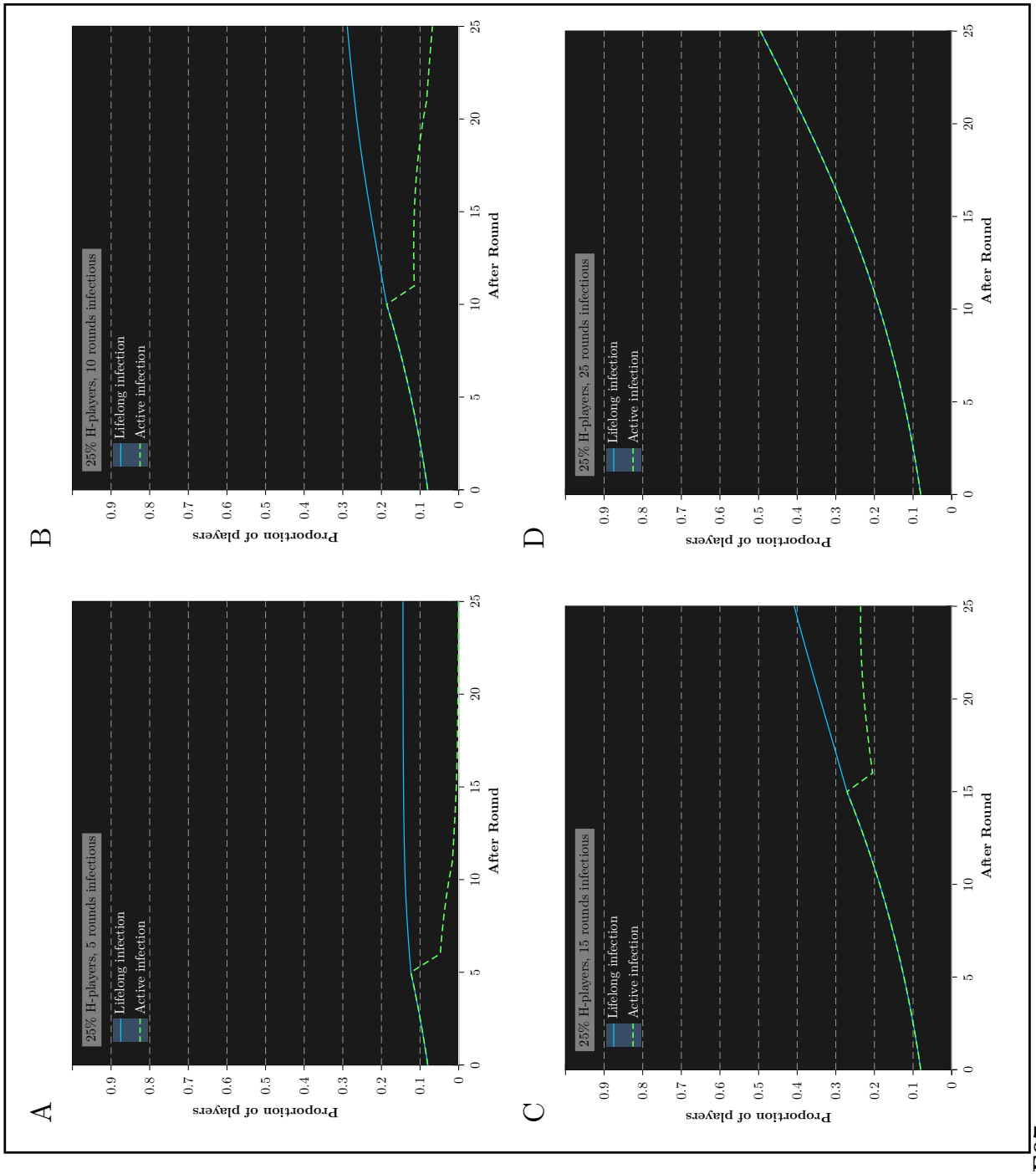


**Figure E97**

*Results for simulations with infectious periods and 25% of players choosing H-actions in each round and a maximum number of infectious rounds set at (A) 5, (B) 10, (C) 15, and (D) 25. Lines represent the total number of infectious and recovered players (blue solid line) and the number of players that can infect others (green dashed line).*

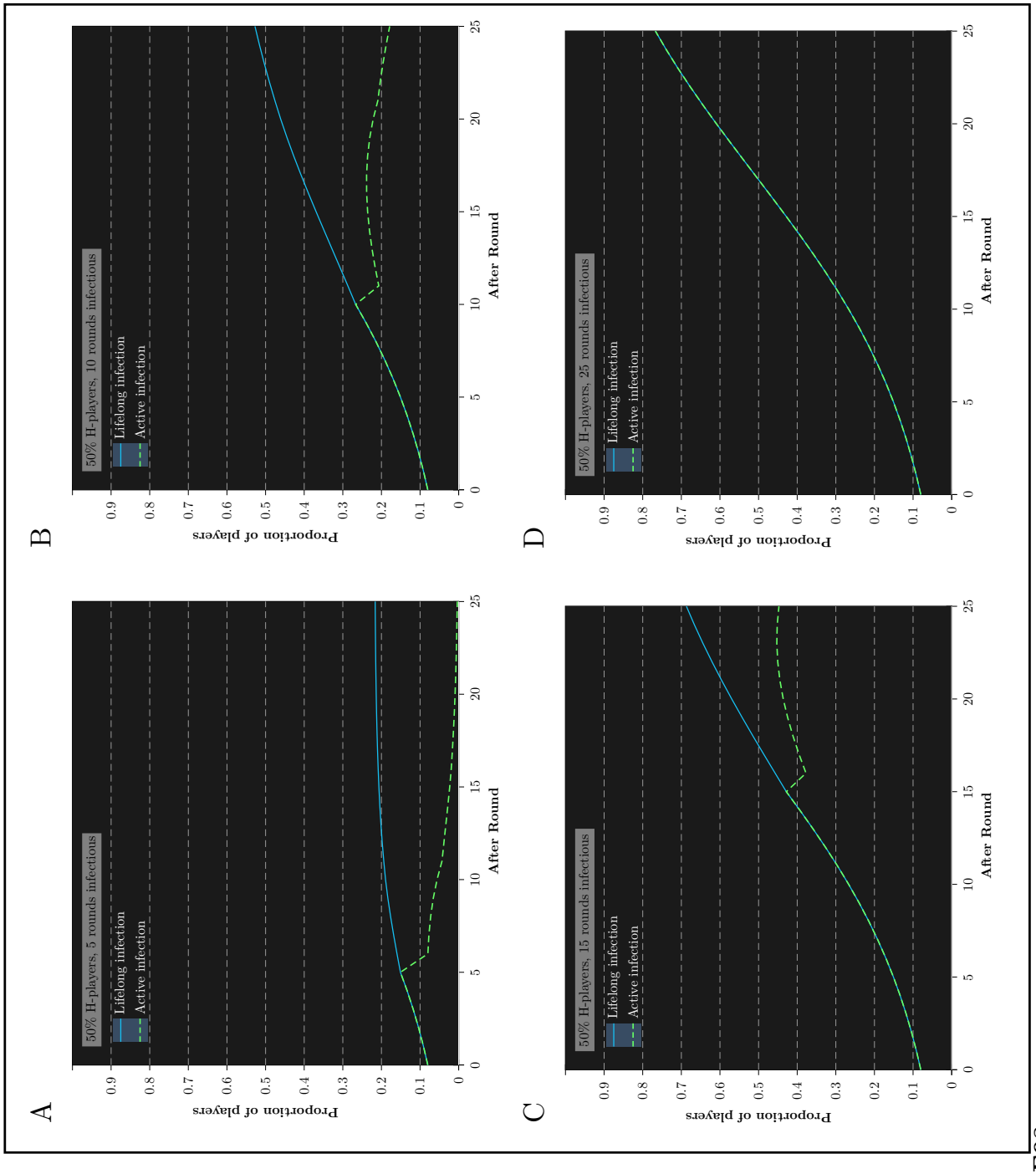


**Figure E98**

*Results for simulations with infectious periods and 50% of players choosing H-actions in each round and a maximum number of infectious rounds set at (A) 5, (B) 10, (C) 15, and (D) 25. Lines represent the total number of infectious and recovered players (blue solid line) and the number of players that can infect others (green dashed line).*

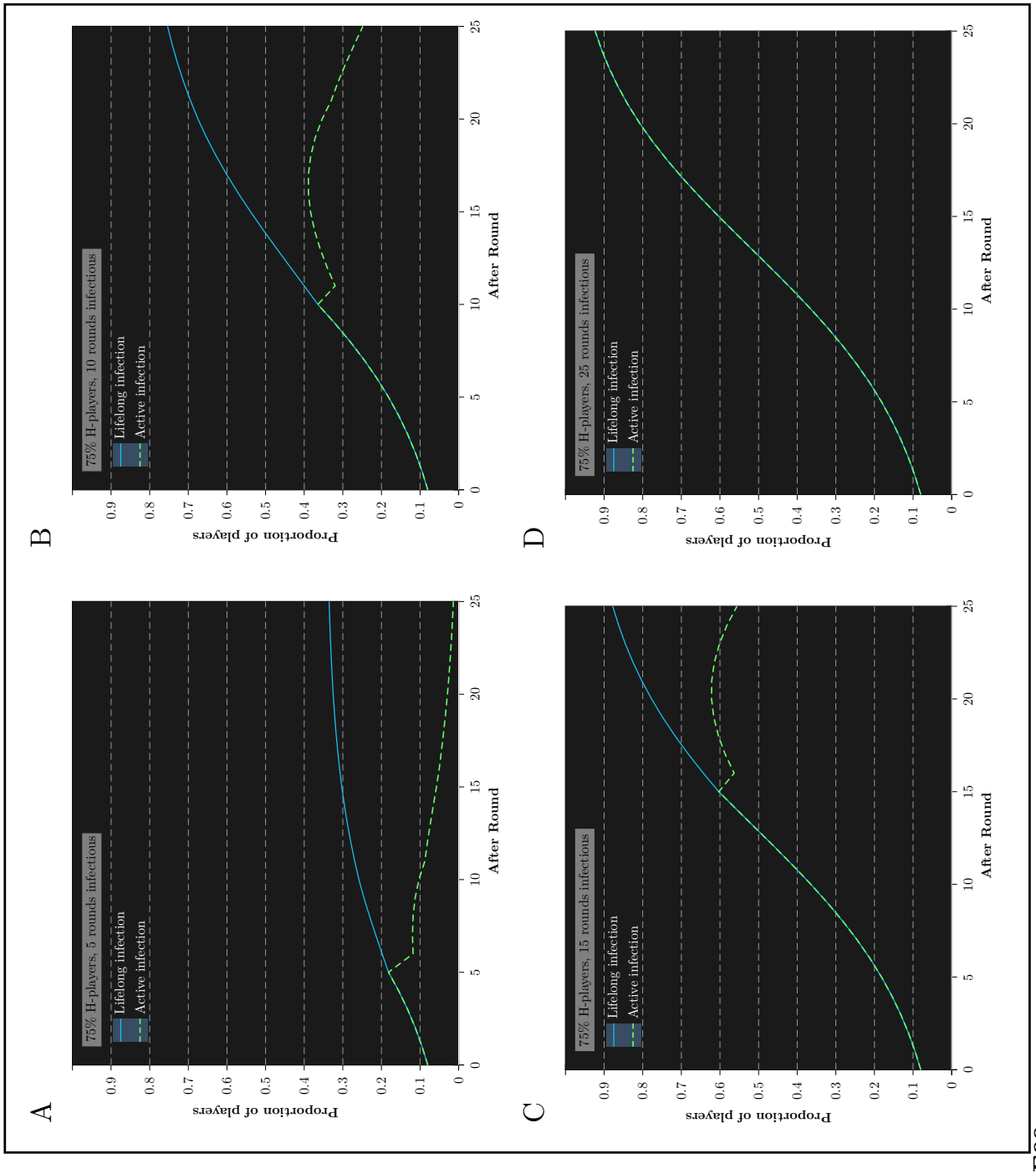


**Figure E99**

*Results for simulations with infectious periods and 75% of players choosing H-actions in each round and a maximum number of infectious rounds set at (A) 5, (B) 10, (C) 15, and (D) 25. Lines represent the total number of infectious and recovered players (blue solid line) and the number of players that can infect others (green dashed line).*

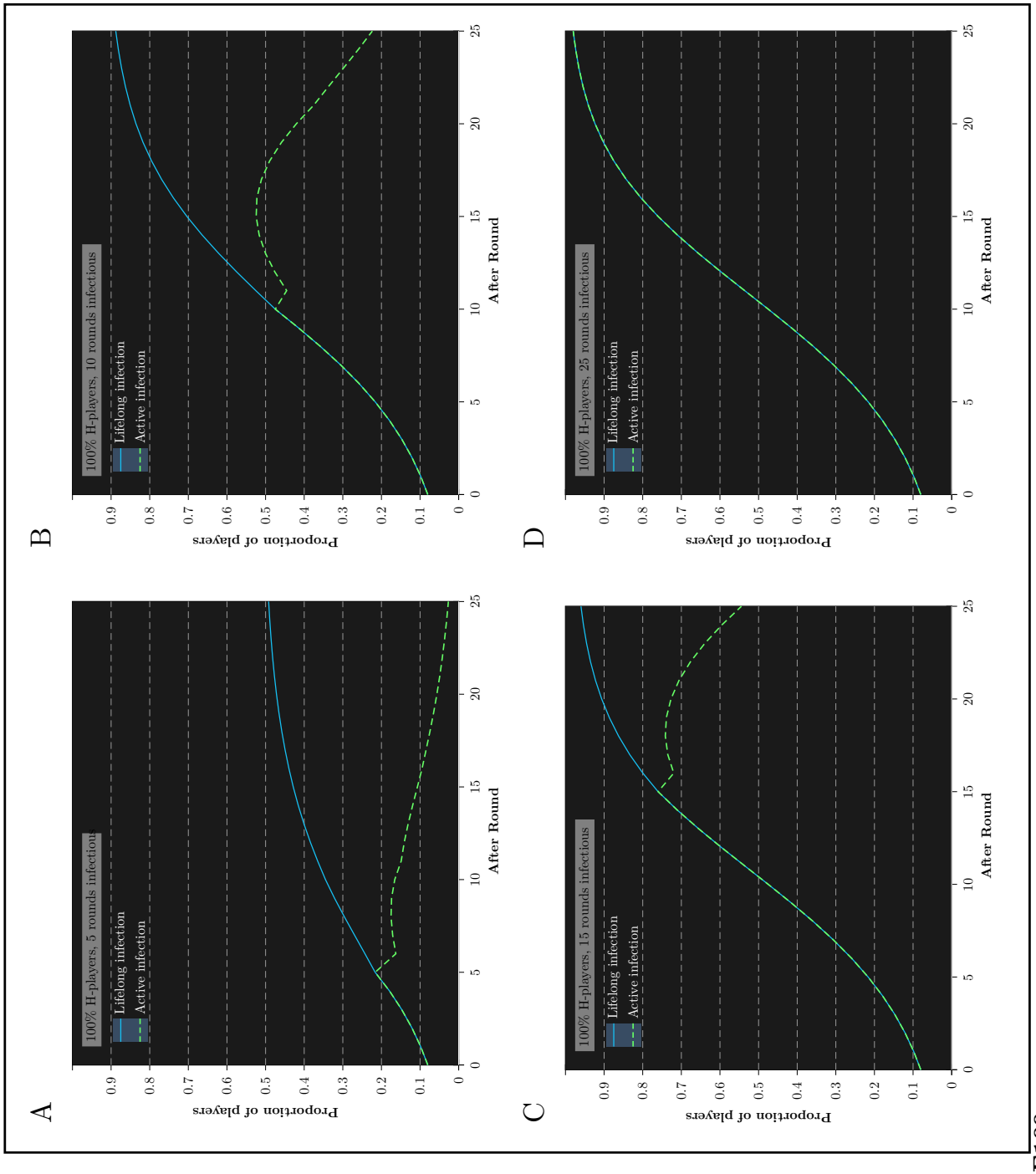


**Figure E100**

*Results for simulations with infectious periods and 100% of players choosing H-actions in each round and a maximum number of infectious rounds set at (A) 5, (B) 10, (C) 15, and (D) 25. Lines represent the total number of infectious and recovered players (blue solid line) and the number of players that can infect others (green dashed line).*

#### *E14.3.3 Discussion*

Not surprisingly, final lifetime infection rates in general are lower when the infectious period is shortened. At the same time, the differences between original and new version are not dramatic for infectious periods of ten rounds or longer, and the shape of the curves does not undergo qualitative changes that would introduce concerns about results obtained with the original game variant.

Adding infectious periods would allow to explore interventions that focus specifically on the timing of critical behaviors. Contact tracing, self-isolation, repeated testing, and other warning systems are created to recommend restrictions for particular points in time. The presented version of the simulation could allow for integrating warning signal (e.g., while using the strategy method and asking participants to indicate behavioral choices contingent on a signal).

This variant might be accompanied by additional behavioral choices, such as an equivalent self-isolation with an even lower payoff than G-actions. We would add the warning, though, that feature proliferation can easily create a scenario whose complexity might overburden the desired target population (see also Section E3.3.6).

### E14.4 A look forward

So what lies ahead for our paradigm? Clear and appropriate messaging and decision-making by authorities was able to change behavior and increase the degree of social distancing in a population (Abouk & Heydari, 2020; Barari et al., 2020; Goldberg et al., 2020; Malik et al., 2020). Yet, much remains unknown about the large differences between countries (Kenyon, 2020), possible negative consequences of improper adherence (Barari et al., 2020), the degree of and predictors of interindividual and intertemporal variation in responding to messages (Pink et al., 2020; Van Bavel et al., 2020). In open-format responses to our intervention, many participants expressed a willingness to follow the recommendation (e.g., "The message served as a cautionary warning for me"), but some participants rejected the statement (e.g., "The message seems pushy"). This indicates that our paradigm might be a stepping stone for obtaining answers to at least some of these open questions. Another promising avenue for future research is a closer analysis of the influence of political ideology and party affiliation on behavior during the pandemic and in our paradigm (compare Clinton et al., 2020; Collins et al., 2020).

Due to its neutral phrasing, the paradigm is not wedded to a single issue or measure. Its abstract nature makes it perfectly translatable into all other languages. Its moderate complexity makes it appropriate for most members of the population. Its flexibility allows for many manipulations and variants for various group sizes, and offline and online contexts. There is thus justified hope that findings are relevant beyond the scope of the pandemic, but also for future challenges within this scope. For example, the uptake of vaccines crucially depends on the effectiveness of messaging (compare Bauch & Earn, 2004). Our paradigm or variants thereof offer the benefit of ethical experimentation focusing on questions of central importance, while simultaneously stepping up from computer simulations to behavioral observation. Thus we are hopeful that it will be added as a useful tool to the toolbox of measures that are applicable in a time of crisis.

## References

Abdelrahman, M. K. (2020). Personality traits, risk perception, and protective behaviors of Arab residents of Qatar during the COVID-19 pandemic (April 14). *PsyArXiv*. https://doi.org/10.1007/s11469-020-00352-7

Abouk, R., & Heydari, B. (2020). The immediate effect of COVID-19 policies on social distancing behavior in the United States. *medRxiv*. https://doi.org/10.1101/2020.04.07.20057356

Adami, C. (1998). *Introduction to artificial life*. Springer Science & Business Media.

Aleta, A., Martín-Corral, D., Pastore y Piontti, A., Ajelli, M., Litvinova, M., Chinazzi, M., Dean, N. E., Halloran, M. E., Longini Jr, I. M., Merler, S., Pentland, A., Vespignani, A., Moro, E., & Moreno, Y. (2020). Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour*, *4*, 964–971. https://doi.org/10.1038/s41562-020-0931-9

Alper, S., Bayrak, F., & Yilmaz, O. (2020). Psychological correlates of COVID-19 conspiracy beliefs and preventive measures: Evidence from Turkey (April 29). *PsyArXiv*. https://doi.org/10.31234/osf.io/mt3p4

Al-Ubaydli, O., Lee, M. S., List, J. A., Mackevicius, C. L., & Suskind, D. (2020). How can experiments play a greater role in public policy? twelve proposals from an economic model of scaling. *Behavioural Public Policy*. https://doi.org/10.1017/bpp.2020.17

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct (2003, amended June 1, 2010 and January 1, 2017). https://www.apa.org/ethics/code

Ammar, A., Trabelsi, K., Brach, M., Chtourou, H., Boukhris, O., Masmoudi, L., Bouaziz, B., Bentlage, E., How, D., Ahmed, M., Mueller, P., Mueller, N., Aloui, A., Hammouda, O., Paineiras-Domingos, L. L., Braakman-jansen, A., Wrede, C., Bastoni, S., Pernambuco, C. S., ... Hoekelmann, A. (2020). Effects of home confinement on mental health and lifestyle behaviours during the COVID-19 outbreak: Insight from the "ECLB-COVID19" multi countries survey. *medRxiv*. https://doi.org/10.1101/2020.05.04.20091017

Ammerman, B. A., Burke, T. A., Jacobucci, R., & McClure, K. (2020). Preliminary investigation of the association between COVID-19 and suicidal thoughts and behaviors in the U.S. (April 6). *PsyArXiv*. https://doi.org/10.31234/osf.io/68djp

Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, *21*(1), 99–131. https://doi.org/10.1007/s10683-017-9527-2

Ariely, D., & Norton, M. I. (2007). Psychology and Experimental Economics: A Gap in Abstraction. *Current Directions in Psychological Science*, *16*(6), 336–339. https://doi.org/10.1111/j.1467-8721.2007.00531.x

Arrow, K. (1971). *Essays in the theory of risk-bearing*. Markham.

Ashton, M. C., & Lee, K. (2009). The HEXACO–60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*(4), 340–345. https://doi.org/110.1080/00223890902935878

Azim, D., Kumar, S., Nasim, S., Arif, T. B., & Nanjiani, D. (2020). Covid-19 as a psychological contagion: A new pandoras box to close? *Infection Control & Hospital Epidemiology*, *41*(9), 989–990. https://doi.org/10.1017/ice.2020.127

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations*, *12*(4), 533–547. https://doi.org/10.1177/1368430209105040

Balliet, D., & Van Lange, P. A. M. (2012). Trust, conflict, and cooperation: A meta-analysis. *Psychological Bulletin*, *139*(5), 1090–1112. https://doi.org/10.1037/a0030939

Bandura, A., & Walters, R. H. (1977). *Social learning theory*. Prentice-Hall.

Barari, S., Caria, S., Davola, A., Falco, P., Fetzer, T., Fiorin, S., Hensel, L., Ivchenko, A., Jachimowicz, J., King, G., Kraft-Todd, G., Ledda, A., MacLennan, M., Mutoi, L., Pagani, C., Reutskaja, E., & Slepoi, F. R. (2020). Evaluating COVID-19 public health messaging in Italy: Self-reported compliance and growing mental health concerns. *medRxiv*. https://doi.org/10.1101/2020.03.27.20042820

Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., & Sugden, R. (2020). *Experimental economics: Rethinking the rules*. Princeton University Press.

Bauch, C. T., & Earn, D. J. (2004). Vaccination and the theory of games. *Proceedings of the National Academy of Sciences*, *101*(36), 13391–13394. https://doi.org/10.1073/pnas.0403823101

Baumgartner, M. T., Lansac-Toha, F. M., Coelho, M. T. P., Dobrovolski, R., & Diniz-Filho, J. A. F. (2020). Social distancing and movement constraint as the most likely factors for COVID-19 outbreak control in Brazil. *medRxiv*. https://doi.org/10.1101/2020.05.02.20088013

Beall, A. T., Hofer, M. K., & Schaller, M. (2016). Infections and elections: Did an Ebola outbreak influence the 2014 US federal elections (and if so, how)? *Psychological Science*, *27*(5), 595–605. https://doi.org/10.1177/0956797616628861

Bertin, P., Nera, K., & Delouvée, S. (2020). Conspiracy beliefs, chloroquine, and the rejection of vaccination: A conceptual replication-extension in the COVID-19 pandemic context (May 22). *PsyArXiv*. https://doi.org/10.31234/osf.io/rz78k

Betsch, C., Korn, L., Sprengholz, P., Felgendreff, L., Eitze, S., Schmid, P., & Böhm, R. (2020). Social and behavioral consequences of mask policies during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *117*(36), 21851–21853. https://doi.org/10.1073/pnas.2011674117

Bicchieri, C., & Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public Choice*. https://doi.org/10.1007/s11127-019-00684-6

Blagov, P. S. (2020). Adaptive and dark personality traits in the Covid-19 pandemic: Predicting health-behavior endorsement and the appeal of public-health messages. *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550620936439

Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour*, *4*, 588–596. https://doi.org/10.1038/s41562-020-0898-6

Böhm, R., Betsch, C., & Korn, L. (2016). Selfish-rational non-vaccination: Experimental evidence from an interactive vaccination game. *Journal of Economic Behavior & Organization*, *131*, 183–195. https://doi.org/10.1016/j.jebo.2015.11.008

Boin, A., Ekengren, M., & Rhinard, M. (2020). Hiding in plain sight: Conceptualizing the creeping crisis. *Risk, Hazards & Crisis in Public Policy*. https://doi.org/10.1002/rhc3.12193

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–298. https://doi.org/10.1038/nature11421

Bostrom, A., Böhm, G., OConnor, R. E., Hanss, D., Bodi-Fernandez, O., & Halder, P. (2020). Comparative risk science for the coronavirus pandemic. *Journal of Risk Research*, *23*(7–8), 902–911. https://doi.org/10.1080/13669877.2020.1756384

Both, F., Adam, M. T. P., Hariharan, A., Dorner, V., Lux, E., & Weinhardt, C. (2016). A randomized tabu search-based approach for perfect stranger matching in economic experiments. *Economics Letters*, *145*, 235–238. https://doi.org/10.1016/j.econlet.2016.06.028

Brandts, J., & Charness, G. (2000). Hot vs. cold: Sequential responses and preference stability in experimental games. *Experimental Economics*, *2*(3), 227–238. https://doi.org/10.1023/A:1009962612354

Brandts, J., & Charness, G. (2011). The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics*, *14*(3), 375–398. https://doi.org/10.1007/s10683-011-9272-x

Brauner, J. M., Mindermann, S., Sharma, M., Stephenson, A. B., Gaveniak, T., Johnston, D., Leech, G., Salvatier, J., Altman, G., Norman, A. J., Monrad, J. T., Besiroglu, T., Ge, H., Mikulik, V., Hartwick, M. A., Teh, Y. W., Chindelevitch, L., Gal, Y., & Kulveit, J. (2020). The effectiveness of eight nonpharmaceutical interventions against COVID-19 in 41 countries. *medRxiv*. https://doi.org/10.1101/2020.05.28.20116129

Briki, W., & Dagot, L. (2020). "Liberate!" - Republicans are more willing to get back to social life because they are less scared of COVID-19 (May 29). *PsyArXiv*. https://doi.org/10.31234/osf.io/thuwd

Brutger, R., Kertzer, J. D., Renshon, J., Tingley, D., & Weiss, C. M. (2020). Abstraction and detail in experimental design. https://www.people.fas.harvard.edu/~jkertzer/Research_files/AbstractConcrete%20091620_Website.pdf

Brzezinski, A., Kecht, V., van Dijcke, D., & Wright, A. L. (2020). Belief in science influences physical distancing in response to COVID-19 lockdown policies. *SSRN*. https://doi.org/10.2139/ssrn.3587990

Burger, J. M., & Shelton, M. (2011). Changing everyday health behaviors through descriptive norm manipulations. *Social Influence*, *6*(2), 69–77. https://doi.org/10.1080/15534510.2010.542305

Campos-Mercade, P., Meier, A., Schneider, F., & Wengström, E. (2020). Prosociality predicts health behaviors during the COVID-19 pandemic (May 17, 2020) [University of Zurich, Department of Economics, Working Paper No. 346]. *SSRN*. https://doi.org/10.2139/ssrn.3604094

Cao, S., Qi, Y., Huang, Q., Wang, Y., Han, X., Liu, X., & Wu, H. (2020). Emerging infectious outbreak inhibits pain empathy mediated prosocial behaviour (March 29). *PsyArXiv*. https://doi.org/10.31234/osf.io/5p8kd

Capraro, V., & Barcelo, H. (2020). The effect of messaging and gender on intentions to wear a face covering to slow down COVID-19 transmission (May 11). *PsyArXiv*. https://doi.org/10.31234/osf.io/tg7vz

Cavojova, V., Srol, J., & Mikusková, E. B. (2020). Scientific reasoning as a predictor of health-related beliefs and behaviors in the time of COVID-19 (April 15). *PsyArXiv*. https://doi.org/10.31234/osf.io/tfy5q

Centers for Disease Control and Prevention (CDC). (2000). Measuring healthy days: Population assessment of health-related quality of life. http://www.cdc.gov/hrqol/pdfs/mhd.pdf

Centola, D. (2011). An experimental study of homophily in the adoption of health behavior. *Science*, *334*(6060), 1269–1272. https://doi.org/10.1126/science.1207055

Cha, A. E. (2021, August 28). A Calif. elementary school teacher took off her mask for a read-aloud. Within days, half her class was positive for delta. https://www.washingtonpost.com/health/2021/08/28/delta-variant-unvaccinated-children-elementary-schools

Chater, N. (2020). Facing up to the uncertainties of COVID-19. *Nature Human Behaviour*, *4*(5), 439–439. https://doi.org/710.1038/s41562-020-0865-2

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., y Piontti, A. P., Mu, K., Rossi, L., Sun, K., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, *368*(6489), 395–400. https://doi.org/10.1126/science.aba9757

Cho, H., Lee, J.-S., & Lee, S. (2013). Optimistic bias about H1N1 flu: Testing the links between risk communication, optimistic bias, and self-protection behavior. *Health Communication*, *28*(2), 146–158. https://doi.org/10.1080/10410236.2012.664805

Christakis, N. A., & Fowler, J. H. (2013). Social contagion theory: Examining dynamic social networks and human behavior. *Statistics in Medicine*, *32*(4), 556–577. https://doi.org/10.1002/sim.5408

Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science*, *12*(4), 105–109. https://doi.org/10.1111/1467-8721.01242

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Clinton, J., Cohen, J., Lapinski, J. S., & Trussler, M. (2020). Partisan pandemic: How partisanship and public health concerns affect individuals' social distancing during COVID-19 (July 9). *SSRN*. https://doi.org/10.2139/ssrn.3633934

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making*, *7*(1), 25–47.

Collins, R. N., Mandel, D. R., & Schywiola, S. S. (2020). Political identity over personal impact: Early US reactions to the COVID-19 pandemic (September 17). *PsyArXiv*. https://doi.org/10.31234/osf.io/jeq6y

Columbus, S. (2020). Who hoards? Honesty-humility and behavioural responses to the 2019/20 coronavirus pandemic (March 24). *PsyArXiv*. https://doi.org/10.31234/osf.io/8e62v

Columbus, S., & Böhm, R. (2021). Norm shifts under the strategy method. *Judgment and Decision Making*, *16*(5), 1267–1289. http://journal.sjdm.org/21/210609b/jdm210609b.pdf

Contreras, S., Dehning, J., Mohr, S. B., Bauer, S., Spitzner, F. P., & Priesemann, V. (2021). Low case numbers enable long-term stable pandemic control without lockdowns. *Science Advances*, *7*(41). https://doi.org/10.1126/sciadv.abg2243

Conway, L. G., Woodard, S. R., Zubrod, A., & Chan, L. (2020). Why are conservatives less concerned about the coronavirus (COVID-19) than liberals? Testing experiential versus political explanations (April 13). *PsyArXiv*. https://doi.org/10.31234/osf.io/fgb84

Davies, N. G., Kucharski, A. J., Eggo, R. M., Gimma, A., & Edmunds, W. J. (2020). Eects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the uk: A modelling study. *The Lancet Public Health*, *5*, e375–e385. https://doi.org/10.1016/S2468-2667(20)30133-X

De Cremer, D., & Van Lange, P. A. M. (2001). Why prosocials exhibit greater cooperation than proselfs: The roles of social responsibility and reciprocity. *European Journal of Personality*, *15*(1_suppl), S5–S18. https://doi.org/10.1002/per.418

de Neys, W., Raoelison, M., Boissin, E., Voudouri, A., Bago, B., & Bialek, M. (2020). Moral outrage and social distancing: Bad or badly informed citizens? (May 4). *PsyArXiv*. https://doi.org/10.31234/osf.io/j9h76

De Vries, R. E. (2013). The 24-item brief HEXACO inventory (BHI). *Journal of Research in Personality*, *47*(6), 871–880. https://doi.org/10.1016/j.jrp.2013.09.003

Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M., & Priesemann, V. (2020). Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*, *369*, eabb9789. https://doi.org/10.1126/science.abb9789

De-Leon, H., & Pederiva, F. (2020). Particle modeling of the spreading of coronavirus disease. *Physical Fluids*, *32*(8), 087113. https://doi.org/10.1063/5.0020565

Dennis, S., Lewandowsky, S., Lorenz-Spreen, P., Oberauer, K., Okan, Y., Goldstone, R., Cheng-Ta, Y., Kashima, Y., Perfors, A., White, J., Garrett, P., Geard, N., Little, D., Mitchel, L., Tomko, M., Kozyreva, A., Herzog, S., Hertwig, R., & Pachur, T. (2020). Social licensing of privacy-encroaching policies to address the COVID-19 pandemic. https://stephanlewandowsky.github.io/UKsocialLicence/index.html

Dickie, R., Rasmussen, S., Cain, R., Williams, L., & MacKay, W. (2018). The effects of perceived social norms on handwashing behaviour in students. *Psychology, Health & Medicine*, *23*(2), 154–159. https://doi.org/10.1080/13548506.2017.1338736

DiStefano III, J. (2015). *Dynamic systems biology modeling and simulation*. Academic Press.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550. https://doi.org/10.1111/j.1542-4774.2011.01015.x

Dorison, C., Lerner, J. S., Heller, B. H., Rothman, A., Kawachi, I. I., Wang, K., Rees, V. W., Gill, B. P., Gibbs, N., Psychological Science Accelerator COVID-19 Rapid Team, & Coles, N. A. (2020). A global test of message framing on behavioural intentions, policy support, information seeking, and experienced anxiety during the COVID-19 pandemic (April 16). *PsyArXiv*. https://doi.org/10.31234/osf.io/sevkf

Doukidis, G. I., & Angelides, M. C. (1994). A framework for integrating artificial intelligence and simulation. *Artificial Intelligence Review*, *8*(1), 55–85.

Drew, D. A., Nguyen, L. H., Steves, C. J., Menni, C., Freydin, M., Varsavsky, T., Sudre, C. H., Cardoso, M. J., Ourselin, S., Wolf, J., et al. (2020). Rapid implementation of mobile technology for real-time epidemiology of COVID-19. *Science*, *368*, 1362–1367. https://doi.org/10.1126/science.abc0473

Du, H., Yang, J., King, R. B., Yang, L., & Chi, P. (2020). COVID-19 increases online emotional and health-related searches (May 2). *PsyArXiv*. https://doi.org/10.31234/osf.io/5gskw

Engel, C. (2011). Dictator games: A meta study. *Experimental economics*, *14*(4), 583–610. https://doi.org/10.1007/s10683-011-9283-7

Erceg, N., Ruoji, M., & Galic, Z. (2020). Misbehaving in the corona crisis: The role of anxiety and unfounded beliefs (April 10). *PsyArXiv*. https://doi.org/10.31234/osf.io/cgjw8

Erev, I., Plonsky, O., & Roth, Y. (2020). Complacency, panic, and the value of gentle rule enforcement in addressing pandemics. *Nature Human Behaviour*, *4*, 1095–1097. https://doi.org/10.1038/s41562-020-00939-z

Evans, A. M., Kogler, C., & Sleegers, W. W. (2021). No effects of synchronicity in online social dilemma experiments: A registered report. *Judgment and Decision Making*, *16*(4), 823–843. http://journal.sjdm.org/18/18503/jdm18503.pdf

Everett, J. A. C. (2013). The 12 item social and economic conservatism scale (SECS). *PLOS ONE*, *8*(12), e82131. https://doi.org/10.1371/journal.pone.0082131

Everett, J. A. C., Colombatto, C., Chituc, V., Brady, W. J., & Crockett, M. (2020). The effectiveness of moral messages on public health behavioral intentions during the COVID-19 pandemic (March 20). *PsyArXiv*. https://doi.org/10.31234/osf.io/9yqs8

Evershed, N., & Ball, A. (2020, April 22). How coronavirus spreads through a population and how we can beat it. https://www.theguardian.com/world/datablog/ng-interactive/2020/apr/22/see-how-coronavirus-can-spread-through-a-population-and-how-countries-flatten-the-curve

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, *133*(4), 1645–1692. https://doi.org/10.1093/qje/qjy013

Fansher, M., Adkins, T., Lalwani, P., Quirk, M., Boduroglu, A., Lewis, R., Shah, P., & Jonides, J. (2020). How well do ordinary americans forecast the growth of COVID-19? (May 19). *PsyArXiv*. https://doi.org/10.31234/osf.io/2d5r9

Farias, J. E. M., & Pilati, R. (2020). Violating social distancing amid the COVID-19 pandemic: Psychological factors to improve compliance (April 29). *PsyArXiv*. https://doi.org/10.31234/osf.io/apg9e

Favero, N., & Pedersen, M. J. (2020). How to encourage "togetherness by keeping apart" amid COVID-19? The ineffectiveness of prosocial and empathy appeals. *Journal of Behavioral Public Administration*, *3*(2). https://doi.org/10.30636/jbpa.32.167

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Ferguson, N. M., Cummings, D. A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S., & Burke, D. S. (2005). Strategies for containing an emerging influenza pandemic in southeast asia. *Nature*, *437*, 209–214. https://doi.org/10.1038/nature04017

Ferguson, N. M., Cummings, D. A., Fraser, C., Cajka, J. C., Cooley, P. C., & Burke, D. S. (2006). Strategies for mitigating an influenza pandemic. *Nature*, *442*(7101), 448–452. https://doi.org/10.1038/nature04795

Ferguson, N., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., A, D., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L. C., van Elsland, S., ... Ghani, A. C. (2020). Impact of non-pharmaceutical interventions (npis) to reduce COVID-19 mortality and healthcare demand [Imperial College COVID-19 Response Team, London, March, 16]. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., & Fraser, C. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, *368*(6491), eabb6936. https://doi.org/10.1126/science.abb6936

Fielding-Miller, R. K., Sundaram, M. E., & Brouwer, K. (2020). Social determinants of COVID-19 mortality at the county level. *medRxiv*. https://doi.org/10.1101/2020.05.03.20089698

Filgueiras, A., & Stults-Kolehmainen, M. (2020). Factors linked to changes in mental health outcomes among brazilians in quarantine due to COVID-19. *medRxiv*. https://doi.org/10.1101/2020.05.12.20099374

Fisher, K. A., Bloomstone, S. J., Walder, J., Crawford, S., Fouayzi, H., & Mazor, K. M. (2020). Attitudes toward a potential SARS-CoV-2 vaccine: A survey of U.S. adults. *Annals of Internal Medicine*. https://doi.org/10.7326/M20-3569

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., Monod, M., Ghani, A. C., Donnelly, C. A., Riley, S., Vollmer, M. A. C., Ferguson, N. M., Okell, L. C., & Bhatt, S. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. https://doi.org/10.1038/s41586-020-2405-7

Flesia, L., Fietta, V., Colicino, E., Segatto, B., & Monaro, M. (2020). Stable psychological traits predict perceived stress related to the COVID-19 outbreak (April 24). *PsyArXiv*. https://doi.org/10.31234/osf.io/yb2h8

Folk, D. P., Okabe-Miyamoto, K., Dunn, E. W., & Lyubomirsky, S. (2020). Have introverts or extraverts declined in social connection during the first wave of COVID-19? (june 7). *PsyArXiv*. https://doi.org/10.31234/osf.io/tkr2b

Fox, J., Shin, Y., & Emamdjomeh, A. (2020, February 19). How epidemics like COVID-19 end (and how to end them faster). https://www.washingtonpost.com/graphics/2020/health/coronavirus-how-epidemics-spread-and-end/

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Galandra, C., Cerami, C., Santi, G., Dodich, A., Cappa, S., Vecchi, T., & Crespi, C. (2020). Covid-19 in mind: How job loss and health threatening events modulate risk-taking behaviours in real-life contexts (may 27). *PsyArXiv*. https://doi.org/10.31234/osf.io/5n942

Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., & Rinaldo, A. (2020). Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences*, *117*(19), 10484–10491. https://doi.org/10.1073/pnas.2004978117

Geldsetzer, P. (2020). Knowledge and perceptions of COVID-19 among the general public in the United States and the United Kingdom: A cross-sectional online survey. *Annals of Internal Medicine*, *173*(2), 157–160. https://doi.org/10.7326/M20-0912

Gelfand, M., Li, R., Stamkou, E., Pieper, D., Denison, E., Fernandez, J., Choi, V. K., Chatman, J., Jackson, J. C., & Dimant, E. (2021). Persuading republicans and democrats to comply with mask wearing: An intervention tournament. *PsyArXiv*. https://doi.org/10.31234/osf.io/6gjh8

Gerhold, L. (2020). COVID-19: Risk perception and coping strategies (March 25). *PsyArXiv*. https://doi.org/10.31234/osf.io/xmpk4

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.

Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., & Colaneri, M. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in italy. *Nature Medicine*, *26*, 855–860.

Globig, L. K., Blain, B., & Sharot, T. (2020). When private optimism meets public despair: Dissociable effects on behavior and well-being (may 29). *PsyArXiv*. https://doi.org/10.31234/osf.io/gbdn8

Glöckner, A., Dorrough, A. R., Wingen, T., & Dohle, S. (2020). The perception of infection risks during the early and later outbreak of COVID-19 in Germany: Consequences and recommendations (May 14). https://doi.org/10.31234/osf.io/wdbgc

Goldberg, M. H., Gustafson, A., Maibach, E., Ballew, M. T., Bergquist, P., Kotcher, J., Marlon, J. R., Rosenthal, S. A., & Leiserowitz, A. (2020). Mask-wearing increased after a government recommendation: A natural experiment in the U.S. during the COVID-19 pandemic (April 20). *PsyArXiv*. https://doi.org/10.31234/osf.io/uc8nz

Götz, F. M., Gvirtz, A., Galinsky, A., & Jachimowicz, J. (2020). How personality and policy predict pandemic behavior: Understanding sheltering-in-place in 55 countries at the onset of COVID-19 (September 10). *PsyArXiv*. https://doi.org/10.31234/osf.io/c7sj2

Gouin, J.-P. (2020). Social, cognitive, and emotional predictors of adherence to physical distancing during the COVID-19 pandemic (May 13). *PsyArXiv*. https://doi.org/10.31234/osf.io/ksj52

Graham, A., Cullen, F. T., Pickett, J. T., Jonson, C. L., Haner, M., & Sloan, M. M. (2020). Faith in trump, moral foundations, and social distancing defiance during the coronavirus pandemic. *Socius*, *6*, 2378023120956815.

Grant, A. M., & Hofmann, D. A. (2011). It's not all about me: Motivating hand hygiene among health care professionals by focusing on patients. *Psychological Science*, *22*(12), 1494–1499. https://doi.org/10.1177/0956797611419172

Hahn, U., Lagnado, D., Lewandowsky, S., & Chater, N. (2020). Crisis knowledge management: Reconfiguring the behavioural science community for rapid responding in the COVID-19 crisis. *PsyArXiv*. https://doi.org/10.31234/osf.io/hsxdk

Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. Random House.

Hamburger, H. (1973). N-person prisoner's dilemma. *Journal of Mathematical Sociology*, *3*(1), 27–48. https://doi.org/10.1080/0022250X.1973.9989822

Harper, C. A., Satchell, L. P., Fido, D., & Latzman, R. D. (2020). Functional fear predicts public health compliance in the COVID-19 pandemic. *International Journal of Mental Health and Addiction*, 1–14. https://doi.org/10.1007/s11469-020-00281-5

Haushofer, J., & Metcalf, C. J. E. (2020). Which interventions work best in a pandemic? *Science*, *368*(6495), 1063–1065. https://doi.org/10.1126/science.abb6144

Heffner, J., Vives, M. L., & FeldmanHall, O. (2020). Emotional responses to prosocial messages increase willingness to self-isolate during the COVID-19 pandemic (April 15). https://doi.org/10.31234/osf.io/qkxvb

Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences*, *33*(2-3), 111–135. https://doi.org/10.1017/S0140525X10000725

Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). Most people are not WEIRD. *Nature*, *466*, 29. https://doi.org/10.1038/466029a

Hertwig, R. (2017). When to consider boosting: Some rules for policy-makers. *Behavioural Public Policy*, *1*(2), 143–161. https://doi.org/10.1017/bpp.2016.14

Hertwig, R., & Ortmann, A. (2008). Deception in experiments: Revisiting the arguments in its defense. *Ethics & Behavior*, *18*(1), 59–92. https://doi.org/10.1080/10508420701712990

Hertwig, R., Pleskac, T. J., Pachur, T., & the Center for Adaptive Rationality. (2019). *Taming uncertainty*. Mit Press.

Hertwig, R., & Ryall, M. D. (2020). Nudge versus boost: Agency dynamics under libertarian paternalism. *The Economic Journal*, *130*(629), 1384–1415. https://doi.org/10.1093/ej/uez054

Holmdahl, I., & Buckee, C. (2020). Wrong but useful—What COVID-19 epidemiologic models can and cannot tell us. *New England Journal of Medicine*, *383*(4), 303–305. https://doi.org/10.1056/NEJMp2016822

Hong, S. M., & Faedda, S. (1996). Refinement of the Hong psychological reactance scale. *Educational and Psychological Measurement*, *56*(1), 173–182. https://doi.org/10.1177/0013164496056001014

Hongbo, L., Hania, A., & Waqas, M. A. (2020). Psychological predictors of anxiety in response to the COVID-19 pandemic: Evidence from Pakistan (May 21). *PsyArXiv*. https://doi.org/10.31234/osf.io/gk6mt

Howard, J., Huang, A., Li, Z., Tufekci, Z., Zdimal, V., van der Westhuizen, H.-M., von Delft, A., Price, A., Fridman, L., Tang, L.-H., Tang, V., Watson, G. L., Bax, C. E., Shaikh, R., Questier, F., Hernandez, D., Chu, L. F., Ramirez, C. M., & Rimoin, A. W. (2021). An evidence review of face masks against COVID-19. *Proc. Natl. Acad. Sci. U.S.A.*, *118*(4), e2014564118. https://doi.org/10.1073/pnas.2014564118

Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., Druckenmiller, H., Huang, L. Y., Hultgren, A., Krasovich, E., et al. (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, *584*(7820), 262–267. https://doi.org/10.1038/s41586-020-2404-8

Huh, D., Mun, E.-Y., Larimer, M. E., White, H. R., Ray, A. E., Rhew, I. C., Kim, S.-Y., Jiao, Y., & Atkins, D. C. (2015). Brief motivational interventions for college student drinking may not be as powerful as we think: An individual participant-level data meta-analysis. *Alcoholism: Clinical and Experimental Research*, *39*(5), 919–931. https://doi.org/10.1111/acer.12714

Hyland, P., Shevlin, M., McBride, O., Murphy, J., Karatzias, T., Bentall, R. P., Martinez, A. P., & Vallières, F. (2020). Anxiety and depression in the Republic of Ireland during the COVID-19 pandemic (April 22). *PsyArXiv*. https://doi.org/10.31234/osf.io/8yqxr

Ibuka, Y., Li, M., Vietri, J., Chapman, G. B., & Galvani, A. P. (2014). Free-riding behavior in vaccination decisions: An experimental study. *PloS one*, *9*(1), e87164. https://doi.org/10.1371/journal.pone.0087164

Imhoff, R., & Lamberty, P. (2020). A bioweapon or a hoax? The link between distinct conspiracy beliefs about the coronavirus disease (COVID-19) outbreak and pandemic behavior (April 14). *PsyArXiv*. https://doi.org/10.31234/osf.io/ye3ma

Jarynowski, A., Wojta-Kempa, M., & Belik, V. (2020). Trends in perception of COVID-19 in Polish internet. *medRxiv*. https://doi.org/10.1101/2020.05.04.20090993

Jordan, J., Yoeli, E., & Rand, D. G. (2020). Don't get it or don't spread it? comparing self-interested versus prosocially framed COVID-19 prevention messaging (April 3). *PsyArXiv*. https://doi.org/10.31234/osf.io/yuq7x

Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, *125*(2), 131–164. https://doi.org/10.1037/rev0000093

Kahn, D. T., Björklund, F., & Hirschberger, G. (2020). The intent and extent of collective threats: A data-driven conceptualization of collective threats and their relation to political preferences (March 17). *PsyArXiv*. https://doi.org/10.31234/osf.io/evxu6

Kanngiesser, P., Sunderarajan, J., & Woike, J. K. (2020). Keeping them honest: Promises reduce cheating in adolescents. *Journal of Behavioral Decision Making*.

Kaufmann, C., Weber, M., & Haisley, E. (2013). The role of experience sampling and graphical displays on one's investment risk appetite. *Management Science*, *59*(2), 323–340. https://doi.org/10.1287/mnsc.1120.1607

Kavanagh, N. M., Goel, R. R., & Venkataramani, A. S. (2020). Association of county-level socioeconomic and political characteristics with engagement in social distancing for COVID-19. *medRxiv*. https://doi.org/10.1101/2020.04.06.20055632

Kazi, D. S., Wadhera, R. K., Shen, C., Ho, K. K., Patell, R., Selim, M. H., Urwin, J. H., Zeidel, M. L., Zimetbaum, P., Tabb, K., & Yeh, R. W. (2020). Decline in emergent and urgent care during the COVID-19 pandemic. *medRxiv*. https://doi.org/10.1101/2020.05.14.20096602

Kenyon, C. (2020). Widespread use of face masks in public may slow the spread of SARS CoV-2: An ecological study. *medRxiv*. https://doi.org/10.1101/2020.03.31.20048652

Kitamura, S., & Yamada, K. (2020). Social comparisons and cooperation during COVID-19 (June 20). *PsyArXiv*. https://doi.org/10.31234/osf.io/rsbmz

Korn, L., Böhm, R., Meier, N. W., & Betsch, C. (2020). Vaccination as a social contract. *Proceedings of the National Academy of Sciences*, *117*(26), 14890–14899. https://doi.org/10.1073/pnas.1919666117

Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., Du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, *368*(6490), 493–497. https://doi.org/10.1126/science.abb4218

Kristof, N., & Thompson, S. A. (2020). Trump wants to 'reopen America.' here's what happens if we do. *New York Times*. https://www.nytimes.com/interactive/2020/03/25/opinion/coronavirus-trump-reopen-america.html

Kroencke, L., Geukes, K., Utesch, T., Kuper, N., & Back, M. (2020). Neuroticism and emotional risk during the Covid-19 pandemic (May 10). *PsyArXiv*. https://doi.org/10.31234/osf.io/8c6nh

Kuiper, M. E., de Bruijn, A. L., Folmer, C. R., Olthuis, E., Brownlee, M., Kooistra, E. B., Fine, A., & van Rooij, B. (2020). The intelligent lockdown: Compliance with COVID-19 mitigation measures in the netherlands. *SSRN*. https://doi.org/10.31234/osf.io/5wdb3

Kuper-Smith, B. J., Doppelhofer, L. M., Oganian, Y., Rosenblau, G., & Korn, C. (2020). Optimistic beliefs about the personal impact of COVID-19 (March 19). *PsyArXiv*. https://doi.org/10.31234/osf.io/epcyb

Lammers, J., Crusius, J., & Gast, A. (2020). Correcting misperceptions of exponential coronavirus growth increases support for social distancing. *Proceedings of the National Academy of Sciences*, *117*(28), 16264–16266. https://doi.org/10.1073/pnas.2006048117

Lan, F.-Y., Wei, C.-F., Hsu, Y.-T., Christiani, D. C., & Kales, S. N. (2020). Work-related COVID-19 transmission. *medRxiv*. https://doi.org/10.1101/2020.04.08.20058297

Lavezzo, E., Franchin, E., Ciavarella, C., Cuomo-Dannenburg, G., Barzon, L., Del Vecchio, C., Rossi, L., Manganelli, R., Loregian, A., Navarin, N., et al. (2020). Suppression of a SARS-CoV-2 outbreak in the italian municipality of Vo'. *Nature*, *584*, 425–429. https://doi.org/10.1038/s41586-020-2488-1

Leary, A. V., & Dvorak, R. D. (2020). Social distancing during the COVID-19 pandemic: Normative interventions and correlates of social distancing (May 13). *PsyArXiv*. https://doi.org/10.31234/osf.io/mszw2

Leder, J., Pastukhov, A., & Schütz, A. (2020). Even prosocially oriented individuals save themselves first: Social value orientation, subjective effectiveness and the usage of protective measures during the COVID-19 pandemic in Germany (March 30). *PsyArXiv*. https://doi.org/10.31234/osf.io/nugcr

Lees, J. M., Cetron, J. S., Vollberg, M. C., Reggev, N., & Cikara, M. (2020). Intentions to comply with COVID-19 preventive behaviors are associated with personal beliefs, independent of perceived social norms (May 20). *PsyArXiv*. https://doi.org/10.31234/osf.io/97jry

Leuker, C., Pachur, T., Hertwig, R., & Pleskac, T. J. (2018). Exploiting risk–reward structures in decision making under uncertainty. *Cognition*, *175*, 186–200.

Leung, K., Wu, J. T., Di Liu, & Leung, G. M. (2020). First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: A modelling impact assessment. *The Lancet*, *395*(10233), 1382–1393. https://doi.org/10.1016/S0140-6736(20)30746-7

Levy, H. (1994). Absolute and relative risk aversion: An experimental study. *Journal of Risk and Uncertainty*, *8*(3), 289–307. https://doi.org/10.1007/BF01064046

Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA faked the moon landing—therefore, (climate) science is a hoax: An anatomy of the motivated rejection of science. *Psychological Science*, *24*(5), 622–633. https://doi.org/10.1177/0956797612457686

Lewandowsky, S., Gignac, G. E., & Oberauer, K. (2015). Correction: The role of conspiracist ideation and worldviews in predicting rejection of science. *PLOS ONE*, *10*(8), e0134773. https://doi.org/10.1371/journal.pone.0134773

Lewandowsky, S., Gignac, G. E., & Vaughan, S. (2013). The pivotal role of perceived scientific consensus in acceptance of science. *Nature Climate Change*, *3*(4), 399–404. https://doi.org/10.1038/nclimate1720

Lewandowsky, S., Woike, J. K., & Oberauer, K. (2020). Genesis or evolution of gender differences? worldview-based dilemmas in the processing of scientific information. *Journal of Cognition*, *3*(1), 9. https://doi.org/10.5334/joc.99

Lewis, M. A., & Neighbors, C. (2006a). Social norms approaches using descriptive drinking norms education: A review of the research on personalized normative feedback. *Journal of American College Health*, *54*(4), 213–218. https://doi.org/10.3200/JACH.54.4.213-218

Lewis, M. A., & Neighbors, C. (2006b). Who is the typical college student? implications for personalized normative feedback interventions. *Addictive behaviors*, *31*(11), 2120–2126. https://doi.org/10.1016/j.addbeh.2006.01.011

Liang, J., Yuan, H.-Y., Wu, L., & Pfeiffer, D. U. (2020). Estimating effects of intervention measures on COVID-19 outbreak in wuhan taking account of improving diagnostic capabilities using a modelling approach. *medRxiv*. https://doi.org/10.1101/2020.03.31.20049387

Liebrand, W. B., Jansen, R. W., Rijken, V. M., & Suhre, C. J. (1986). Might over morality: Social values and the perception of other players in experimental games. *Journal of Experimental Social Psychology*, *22*(3), 203–215. https://doi.org/10.1016/0022-1031(86)90024-7

Lim, W., & Zhang, P. (2020). Herd immunity and a vaccination game: An experimental study. *PLOS ONE*, *15*(5), e0232652. https://doi.org/10.1371/journal.pone.0232652

Litvinova, M., Liu, Q.-H., Kulikov, E. S., & Ajelli, M. (2019). Reactive school closure weakens the network of social interactions and reduces the spread of influenza. *Proceedings of the National Academy of Sciences*, *116*(27), 13174–13181. https://doi.org/10.1073/pnas.1821298116

Lofgren, E. T., & Fefferman, N. H. (2007). The untapped potential of virtual game worlds to shed light on real world epidemics. *The Lancet Infectious Diseases*, *7*(9), 625–629. https://doi.org/10.1016/S1473-3099(07)70212-8

López, L., & Rodó, X. (2020). The end of social confinement and COVID-19 re-emergence risk. *Nature Human Behaviour*, *4*(7), 746–755. https://doi.org/10.1038/s41562-020-0908-8

Lunn, P. D., Timmons, S., Julienne, H., Belton, C., Barjaková, M., Lavin, C., & McGowan, F. P. (2020). Using decision aids to support self-isolation during the COVID-19 pandemic (May 29). *PsyArXiv*. https://doi.org/10.31234/osf.io/fngx5

Luttrell, A., Petty, R. E., Briñol, P., & Wagner, B. C. (2016). Making it moral: Merely labeling an attitude as moral increases its strength. *Journal of Experimental Social Psychology*, *65*, 82–93. https://doi.org/10.1016/j.jesp.2016.04.003

Maier, B. F., & Brockmann, D. (2020). Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*, *368*(6492), 742–746. https://doi.org/10.1126/science.abb4557

Malik, A. A., Couzens, C., & Omer, S. B. (2020). COVID-19 related social distancing measures and reduction in city mobility. *medRxiv*. https://doi.org/10.1101/2020.03.30.20048090

Manjoo, F. (2020, September 30). Opinion | Trump Can Still Make a Difference on Masks. https://www.nytimes.com/2020/09/30/opinion/sunday/coronavirus-masks.html

Mao, A., Parkes, D. C., Procaccia, A. D., & Zhang, H. (2011). Human computation and multiagent systems: An algorithmic perspective. *In Proc. AAAI 11*.

March, J. G., & Olsen, J. P. (2013). The logic of appropriateness. In R. E. Goodin (Ed.), *The oxford handbook of political science*. https://doi.org/10.1093/oxfordhb/9780199604456.013.0024

Marinthe, G., Brown, G., Delouvée, S., & Jolley, D. (2020). Looking out for myself: Exploring the relationship between conspiracy mentality, perceived personal risk and COVID-19 prevention measures (May 29). *PsyArXiv*. https://doi.org/10.31234/osf.io/cm9st

Martin, C. C. (2020). HEXACO traits, big five traits, and COVID-19 (May 20)). *PsyArXiv*. https://doi.org/10.31234/osf.io/c9gxe

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavioral Research Methods*, *44*(1), 1–23. https://doi.org/10.3758/s13428-011-0124-6

McElroy, E., Patalay, P., Moltrecht, B., Shevlin, M., Shum, A., Creswell, C., & Waite, P. (2020). Demographic and health factors associated with pandemic anxiety in the context of COVID-19 (May 8). *PsyArXiv*. https://doi.org/10.31234/osf.io/2eksd

McNaughton, C. D., Cavanaugh, K. L., Kripalani, S., Rothman, R. L., & Wallston, K. A. (2015). Validation of a short, 3-item version of the Subjective Numeracy Scale. *Medical Decision Making*, *35*(8), 932–936. https://doi.org/10.1177/0272989X15581800

Metcalf, C. J. E., Morris, D. H., & Park, S. W. (2020). Mathematical models to guide pandemic response. *Science*, *369*(6502), 368–369. https://doi.org/10.1126/science.abd1668

Michie, S., West, R., Rogers, M. B., Bonell, C., Rubin, G. J., & Amlôt, R. (2020). Reducing SARS-CoV-2 transmission in the uk: A behavioural science approach to identifying options for increasing adherence to social distancing and shielding vulnerable people. *British Journal of Health Psychology*. https://doi.org/10.1111/bjhp.12428

Mittal, C., Griskevicius, V., & Haws, K. L. (2020). From cradle to grave: How childhood and current environments impact consumers' subjective life expectancy and decision-making. *Journal of Consumer Research*. https://doi.org/10.1093/jcr/ucaa003

Molnar, A. (2019). SMARTRIQS: A simple method allowing real-time respondent interaction in Qualtrics surveys. *Journal of Behavioral and Experimental Finance*, *22*, 161–169.

Moore, R. C., Lee, A., Hancock, J. T., Halley, M., & Linos, E. (2020). Experience with social distancing early in the COVID-19 pandemic in the united states: Implications for public health messaging. *medRxiv*. https://doi.org/10.1101/2020.04.08.20057067

Moshagen, M., Zettler, I., & Hilbig, B. E. (2019). Measuring the dark core of personality. *Psychological Assessment*, *32*(2), 182–196. https://doi.org/10.1037/pas0000778

Mukherji, N. (2020). The social and economic factors underlying the incidence of COVID-19 cases and deaths in US counties. *medRxiv*. https://doi.org/10.1101/2020.05.04.20091041

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision making*, *6*(8), 771–781.

Muto, K., Yamamoto, I., Nagasu, M., Tanaka, M., & Wada, K. (2020). Japanese citizens behavioral changes and preparedness against COVID-19: How effective is Japan's approach of self-restraint? *medRxiv*. https://doi.org/10.1101/2020.03.31.20048876

Nacoti, M., Ciocca, A., Giupponi, A., Brambillasca, P., Lussana, F., Pisano, M., Goisis, G., Bonacina, D., Fazzi, F., Naspro, R., et al. (2020). At the epicenter of the COVID-19 pandemic and humanitarian crises in Italy: Changing perspectives on preparation and mitigation. *NEJM Catalyst Innovations in Care Delivery*, *1*(2). https://doi.org/10.1056/CAT.20.0080

Nozick, R. (1993). *The nature of rationality*. Princeton University Press.

Nyhan, B. (2014, October 16). The Partisan Divide on Ebola Preparedness. https://www.nytimes.com/2014/10/17/upshot/the-partisan-divide-on-ebola-preparedness.html

Okruszek, L., Aniszewska-Staczuk, A., Piejka, A., Winiewska, M., & urek, K. (2020). Safe but lonely? loneliness, mental health symptoms and COVID-19 (April 10). *PsyArXiv*. https://doi.org/10.31234/osf.io/9njps

Oosterhoff, B., & Palmer, C. (2020). Psychological correlates of news monitoring, social distancing, disinfecting, and hoarding behaviors among US adolescents during the COVID-19 pandemic (March 23). *PsyArXiv*. https://doi.org/10.31234/osf.io/rpcy4

Ortmann, A., & Hertwig, R. (2002). The costs of deception: Evidence from psychology. *Experimental Economics*, *5*(2), 111–131. https://doi.org/10.1023/A:1020365204768

Pan, A., Liu, L., Wang, C., Guo, H., Hao, X., Wang, Q., Huang, J., He, N., Yu, H., Lin, X., et al. (2020). Association of public health interventions with the epidemiology

of the COVID-19 outbreak in Wuhan, China. *JAMA*, *323*(19), 1915–1923. https://doi.org/10.1001/jama.2020.6130

Peeples, L. (2021, September 9). Face masks for COVID pass their largest test yet. https://doi.org/10.1038/d41586-021-02457-y

Pennycook, G., McPhetres, J., Bago, B., & Rand, D. G. (2020a). Beliefs about COVID-19 in Canada, the U.K., and the U.S.A.: A novel test of political polarization and motivated reasoning (February 1). *PsyArXiv*. https://doi.org/10.31234/osf.io/zhjkp

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy nudge intervention (March 17). *PsyArXiv*. https://doi.org/10.31234/osf.io/uhbk9

Pennycook, G., McPhetres, J., Bago, B., & Rand, D. G. (2020b). Predictors of attitudes and misperceptions about COVID-19 in canada, the uk, and the usa (April 14). *PsyArXiv*. https://doi.org/10.31234/osf.io/zhjkp

Perrotta, D., Grow, A., Rampazzo, F., Cimentada, J., Del Fava, E., Gil-Clavel, S., & Zagheni, E. (2020). Behaviors and attitudes in response to the COVID-19 pandemic: Insights from a cross-national Facebook survey. *medRxiv*. https://doi.org/10.1101/2020.05.09.20096388

Peyton, K., Huber, G. A., & Coppock, A. (2021). The generalizability of online experiments conducted during the COVID-19 pandemic. *Journal of Experimental Political Science*, 1–16. https://doi.org/10.1017/XPS.2021.17

Pfattheicher, S., Nockur, L., Böhm, R., Sassenrath, C., & Petersen, M. B. (2020). The emotional path to action: Empathy promotes physical distancing during the COVID-19 pandemic (March 23). https://doi.org/10.31234/osf.io/y2cg5

Pink, S. L., Stagnaro, M., Chu, J., Mernyk, J., Voelkel, J. G., & Willer, R. (2020). Short messages encouraging compliance with COVID-19 public health guidelines have minimal persuasive effects (August 10). *PsyArXiv*. https://doi.org/10.31234/osf.io/g93zw

Plohl, N., & Musil, B. (2021). Modeling compliance with COVID-19 prevention guidelines: The critical role of trust in science. *Psychology, Health & Medicine*, *26*(1), 1–12. https://doi.org/10.1080/13548506.2020.1772988

Plonsky, O., Roth, Y., & Erev, I. (2021). Underweighting of rare events in social interactions and its implications to the design of voluntary health applications. *Judgment and Decision Making*, *16*(2), 267. https://sjdm.org/journal/20/201217/jdm201217.pdf

Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N., Flasche, S., Clifford, S., Pearson, C. A., Munday, J. D., et al. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *The Lancet Public Health*, *5*, e261–e270. https://doi.org/10.1016/S2468-2667(20)30073-6

Rader, B., White, L. F., Burns, M. R., Chen, J., Brilliant, J., Cohen, J., Shaman, J., Brilliant, L., Kraemer, M. U. G., Hawkins, J. B., Scarpino, S. V., Astley, C. M., & Brownstein, J. S. (2021). Mask-wearing and control of SARS-CoV-2 transmission in the USA: a cross-sectional study. *Lancet Digital Health*, *3*(3), e148–e157. https://doi.org/10.1016/S2589-7500(20)30293-4

Raihani, N., & de-Wit, L. (2020). Factors associated with concern, behaviour & policy support in response to SARS-CoV-2 (April 21). *PsyArXiv*. https://doi.org/10.31234/osf.io/8jpzc

Ratcliff, R., & Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. *Psychological Science*, *9*(5), 347–356. https://doi.org/10.1111/1467-9280.00067

Raude, J. (2020). Determinants of preventive behaviors in response to the COVID-19 pandemic in france: Comparing the sociocultural, psychosocial and social cognitive explanations (May 31). *PsyArXiv*. https://doi.org/10.31234/osf.io/4yvk2

Rentsch, C. T., Kidwai-Khan, F., Tate, J. P., Park, L. S., King, J. T., Skanderson, M., Hauser, R. G., Schultze, A., Jarvis, C. I., Holodniy, M., Lo Re, V., Akgun, K. M., Crothers, K., Taddei, T. H., Freiberg, M. S., & Justice, A. C. (2020). Covid-19 by race and ethnicity: A national cohort study of 6 million United States Veterans. *medRxiv*. https://doi.org/10.1101/2020.05.12.20099135

Richter, I., Thøgersen, J., & Klöckner, C. A. (2018). A social norms intervention going wrong: Boomerang effects from descriptive norms information. *Sustainability*, *10*(8). https://doi.org/10.3390/su10082848

Robert, C. P., Casella, G., & Casella, G. (2004). *Monte carlo statistical methods* (Vol. 2). Springer.

Roberts, S. (2020a, March 13). The exponential power of now. https://www.nytimes.com/2020/03/13/science/coronavirus-math-mitigation-distancing.html

Roberts, S. (2020b, March 19). You can help break the chain of transmission. https://www.nytimes.com/2020/03/19/health/coronavirus-distancing-transmission.html

Romano, A., Sotis, C., Dominioni, G., & Guidi, S. (2020). Covid-19 data: The logarithmic scale misinforms the public and affects policy preferences. *Health Economics*. https://doi.org/10.1002/hec.4143

Rosenfeld, D. L. (2020). Political ideology and the outbreak of COVID-19 in the United States (April 22). *PsyArXiv*. https://doi.org/10.31234/osf.io/jrpfd

Rosenfeld, D. L., Balcetis, E., Bastian, B., Berkman, E. T., Bosson, J. K., Brannon, T. N., Burrow, A. L., Cameron, C. D., Chen, S., Cook, J. E., et al. (2021). Psychological science in the wake of COVID-19: Social, methodological, and meta-scientific considerations. *Perspectives on Psychological Science*.

Roth, Y., Plonsky, O., Shalev, E., & Erev, I. (2020). On the value of alert systems and gentle rule enforcement in addressing pandemics. *Frontiers in Psychology*, *11*, 577743. https://doi.org/10.3389/fpsyg.2020.577743

Rothgerber, H., Wilson, T., Whaley, D., Rosenfeld, D. L., Humphrey, M., Moore, A. L., & Bihl, A. (2020). Politicizing the COVID-19 pandemic: Ideological differences in adherence to social distancing (april 22). *PsyArXiv*. https://doi.org/10.31234/osf.io/k23cv

Rudisill, C. (2013). How do we handle new health risks? Risk perception, optimism, and behaviors regarding the H1N1 virus. *Journal of Risk Research*, *16*(8), 959–980. https://doi.org/10.1080/13669877.2012.761271

Saha, A. (1993). Expo-power utility: A flexible form for absolute and relative risk aversion. *American Journal of Agricultural Economics*, *75*(4), 905–913. https://doi.org/10.2307/1243978

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2018). The constructive, destructive, and reconstructive power of social norms: Reprise. *Perspectives on Psychological Science*, *13*(2), 249–254. https://doi.org/10.1177/1745691617693325

Schulze, C., & Pachur, T. (2019). Going round in circles: How social structures guide and limit search. In R. Hertwig, T. J. Pleskac, T. Pachur, & the Center for Adaptive Rationality (Eds.). MIT Press.

Seale, H., Heywood, A. E., Leask, J., Steel, M., Thomas, S., Durrheim, D. N., Bolsewicz, K., & Kaur, R. (2020). COVID-19 is rapidly changing: Examining public perceptions and behaviors in response to this evolving pandemic. *medRxiv*. https://doi.org/10.1101/2020.05.04.20091298

Shea, K., Runge, M. C., Pannell, D., Probert, W. J., Li, S.-L., Tildesley, M., & Ferrari, M. (2020). Harnessing multiple models for outbreak management. *Science*, *368*(6491), 577–579. https://doi.org/10.1126/science.aba3763

Siedner, M. J., Harling, G., Reynolds, Z., Gilbert, R. F., Haneuse, S., Venkataramani, A., & Tsai, A. C. (2020). Social distancing to slow the U.S. COVID-19 epidemic: Longitudinal pretest-posttest comparison group study. *medRxiv*. https://doi.org/10.1101/2020.04.03.20052373

Simione, L., & Gnagnarella, C. (2020). Differences between health workers and general population in risk perception, behaviors, and psychological distress related to COVID-19 spread in Italy (April 3). *PsyArXiv*. https://doi.org/10.31234/osf.io/84d2c

Sparkman, G., & Walton, G. M. (2017). Dynamic norms promote sustainable behavior, even if it is counternormative. *Psychological Science*, *28*(11), 1663–1674. https://doi.org/10.1177/09567976177199500

Stanley, M., Barr, N., Peters, K., & Seli, P. (2020). Analytic-thinking predicts hoax beliefs and helping behaviors in response to the COVID-19 pandemic (March 30). https://doi.org/10.31234/osf.io/m3vth

Sternisko, A., Cichocka, A., Cislak, A., & Van Bavel, J. J. (2020). Collective narcissism predicts the belief and dissemination of conspiracy theories during the COVID-19 pandemic (May 21). *PsyArXiv*. https://doi.org/10.31234/osf.io/4c6av

Stevens, H. (2020a, August 7). A vaccine, or a spike in deaths: How america can build herd immunity to the coronavirus. https://www.washingtonpost.com/graphics/2020/health/coronavirus-herd-immunity-simulation-vaccine/

Stevens, H. (2020b, March 14). Why outbreaks like coronavirus spread exponentially, and how to "flatten the curve". https://www.washingtonpost.com/graphics/2020/world/corona-simulator/

Stevens, H., & Muyskens, J. (2020, May 14). Disease modelers are wary of reopening the country. here's how they arrive at their verdict. https://www.washingtonpost.com/graphics/2020/health/disease-modeling-coronavirus-cases-reopening/

Summerville, A., & Chartier, C. R. (2013). Pseudo-dyadic interaction on Amazon's Mechanical Turk. *Behavioral Research Methods*, *45*(1), 116–124. https://doi.org/10.3758/s13428-012-0250-9

Suri, S., & Watts, D. J. (2011). Cooperation and Contagion in Web-Based, Networked Public Goods Experiments. *PLoS One*, *6*(3), e16836. https://doi.org/10.1371/journal.pone.0016836

Teovanovic, P., Lukic, P., Zupan, Z., Lazi, A., Ninkovi, M., & Zezelj, I. (2020). Irrational beliefs differentially predict adherence to guidelines and pseudoscientific practices during the COVID-19 pandemic (May 20). *PsyArXiv*. https://doi.org/10.31234/osf.io/gefhn

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision making*, *11*(1), 99–113.

Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U., Li, B., Cai, J., Xu, B., Yang, Q., et al. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, *368*(6491), 638–642. https://doi.org/10.1126/science.abb6105

Trueblood, J., Sussman, A., O'Leary, D., & Holmes, W. (2020). A tale of two crises: Financial fragility and beliefs about the spread of COVID-19 (April 21). *PsyArXiv*. https://doi.org/10.31234/osf.io/xfrz3

Tunçgenç, B., El Zein, M., Sulik, J., Newson, M., Zhao, Y., Dezecache, G., & Deroy, O. (2020). We distance most when we believe our social circle does (May 8). *PsyArXiv*. https://doi.org/10.31234/osf.io/u74wc

Utych, S. M., & Fowler, L. (2020). Age-based messaging strategies for communication about COVID-19. *Journal of Behavioral Public Administration*, *3*(1). https://doi.org/10.30636/jbpa.31.151

Van Bavel, J. (2020, March 22). In a pandemic, political polarization could kill people. https://www.washingtonpost.com/outlook/2020/03/23/coronavirus-polarization-political-exaggeration/

Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., Drury, J., Dube, O., Ellemers, N., Finkel, E. J., Fowler, J. H., Gelfand, M., Han, S., Haslam, S. A., Jetten, J., ... Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, *4*(5), 460–471. https://doi.org/10.1038/s41562-020-0884-z

van der Linden, S., Leiserowitz, A., & Maibach, E. (2018). Scientific agreement can neutralize politicization of facts. *Nature Human Behaviour*, *2*(1), 2–3. https://doi.org/10.1038/s41562-017-0259-2

Velásquez, N., Leahy, R., Johnson Restrepo, N., Lupu, Y., Sear, R., Gabriel, N., Jha, O., & Johnson, N. (2020). Hate multiverse spreads malicious COVID-19 content online beyond individual platform control. *arXiv*, 2004.00673 [physics.soc–ph]. https://arxiv.org/abs/2004.00673

Vischer, T., Dohmen, T., Falk, A., Huffman, D., Schupp, J., Sunde, U., & Wagner, G. G. (2013). Validating an ultra-short survey measure of patience. *Economics Letters*, *120*(2), 142–145. https://doi.org/10.1016/j.econlet.2013.04.007

Walker, P. G., Whittaker, C., Watson, O. J., Baguelin, M., Winskill, P., Hamlet, A., Djafaara, B. A., Cucunubá, Z., Mesa, D. O., Green, W., et al. (2020). The impact of COVID-19 and strategies for mitigation and suppression in low-and middle-income countries. *Science*, *369*, 413–422. https://doi.org/10.1126/science.abc0035

West, R., Michie, S., Rubin, G. J., & Amlôt, R. (2020). Applying principles of behaviour change to reduce SARS-CoV-2 transmission. *2397-3374*, *4*(5), 451–459. https://doi.org/10.1038/s41562-020-0887-9

Westgard, B. C., Morgan, M. W., Vazquez-Benitez, G., Erickson, L. O., & Zwank, M. D. (2020). Characteristics of patients presenting, and not presenting, to the emergency department during the early days of COVID-19. *medRxiv*. https://doi.org/10.1101/2020.05.05.20090795

Whitt, Z. T., Bernstein, M., Spillane, N., Stein, L., Suffoletto, B., Neighbors, C., Schick, M. R., & Cyders, M. A. (2019). Positive urgency worsens the impact of normative feedback on 21st birthday drinking. *Drug and Alcohol Dependence*, *204*, 107559. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2019.107559

Wise, T., Zbozinek, T. D., Michelini, G., Hagan, C. C., & Mobbs, D. (2020). Changes in risk perception and protective behavior during the first week of the COVID-19 pandemic in the united states (March 19). *PsyArXiv*. https://doi.org/10.31234/osf.io/dz428

Wissmath, B., Mast, F. W., Kraus, F., & Weibel, D. (2020). Understanding the psychological impact of the COVID-19 pandemic and containment measures: An empirical model of stress. *medRxiv*. https://doi.org/10.1101/2020.05.13.20100313

Woike, J. K. (2019). Upon repeated reflection: Consequences of frequent exposure to the Cognitive Reflection Test for Mechanical Turk participants. *Frontiers in Psychology*, *10*, 2646. https://doi.org/10.3389/fpsyg.2019.02646

Woike, J. K., & Hafenbrädl, S. (2020). Rivals without a cause? Relative performance feedback creates destructive competition despite aligned incentives. *Journal of Behavioral Decision Making*, *33*(4), 523–537.

Wood, W. (2000). Attitude change: Persuasion and social influence. *Annual Review of Psychology*, *51*, 539–570. https://doi.org/10.1146/annurev.psych.51.1.539

Xu, P., & Cheng, J. (2021). Individual differences in social distancing and mask-wearing in the pandemic of COVID-19: The role of need for cognition, self-control and risk attitude. *Personality and Individual Differences*, *175*, 110706. https://doi.org/10.1016/j.paid.2021.110706

Yamada, Y., Cepuli, D.-B., Coll-Martín, T., Debove, S., Gautreau, G., Han, H., Rasmussen, J., Tran, T. P., Travaglino, G. A., & Lieberoth, A. (2020). COVIDiSTRESS Global Survey dataset on psychological and behavioural consequences of the COVID-19 outbreak (June 2). *PsyArXiv*. https://doi.org/10.31234/osf.io/v7cep

Yoeli, E., & Rand, D. G. (2020). A checklist for prosocial messaging campaigns such as COVID-19 prevention appeals (April 17). *PsyArXiv*. https://doi.org/10.31234/osf.io/rg2x9

Zettler, I., Schild, C., Lilleholt, L., & Böhm, R. (2020). Individual differences in accepting personal restrictions to fight the COVID-19 pandemic: Results from a Danish adult sample (March 23). *PsyArXiv*. https://doi.org/10.31234/osf.io/pkm2a

Zhang, J., Litvinova, M., Liang, Y., Wang, Y., Wang, W., Zhao, S., Wu, Q., Merler, S., Viboud, C., Vespignani, A., Ajelli, M., & Yu, H. (2020). Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*, *368*(6498), 1481–1486. https://doi.org/10.1126/science.abb8001

Zhang, Y., Johnson, K., Lich, K. H., Ivy, J., Keskinocak, P., Mayorga, M., & Swann, J. L. (2021). COVID-19 Projections for K12 Schools in Fall 2021: Significant Transmission

without Interventions. *medRxiv*, 2021.08.10.21261726. https://doi.org/10.1101/2021.08.10.21261726

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, *111*(4), 493–504. https://doi.org/10.1037/pspa0000056

Zizzo, D. J. (2003). Money burning and rank egalitarianism with random dictators. *Economics Letters*, *81*(2), 263–266. https://doi.org/10.1016/S0165-1765(03)00190-3