App-based Epidemic Game to Model Belief-Behavior Mapping and Cost Incentives in Voluntary Quarantine: A Randomized Controlled Trial

Andrés Colubri[1,2*], Andonaq Grozdani[2], Mansi Khandpekar[2], Yousif Graytee[3], Omar Al-Mohammedi[3], Ahmed Ayden Al-Shabandar[3], Wid Yasir Shabeeb[3], Yaqoob Ghassan[3], Hayder Swayedi[3], Chris T. Bauch[4], John Drury[5], Jasmina Panovska-Griffiths[6,7], Dmitri Williams[8], Dale King[3*]

[1] Department of Genomics and Computational Biology, UMass Chan Medical School, Worcester MA 01605, United States of America
[2] Broad Institute of Harvard and MIT, Cambridge MA 02142, United States of America
[3] American University of Iraq - Baghdad, Baghdad Governorate 77HC+FF, Iraq
[4] Department of Applied Mathematics, University of Waterloo, Waterloo ON N2L 3G1, Canada
[5] School of Psychology, University of Sussex, Brighton BN1 9QH, United Kingdom
[6] Pandemic Sciences Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7LF, United Kingdom
[7] Queen's College, University of Oxford, Oxford OX1 4AW, United Kingdom
[8] USC Annenberg School for Communication and Journalism, University of Southern California, Los Angeles CA 90007, United States of America
[*]=corresponding authors

Abstract

Background: Understanding the drivers of protective behavior during infectious disease outbreaks is critical for public health policy. App-based experimental epidemic games (epigames) offer a novel method to study these behaviors empirically, but their external validity—how well in-game choices reflect real-life beliefs—still needs to be rigorously tested.

Methods: We conducted a two-week randomized controlled trial (N=567) using the *Epigames* smartphone app at the American University of Iraq – Baghdad (AUIB) campus in the Middle East. This app used Bluetooth communication to sample the contact network between participants in sub-minute resolution and simulated the spread of a hypothetical respiratory virus through this network. Participants were randomized into two groups with differing opportunity costs (in-game points) for adopting voluntary quarantine within the game that would protect them from the simulated infection: Group 1 (Low Barrier) faced a small point difference between quarantine and non-quarantine choices, while Group 2 (High Barrier) faced a much larger difference. The optimal point difference between groups was determined by a Willing to Accept (WTA) pilot survey prior to the epigame. We measured real-life health beliefs and in-game beliefs through surveys at the beginning of the epigame, including questions on susceptibility, severity, self-efficacy, and benefits, to calculate the correlation between the two and to construct a Health Belief Model (HBM) parameterized by game data. We also measured self-assessment of game realism and behavior via an exit survey.

Results: Real-life and in-game beliefs showed moderate positive (Spearman's coefficient ρ from 0.13 to 0.36) and statistically significant (p-value < 0.05) correlation across all survey measures, suggesting indicator (behavior) parallelism between real-life and the epigame. The exit survey also yielded positive self-assessment of context and behavior realism during the epigame. While simple aggregate analysis showed no significant difference in quarantine rates between Groups 1 and 2, a Poisson regression model revealed a significant crossover interaction. High economic barriers significantly reduced quarantine adoption among participants with low in-game motivation (interaction coef. = -2.22, $p < 0.01$). Perceived benefits appeared to moderate this effect at a near significance level (coef. = +0.33, $p < 0.08$). Demographic factors such as gender

appeared to be significantly correlated with quarantine choice ($p < 0.01$). Analysis of the contact network measured with the app through Bluetooth showed assortative properties of several belief variables.

Conclusion: This is the first systematic pre-registered study on the external validity of app-based epigames and the impact of economic cost and individual beliefs on protective behaviors. We found that economic barriers act as a "gatekeeper" for quarantine during the game, suppressing action among the skeptical while allowing highly motivated individuals to act. The significant correlation between real-life, in-game beliefs, and network structure suggests that epigames are a valid experimental tool for network-aware behavioral epidemiology. This warrants further studies for replication, examining variance across settings, and addressing limitations (e.g., crosstalk between groups) and technical issues (e.g., Bluetooth connectivity) in the study design.

Keywords: Experimental game, epidemiology, epigame, external validity, indicator parallelism, health behavior, belief-behavior mapping, quarantine adoption, Health Belief Model

Study pre-registration: https://osf.io/qev6n/su

1. Introduction

Human behavior is a decisive factor in the dynamics of infectious disease transmission [1-6]. However, quantifying the causal impact of behavioral changes, such as voluntary quarantine, on pathogen spread remains a fundamental challenge in epidemiology. Observational data from real-world outbreaks are often confounded by simultaneous interventions and varying compliance, while mathematical models frequently rely on assumed rather than empirical parameters for human decision-making. True experiments are impossible for clear ethical reasons. Consequently, there is an urgent need for experimental frameworks that can isolate and quantify the drivers of protective behaviors under controlled yet realistic conditions.

Experimental epidemic games, or "epigames," conducted in naturalistic settings, are inspired by behavioral economics and game theory, and offer a promising avenue to fill this gap [7-9]. By placing participants in social dilemmas within a realistic framing where personal decisions impact both individual and collective outcomes, researchers can simulate the incentives inherent in disease prevention and mitigation [10-14]. While controlled behavioral experiments, often relying on web-based surveys and incentivized choice games [15-17] are designed to satisfy internal validity, they may lack external validity, specifically, by not being able to replicate the complex social contexts and realistic disease exposure patterns of daily life given their simplified framing and the fundamental issue of different real costs such as sickness and death. Conversely, early digital epidemiology projects like the "Corrupted Blood" incident in World of Warcraft [18] or the FluPhone project [19] demonstrated that highly immersive or sensor-based environments could generate realistic contact networks and rich behavioral data, and provide the basis for the epigames approach developed here. However, critical questions remains: *do the decisions players make in these gamified proxies "map" over to their real-life health beliefs, or are they merely artifacts of the game mechanics?* [20] *Can app-based gamified experiments reach a level of immersion and incentive that is verifiably parallel to their real-life counterparts?*

Despite the potential of the experimental games in general to provide insights into real-world behavior and cross-cultural differences often not captured by traditional ethnographic or survey methods [21], they have also been criticized for lacking external validity [22-26]. Anthropologists have responded to these critiques by testing framing effects [25, 27] and minimalism in experimental designs [23]. Although debates continue about the universal applicability of experimental game results, these instruments can provide meaningful insights into human responses under well-defined experimental conditions. Much like accepted disaster preparedness exercises, including the "Great ShakeOut" earthquake drills [28] and the Humanitarian Response Intensive Course [29], experimental games serve as performative "stress tests" that reveal behavioral patterns and systemic vulnerabilities within a safe, controlled environment.

It's important to distinguish two dimensions of external validity: generalizability, whether observed behaviors hold across populations and settings, and parallelism, the extent to which in-game mechanisms (context parallelism, or ecological validity) and behaviors (indicator parallelism) reflect real-world counterparts [30, 31]. Given the question of whether epigames can be used to study preventive health behaviors, we are interested in maximizing context and indicator parallelism as critical measures of external validity. More specifically, indicator parallelism is vital for establishing epigames as a rigorous scientific tool. If epigames are to serve as "laboratories" for public health policy, we must demonstrate that a participant's in-game choice to quarantine is driven by the same psychological factors (e.g., perceived susceptibility, severity, and efficacy) that drive real-world compliance. This validation is particularly crucial for understanding high-stakes interventions like quarantine, where individuals must weigh the collective public health benefit against immediate personal and economic costs.

To address these challenges, we developed the *Epigames* app [32] based on the *Operation Outbreak* (OO) app, which introduced the use of Bluetooth proximity sensing in experiential learning simulations of infectious disease spread [33]. Building on this foundation, the *Epigames* app allows us to conduct large-scale, randomized trial studies within a high-resolution, large-scale epidemic simulation. Unlike previous retrospective or exploratory analyses of contact networks and human behavior [34, 35], this study was explicitly designed to investigate the relationship between pre-existing real-life health beliefs and in-game protective choices, specifically examining how economic barriers (opportunity costs) moderate this relationship in the case of voluntary quarantine.

We designed the study as a randomized controlled trial (RCT). The app recorded real-world contact networks via Bluetooth to simulate the spread of a "digital" respiratory virus and represented health status via an animated avatar. However, "quarantine" was operationalized as a digital decision rather than physical isolation (which would have been difficult to implement): participants could choose to suspend their Bluetooth contacts via an in-app button, thereby preventing infection. To simulate real-world economic and social costs, choosing to quarantine incurred a point penalty compared to remaining active. We acknowledge that this digital abstraction is a much-simplified proxy for the complex burden of physical isolation that limits context parallelism of our study, but it allows for the precise experimental manipulation of the economic barrier. Participants were randomized into two arms: a Low Barrier Group (G1) facing a nominal penalty, and a High Barrier Group (G2) facing a significant penalty. Participants completed two surveys measuring health belief constructs, first regarding real-life respiratory diseases (S1), and subsequently regarding the game context (S2), to assess the external validity

of their in-game beliefs and behaviors. Consistent with our pre-registration on the Open Science Framework (OSF) website (https://osf.io/qev6n/) [36], we tested three hypotheses: (H1) Main Effect: G1 will exhibit higher quarantine rates than G2; (H2) Moderation by Cost: The correlation between personal beliefs and behavior will be stronger in G1, as high costs in G2 overpower personal motivations; and (H3) Interaction by Efficacy: The lower quarantine rates due to high barrier in G2 will be most pronounced among participants with low self-efficacy. In addition to these hypotheses, surveys S1 and S2 allowed us to investigate several research questions. (RQ1) Real-Life to Game mapping: is there a significant correlation between S1 and S2 variables? (RQ2) Demographics and Quarantine: are demographic variables collected in S1, such as gender, associated with the quarantine decision during the game? Finally, the measurement of the participants' contact network via Bluetooth enabled an additional research question. (RQ3) Network Structure: what topology best describes the contact network and whether the properties of its nodes exhibit assortative mixing.

We adopted the Health Belief Model (HBM) [37, 38] as an initial, although not definitive, theoretical framework for this investigation. A reason for doing so is HBM's distinction of being one of the most utilized models for predicting health behaviors [39], which makes it a well-known and understood model in the field of health behavioral research. HBM posits that an individual's decision to adopt a protective action such as voluntary quarantine is determined by the interplay of specific cognitive perceptions: perceived susceptibility to the illness, perceived severity of its consequences, perceived benefits of the action, and perceived barriers to performing it. In the context of infectious disease, the model suggests that compliance is a calculated trade-off where the threat of infection is weighed against the tangible costs of safety measures. In our study, we specifically operationalize "perceived barriers" through the game's economic opportunity costs, allowing us to experimentally test how high-cost barriers suppress protective behaviors even when perceived risk is held constant. By constructing a conceptual HBM using the data from the AUIB epigame, we aimed at demonstrating the use of the novel epigame approach to quantify how economic costs act as a barrier to protective behaviors [40, 41], providing empirical evidence to inform pandemic preparedness strategies.

2. Methods

2.1 Study Design and Participants

We recruited 567 participants from the American University of Iraq – Baghdad (AUIB) in the Middle East. A target sample size of N=500 was determined as a conveniency sample based on earlier pilots and total school enrollment (~2,500 students). Eligibility criteria included being a student, faculty, or staff member at AUIB, being at least 18 years of age, and possessing an iOS or Android smartphone capable of running the *Epigames* app. The study ran for 15 days, with six participants withdrawing during this period, leaving a final sample of 561. Among those who completed demographic survey S1, 314 identified as students and 23 as faculty or staff; 161 identified as male, 169 as female, and 4 as other. The study was designed as a two-arm, parallel-group, single-blind randomized controlled trial (RCT) with a 1:1 allocation ratio. Randomization was performed by the app using a coin-flip algorithm upon participant registration, ensuring that allocation had no inference from researchers and that it was concealed from participants until registration was completed. Refer to **Figure 1A** for the CONSORT Flow Diagram.

A month-long recruitment campaign comprising an information desk with physical fliers, an informational website and emails, class presentations, and a Telegram chat group, was conducted on campus by a local team of student volunteers (**Figure 1B**). Participants provided informed consent after installing the app to join the epigame and were incentivized via daily raffles for local store vouchers to maintain engagement. Points collected through the app were used in a lottery for gift cards, with the number of tickets proportional to the participants' final point score. The study protocol was reviewed and approved by UMass Chan IRB (ID 2465) and received clearance by AUIB IRB.

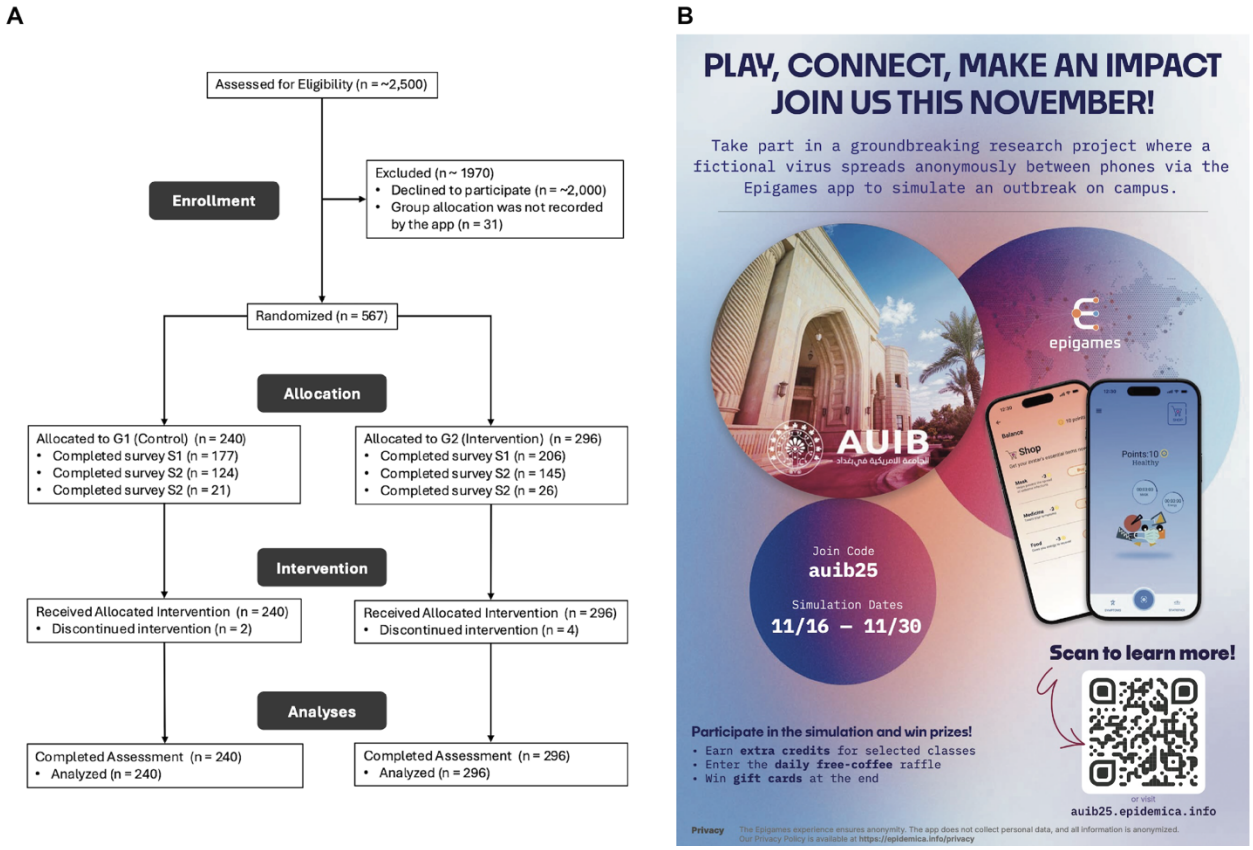


Figure 1: Study implementation and participant recruitment. *The panel on the left (A) contains the CONSORT Flow Diagram describing how the RCT study design was implemented and all corresponding numbers of participants in each stage of the study. The exclusion of 31 participants from randomization was due to a connectivity issue that resulted in the allocated group not to be recorded by the app to the database. The panel on the right (B) shows the flier that was used for promoting the epigame on the AUIB campus.*

2.2 The Epigames app

The study utilized the *Epigames* app, which simulates a "digital pathogen" transmitting between nearby users of the app. This app was built upon the existing OO app [15] for experiential learning simulations on infectious disease [16]. The *Epigames* app uses the same Bluetooth proximity sensing technology as OO, based on the open-source Herald project [17], to sample high-resolution contact networks between participants, while adding additional infrastructure for gamified behavioral studies, including the assignment of participants to randomized groups, pre-, mid-, and post-game surveys, and gamified interventions such as quarantine and contact tracing (whose parameters may depend on group randomization).

For the epigame at AUIB, the digital pathogen in the app was parametrized to mirror the properties of Omicron variant of SARS-CoV-2 with an R0 of ~7 [42]. Duration of incubation period, serial interval, and recovery rate were calculated to fit approximately three consecutive waves of disease spread in the two weeks of the epigame. Similar parameters were used in an earlier pilot outbreak simulation with the OO app at an international college campus [35]. This pilot yielded highly realistic transmission trees with super-spreader statistics closely matching those of real-life COVID-19 transmission trees [43]. The epigame was seeded by the app randomly choosing several index cases among the participants.

2.3 Belief Measures

The app allows multiple-choice surveys to be administered to participants at any point during the study. For the epigame at AIUB, we constructed three surveys with the purpose of evaluating relationship between prior real-life beliefs and in-game behaviors (pre-registered hypothesis H1-H3), calculating correlation between real-life and in-game beliefs, and between demographic variable and in-game behavior (research questions RQ1 and RQ2), and obtaining self-assessment from participants after the epigame ended:

- **S1 (Real-Life Beliefs):** Baseline survey measuring perceived susceptibility (Q1) and severity (Q2) about respiratory viruses, and self-efficacy (Q3) and benefits (Q4) regarding voluntary quarantine, both in real-world settings. It was administered to participants shortly after they provided informed consent and joined the epigame and before any transmission has occurred. This was implemented in this way since no data collection was possible before consent, but early enough to minimize any bias resulting from the study itself.
- **S2 (In-Game Beliefs):** Survey measuring the same constructs as in S1 (susceptibility, severity, self-efficacy, and benefits) but now reframed in the context of the epigame (*"If your avatar gets infected with the virtual pathogen…"*) to provide the necessary data for the real-life-to-game mapping calculations. It was administered at the end of the third day of the epigame, when index cases were already being assigned, but symptoms were still to be shown through the *Epigames* app. In this way, participants had time to familiarize themselves with the app so they could provide meaningful responses to this survey without being influenced by the simulated transmission or disease events.
- **S3 (Exit Survey):** Administered on the final day, this survey assessed the epigame's external validity and user experience. Participants evaluated the realism of disease transmission and their own quarantine decisions, reported their anxiety regarding virtual infection, and provided usability feedback and demographic data.

These surveys are provided in the supplementary materials, and all used a Likert scale between 1 and 6, except for demographic questions (gender, school affiliation), to avoid a neutral midpoint and therefore requiring respondents to give a non-neutral view. They were also available in Arabic if the participants had their phone's user interface set to that language as default.

2.4 User Interface and Gamified Quarantine Choice

Participants joined the epigame by first installing the *Epigames* app, available free of charge from the Apple and Google app stores, then entering a code that uniquely identified the AUIB epigame, granting access to Bluetooth and other services required by the app, and finally providing informed consent digitally (**Figure 2**). Throughout the duration of the epigame, participants were updated about their simulated health status through an animated avatar on the

main screen on the app. Each participant was randomly assigned a distinct character for their avatar from a set of five different, pre-determined characters. The whimsical design of these characters was inspired by the concept of a “digital pet” [44, 45] to elicit user attachment with their avatar, increase engagement during the game, and potentially strengthen the mapping between real-life and in-game behaviors [46]. Feedback from earlier pilots suggested that users found these avatars pleasant and some requested making them more interactive and customizable, which opens possible paths for future improvements in the app’s user experience. Beyond this aspect, the *Epigames* app offered other interface features, including a screen displaying nearby contacts and their respective health and quarantine states. The surveys were administered through the app via push notifications. Upon tapping on these notifications, participants had to complete the associated survey to continue using the app.

A key feature in this epigame was the quarantine choice. This was implemented with a popup in the app that was presented daily to all participants between 8:00 AM and 1:00 PM. To model the burden of real-world quarantine without requiring physical isolation, we operationalized “cost” through a differential point-based incentive structure:

1. **Quarantine:** Isolate the avatar to prevent infection due to interactions in the real world (by suspending the app’s Bluetooth proximity sensing). The participant receives the lowest number of points by making this choice (Low Reward). This point penalty serves as a digital proxy for the real-life costs of quarantining (e.g., lost income, social isolation), operationalizing the "barrier" to safety.
2. **Do Not Quarantine:** Risk infection of the avatar by allowing interactions in the real-world as measured via Bluetooth, but the participant receives a larger number of points (High Reward), representing the immediate utility and benefit of maintaining normal social and economic activity.

Infection resulted in “illness” (during which point rewards were reduced to a minimum irrespective of quarantine choice), clearly represented by changes in app’s UI color and avatar appearance. Illness could lead to “death” of the avatar and loss of all accumulated points. Total points for each participant were displayed on a public leaderboard in the app that updated daily as participants acquired points by making their quarantine choices or lost points due to illness.

We expressly designed this mechanic to abstract the multifaceted burden of quarantine into a quantifiable economic decision. While we recognize that a reduction in app points cannot replicate the physical, psychological, and social costs of actual isolation, thereby limiting the context parallelism of the study, this abstraction allowed us to experimentally manipulate the cost variable across treatment groups (G1 vs. G2) to measure behavioral elasticity.

2.5 Outcome measures

The primary outcome was the daily quarantine decision (binary), total days quarantined (integer), and overall quarantine rate (continuous) over the 15-day period, all per participant.

Secondary outcomes included the correlation between real-life (S1) and in-game (S2) belief scores, and the interaction effect of economic cost (represented by the group assignment, G1 or G2) on these behaviors, also per participant. Number of daily contacts and simulated infections were used for descriptive analysis, and final point scores of each participant were used in the lottery for gift cards after the epigame ended.

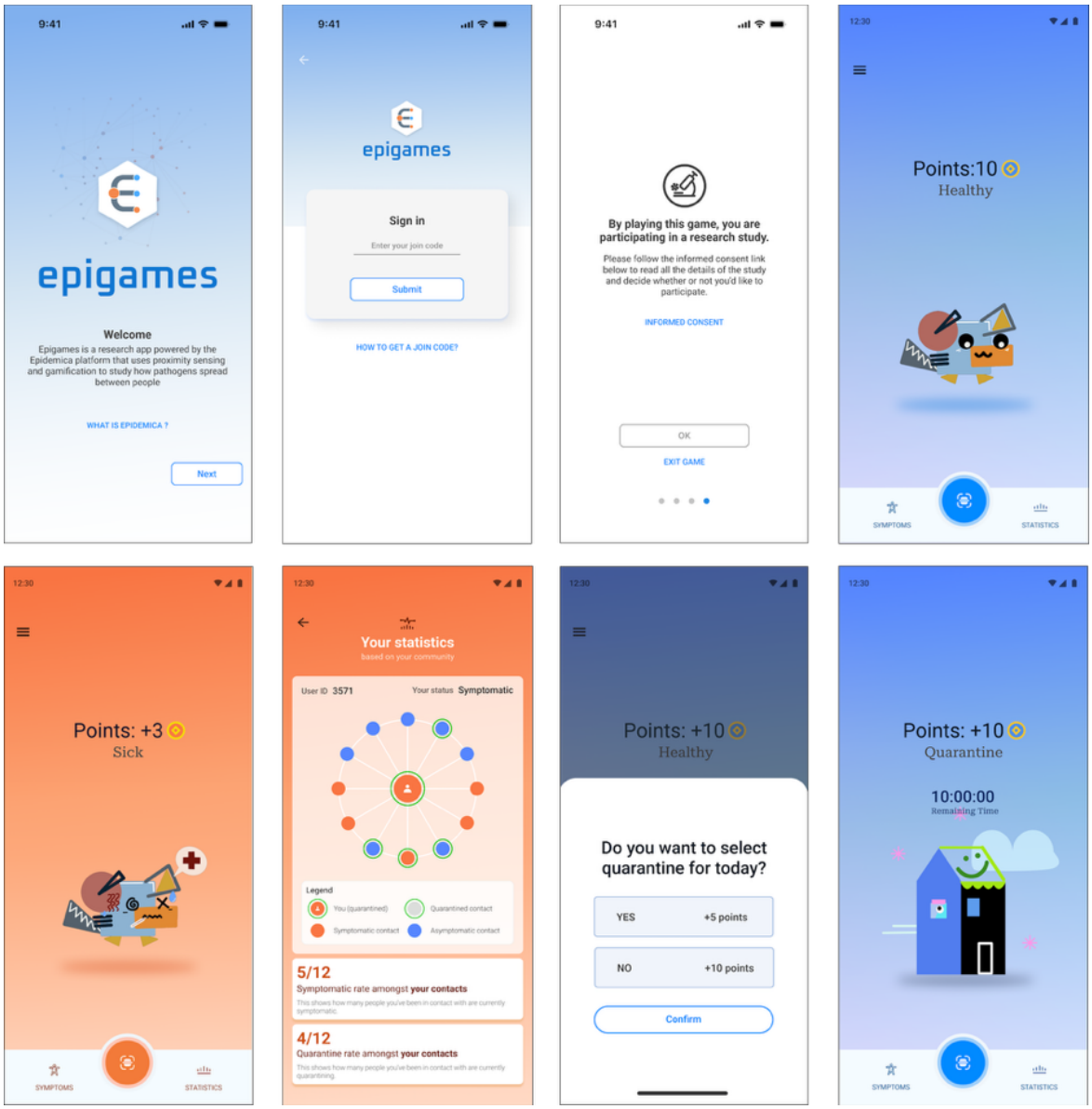


Figure 2: Epigames app. *The images in the top row depict, from left to right, the welcome, sign in, informed consent, and home screens. The avatar in the home screen conveys the health and quarantine status of the user. In the bottom row, from left to right, the home screen displaying the avatar in the sick state, the statistic screen with information of neighboring players and their respective statuses, the popup with daily quarantine choice, and home screen with avatar in quarantined state.*

2.5 Determining the Economic Barriers to Quarantine

A "Willingness to Accept" (WTA) [47] pilot survey (N=50) was conducted prior to the main experiment. We shared an anonymous Google form via a Telegram chat regularly used by AUIB students, which presented with the following scenario and questions:

- Scenario: "In the upcoming epigame, you will make a daily choice about your in-game avatar:
  - Choice A: 'Quarantine' your avatar. You get 5 points, guaranteed.

- Choice B: ‘Do Not Quarantine’ your avatar. This is the risky choice. You get more points, but only if you avoid your avatar from getting virtually ‘infected’. If your avatar gets infected, you risk losing all your points.”

- Questions:
  - “If Choice B gave you 10 points (a 5-point gap), which would you choose? (A or B)”
  - “If Choice B gave you 15 points (a 10-point gap), which would you choose? (A or B)”
  - “If Choice B gave you 20 points (a 15-point gap), which would you choose? (A or B)”
  - “If Choice B gave you 25 points (a 20-point gap), which would you choose? (A or B)”
  - “If Choice B gave you 30 points (a 25-point gap), which would you choose? (A or B)”

Analysis of the responses was conducted to find the median acceptance threshold, defined as the point gap at which approximately 50% of the respondents “flip” from Choice A to Choice B. This identified the point of maximum behavioral tension at a gap of 12 points (see **Suppl. Figure 1**). Consequently, participants were randomized into:

- **Group 1 (Low Barrier):** In this group, participant who quarantined received 5 points (Low Reward), those who did not received 10 points (High Reward), therefore the economic barrier to quarantine was 5 points (nominal/low).
- **Group 2 (High Barrier):** In this group, participant who quarantined received 5 points (Low Reward) as before, but those who did not, received 17 points instead (High Reward), therefore the economic barrier to quarantine was 12 points (high).

2.6 Statistical Analysis

The mapping between real-life and game beliefs (as measured by surveys S1 and S2) was quantified by calculating the Spearman's rank correlation coefficient between the corresponding survey responses: susceptibility (Q1), severity (Q2), self-efficacy (Q3), and benefits (Q4).

We employed Generalized Linear Models (GLM) with a Poisson regression to model the count of quarantine actions. The Poisson regression is designed for count data, ensuring predictions are positive and handling the specific distribution of counts better [48]. To address the “zero-inflation” of Likert scales, all belief scores (1-6) were anchored to 0 (score - 1). To model the quarantine rate as a function of the belief constructs in our HBM, we used a Binomial GLM instead with a logit link function (logistic regression) as in other recent studies [49], since in this case we are predicting the probability of adopting the quarantine behavior, not the counts.

Network analysis, albeit not the focus of this analysis, was conducted to characterize the global structure of the contact network of the participants of the AUIB epigame, as well as to calculate the assortativity of participants’ reported beliefs. Very briefly, assortativity measures the extent to which individuals with similar node properties (such as their survey responses) have higher number of contacts between them than the average, or lower number of contacts (negative assortativity) [50].

All statistical analyses and plots were performed with the Python programming language version 3.13.5 and the following packages: numpy 2.3.0, pandas 2.3.0, pyarrow 20.0.0, pygraphviz 1.14, matplotlib 3.10.3, network 3.5, scikit-learn 1.7.0, scipy 1.15.2, statsmodels 0.14.4, and seaborn 0.13.2. Link to code repository is provided in the supplementary materials.

3. Results

3.1 Descriptive Statistics & Epigame Dynamics

The daily contacts logged by the app over the 15-day period of the epigame revealed weekly patterns of interaction on the AUIB campus, including reduced attendance on Thursdays (reserved for biology labs) and the Friday-Saturday weekend observed in Iraq. The simulated pathogen successfully seeded, resulting in 52 infections, with the peaks in daily number of new "cases" following the peaks in daily contacts (**Figure 3A**). A majority of participants chose not to quarantine on most days (**Figure 3B**). Visual inspection of survey histograms (**Figure 4A** and **4B**, S1 vs. S2) indicates that the distribution of in-game beliefs closely mirrored real-life beliefs and the exit survey S3 (although it was affected by attrition as only 47 participants completed it) showed high scores for the self-assessed realism of behavior during the study (**Figure 4C**).

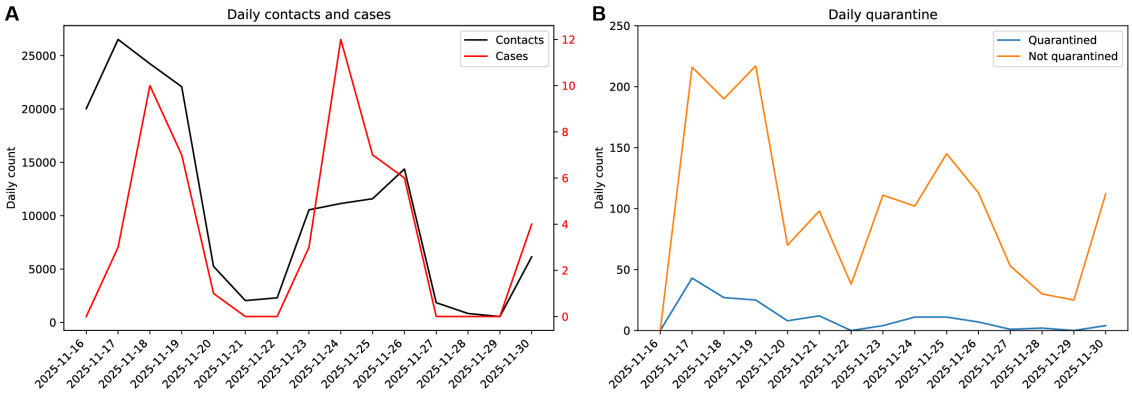


Figure 3: Epigame dynamics. *The plot in the left shows the number of contacts detected via Bluetooth (black line) and simulated new cases (red line) each day of the epigame, while the plot on the right shows the number of participants who chose to quarantine/not to quarantine (blue and orange, respectively).*

3.2 External Validity: Real-Life to In-Game Belief Mapping

A central research question of this study was to determine the extent to which users mapped their real-world beliefs into the epidemic game (RQ1). The surveys S1 and S2 were precisely constructed with the purpose of answering RQ1 in mind, as they asked the same questions, but framed in the real-life and in-game settings, respectively:

- **Correlation:** All corresponding pairs of real-life (S1) and in-game (S2) variables exhibited significant positive correlation, albeit of moderate magnitude (**Figure 5**). For perceived susceptibility (Q1) the Spearman rank correlation coefficient $\rho=0.2$ between S1 and S2 (p-value=0.0012), for perceived severity, $\rho=0.13$ ($p = 0.037$), for perceived quarantine self-efficacy, $\rho=0.27$ ($p < 0.001$), and for perceived quarantine benefits, $\rho=0.36$ ($p < 0.001$).
- **Prediction Models:** OLS regression showed that real-life beliefs were significant predictors of most in-game beliefs. This was strongest for perceived benefits ($R^2 = 0.142$, $\beta = 0.322$, $p < 0.001$) and self-efficacy ($R^2 = 0.081$, $\beta = 0.215$, $p = 0.001$), while perceived susceptibility showed a weaker association ($R^2 = 0.058$, $\beta = 0.149$, $p = 0.004$) and not significant for perceived severity ($R^2 = 0.042$, $\beta = 0.112$, $p = 0.127$) in the OLS model.

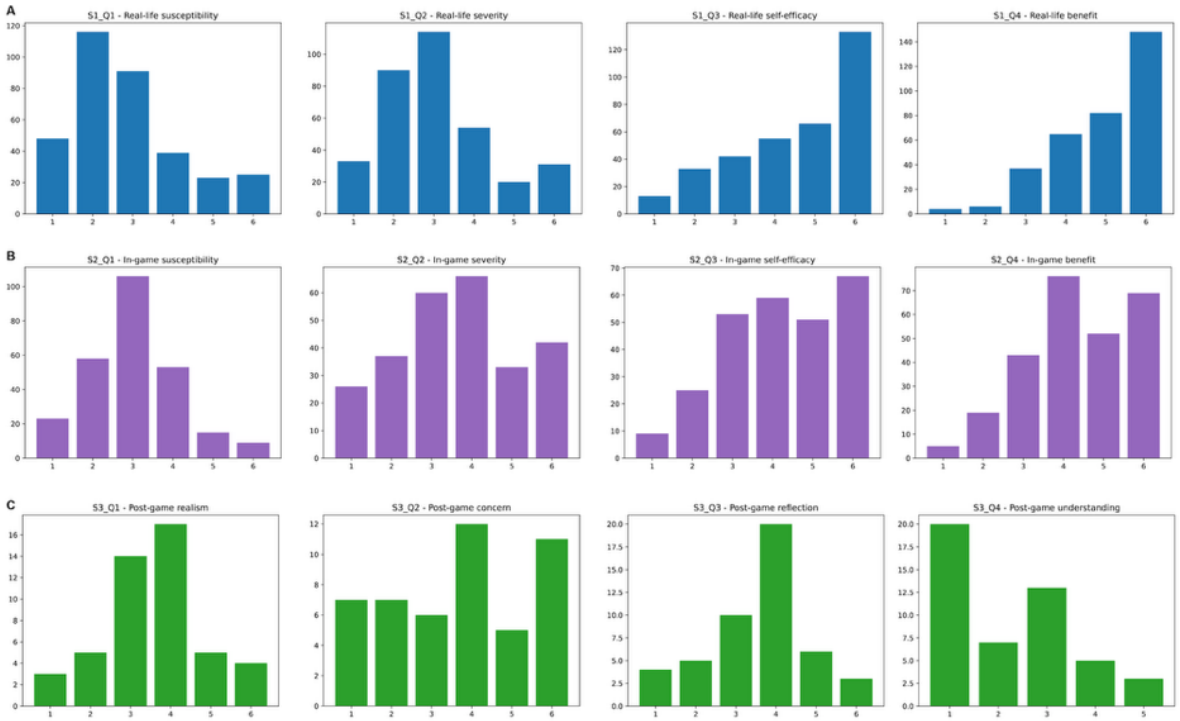


Figure 4: Survey responses. *The top row of histograms (A) shows the counts of the answers on a 1-6 Likert scale to the first 4 questions of survey S1 administered shortly after participants joined the AUIB epigame on perceived susceptibility and severity about respiratory viruses, and self-efficacy and benefits regarding voluntary quarantine in real-life settings. The middle row of histograms (B) shows the answers to the survey S2, administered on the third day of the epigame and on the same beliefs as S1 but reframed in the context of the game (see Supplementary Materials for the full list of questions in each survey). The bottom row of histograms (C) shows the responses to the exit survey (S3) on the final day of the epigame where participants self-assessed realism of the simulated epidemic, concern about virtual infection, realism of their own quarantine decisions, and difficulty in understanding the information presented through the app.*

3.3 Correlation between Gender and Quarantine

Exploration of RQ2 yielded gender as the only demographic variable associated with the quarantine behavior during the epigame. Female participants showed significantly higher quarantine adoption than males: $\rho=0.2$ (p-value < 0.001) between overall quarantine count over the entire epigame and gender reported in S1, consistent with real-world epidemiological literature [51].

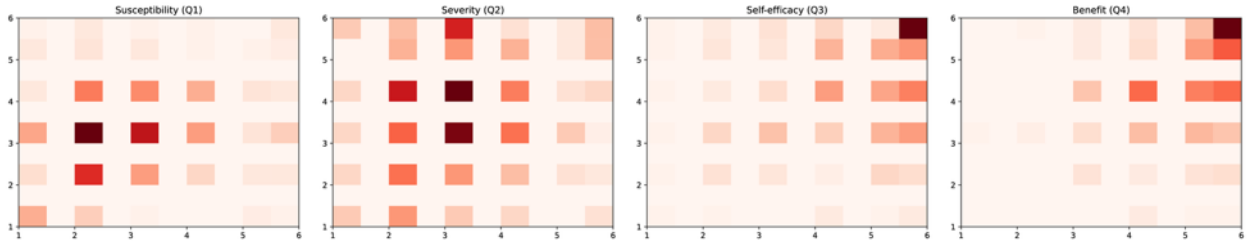


Figure 5: S1-S2 correlation. *These 2D histogram plots correspond to the susceptibility, severity, self-efficacy, and benefit question in S1 and S2. The X axis is the value of the response in S1, while the Y axis, the response in S2.*

3.4 The Effect of Economic Barriers (Hypothesis Testing)

The hypothesis registered on OSF prior to the study included:

- **H1 (Main Effect):** Participants in Group 1 (Low Barrier) will have a significantly higher overall quarantine rate than participants in Group 2 (High Barrier).
- **H2 (Moderation by Cost):** The correlation between S1 beliefs (e.g., S1-Q2 perceived severity, S1-Q4 perceived effectiveness) and quarantine behavior will be significantly stronger in Group 1 (Low Barrier) than in Group 2 (High Barrier). We hypothesized the high economic cost in G2 will reduce the influence of personal beliefs.
- **H3 (Interaction by Efficacy):** The difference in quarantine rates between G1 and G2 (the main effect of the barrier) will be moderated by baseline self-efficacy. Specifically, the effect of the higher barrier in G2 will be largest for participants who report low S1 self-efficacy (S1-Q3). Participants with high self-efficacy may be unaffected by the cost manipulation, as their behavior is driven by their confidence, not by the barrier.

Initial aggregate comparisons showed no significant difference in average quarantine rates between G1 and G2 ($p > 0.05$). Further examination with a multivariate Poisson model using the belief variables in S1 and group variable (G1 or G2) as the interaction term yielded no significant effects (**Suppl. Table 1**), resulting in the rejection of all three proposed hypothesis.

However, turning into the game beliefs (S2), the same multivariate Poisson regression model with group interaction showed a significant crossover interaction (pseudo $R^2 \sim 0.09$). Results of the Poisson regression are shown in **Table 1**.

Table 1: Poisson Regression Results (quarantine count, S2 beliefs, and group interaction)

Predictor	Coefficient	p-value	Interpretation
Intercept	-1.28	0.008	Baseline quarantine for G1 (Low Cost) with min. motivation.
Group 2 (High Cost)	**-2.22**	**0.006**	**Barrier Effect.** For "skeptical" users (those with low perceptions of susceptibility, severity, self-efficacy, and benefit in S2), high cost reduces quarantine to ~11% of G1 levels.
S2-Benefit (Q4)	-0.03	0.819	In Low Cost (G1), beliefs don't drive action (everyone acts).
S2-Benefit Group 2	**+0.34**	**0.076**	**Interaction.** In High Cost (G2), higher belief in benefit of quarantine benefit increases action at a close-to-significance level.

We interpret these results by concluding that the high cost or "economic" barrier acted as a filter. The cost barrier reduced quarantine among unmotivated users in the game, who are those reporting low perceptions of susceptibility, severity, self-efficacy, and benefit in S2 (significant negative barrier effect). However, as perceived benefit of quarantine increases, users in the High Cost group become more willing to pay the cost (positive interaction), reducing the gap with the Low Cost group, at a close-to-significance level.

3.5 Health Belief model

To quantify how specific individual beliefs drive protective behavior, we constructed a Binomial Generalized Linear Model (GLM) based on the Health Belief Model (HBM) framework. In this model, the dependent variable was the participant's Quarantine Rate (probability of action) during the epigame, weighted by their total number of decision trials.

The predictors included the in-game belief scores (S2, centered to 0), gender, and the experimental group. In this formulation, the Group variable acts as a proxy for the "Perceived Barrier" (Cost), which typically enters HBM equations as a subtraction from perceived benefits.

Table 2: Health Belief Model parametrized with in-game beliefs

Predictor	Coefficient	Std. Err	P-value	Interpretation
Intercept	**-3.99**	0.476	< 0.001	Baseline probability for min. score & low cost.
Susceptibility (S2-Q1)	**+0.27**	0.100	**0.007**	Strongest belief driver of behavior.
Severity (S2-Q2)	-0.03	0.076	0.647	Perceived severity did not drive action.
Self-Efficacy (S2-Q3)	+0.15	0.084	0.067	Marginally significant positive driver.
Benefits (S2-Q4)	+0.11	0.098	0.281	Positive direction, but not significant in this additive model.
Gender (Female)	**+0.37**	0.175	**0.035**	Females significantly more likely to quarantine.
Barrier (Cost/Group)	-0.15	0.211	0.475	Negative direction (consistent with theory) but insignificant as a main effect.*

**Note: While the high cost (Group) was a significant driver in the interaction model (Section 3.3), it is not significant as a simple additive term here, likely due to the crossover effect described previously.*

The model achieved a pseudo R^2 (CS) of 0.072. The results highlight that, within the context of the game, perceived susceptibility to infection and gender were the most distinct statistical drivers of choosing the quarantine option in the app, while self-efficacy played a supporting role. We also fitted a similar HBM but on the S1 (real-life belief) variables, to predict the same in-game quarantine rate (**Suppl. Table 2**). This HBM had a lower pseudo R^2 of 0.039.

3.7 Analysis of Network Structure

The contact network aggregated over the 15 days of the epigame included 453 nodes and 3372 edges (nodes with no edges were removed, most likely caused by limitation in Bluetooth contact detection on some phones, see limitations section). The network exhibited the scale-free property (degree distribution following a power law) (**Suppl. Figure 2A**).

In terms of assortative mixing, we observed statistically significant positive assortativity for several node (participant) properties: days not in quarantine (Pearson's r=0.46, p-value=0.006), perceived real-life susceptibility (S1-Q1) (r=0.23, p=0.05), perceived in-game severity (S2-Q2) (r=0.40, p=0.007), and perceived in-game benefit (S2-Q4) (r=0.48, p=0.002), suggesting that

participants with similar beliefs tended to cluster in the network, which could be caused by several processes, including social proof (people copying the actions of others) and the "birds of a feather" effect (connection promoted by similarity) often found in network science. Our demographic variable, gender, also showed significant assortativity, ($r=0.34$, $p=0.008$). See **Suppl. Figure 2B** for network graphs colored by these assortative variables.

4. Discussion

4.1 Epigames as a Valid Proxy for Real-Life Protective Behavior

The significant correlations observed between real-life and in-game beliefs regarding disease susceptibility, severity, and quarantine self-efficacy and benefit, provide initial evidence for the external validity of epigames. Furthermore, the replication of well-established real-world gender effects, with female participants exhibiting significantly higher quarantine adoption than males, suggests that participants did not merely role-play arbitrary personas but projected their authentic risk profiles into the simulated scenario. These findings justify the use of gamified experiments as a "epidemiology laboratory" for probing the behavioral factors influencing protective health behaviors during infectious disease outbreaks.

However, we must acknowledge that this real-life-to-game mapping, while statistically significant, was relatively noisy (Pearson's $r < 0.4$ across all health belief constructs). This indicates that while the game successfully elicits real-world values, it also introduces a layer of game-specific context that weakens these values. Future iterations of the platform must focus on improving this mapping, perhaps through more immersive narratives or higher-stakes incentives, to increase the predictive fidelity of in-game actions.

4.2 Cost of Behavior in Determinant for the Skeptical

Our findings challenge simple "one-size-fits-all" behavioral models. The high economic barrier in Group 2 did not uniformly suppress quarantine behavior during the epigame; rather, it acted as a gatekeeper, specifically suppressing action among those with low health beliefs (low susceptibility/benefit) for the game setting. In the low-barrier group (Group 1), quarantine was "cheap," meaning even participants with low motivation adopted the behavior. This reveals a critical nuance for public health interventions, at least in gamified scenarios like this one: economic subsidies (lowering the barrier) are likely most effective for capturing the "skeptical tail" of the population, whereas educational interventions (increasing perceived benefit) matter most when the behavior is costly to the individual. But since these conclusions are based on in-game, not real-life beliefs & behaviors, extrapolation to real-life settings is limited, at least with the current implementation of the *Epigames* app and gamified quarantine intervention.

4.3 A Framework for Predicting Adoption of Protective Behavior

A key contribution of this study is the parameterization of an in-game Health Belief Model (HBM) that can potentially forecast real-world compliance. While "in-game beliefs" (S2) are unknown prior to an experiment, "real-life attitudes" (S1) can be easily measurable via psychosocial surveys. Our analysis supports a two-step prediction framework: (1) measuring real-life beliefs in a target population, (2) imputing expected in-game beliefs using the probabilistic relationship between the two established here, and (3) inputting these values into

the parameterized game HBM to forecast behavior under different cost constraints. While the current mapping introduces variability ($R^2 < 0.15$), this framework demonstrates how epigames can serve as a "behavioral calibration layer," translating raw survey data into more accurate probabilities of action under specific cost and potentially other constraints.

4.3 Limitations and Future Directions

While our results are promising, this study has several limitations that warrant careful consideration moving forward.

First, technical and design constraints. The study faced technical challenges, including Bluetooth connectivity issues on certain Android/iOS versions, which likely undercounted contacts and simulated infections. Additionally, we could not enforce network-aware clustering for randomization in the AUIB epigame. This created a risk of "information leakage," where participants in the Low Barrier group may have communicated point values to those in the High Barrier group, potentially dampening the perceived contrast of the intervention. Features such as the point leaderboard and the gift card lottery were implemented to create tangible motivations for strategic decision-making but may have over-incentivized participants to simply try to maximize their point score regardless of other considerations, biasing the results and reducing behavioral parallelism. Participant dropout was also significant towards the end of the AUIB epigame, with fewer than 50 participants completing the exit survey (S3), highlighting the need for stronger incentives for post-game engagement, or alternatively, shorter epigames.

Second, the reliance on a single epigame instance to draw conclusions about external validity, which is a fundamental limitation of this study and the field of experimental games more broadly. Without running multiple, distinct epigames with varying parameters (e.g., different narratives, mechanics, or populations), it is difficult to disentangle which validity failures are due to the specific design of this game versus a fundamental disconnect between games and reality. Confidence in validity will increase not just by seeing what works, but by observing variance across different contexts. Future research should prioritize multi-game comparisons to identify the specific features that contribute to or detract from validity.

Third, magnitude vs. significance. While the key correlations we found in this study were statistically significant, their magnitude was modest. This "noisy signal" is characteristic of behavioral data but suggests that individual-level prediction remains challenging. The game successfully captured aggregate trends (e.g., gender effects, cost moderation) but was less effective at predicting the exact actions of specific individuals.

Finally, regarding context parallelism: The quarantine mechanic as currently implemented—a digital button press—was designed to achieve indicator parallelism (a clear binary choice) while accounting for practical considerations (participants cannot be asked to physically quarantine) but because of this, it lacked the experiential weight of real-world isolation, thereby limiting context parallelism. To address this, future studies should focus on enhancing the salience of behavioral choices, ensuring that the decision "matters" psychologically to the player. This could be achieved by designing interactions that enforce "digital isolation" (e.g., temporary app lockout) or by fostering a deeper emotional attachment to the avatar. Such mechanisms would better mimic the non-monetary costs of quarantine (e.g., boredom, FOMO) and the emotional consequences of infection, thereby narrowing the gap between a gamified decision and a real-world dilemma.

5. Conclusion

This study demonstrates that app-based epigames are plausible experimental tools for behavioral economics and network epidemiology, capable of revealing the "price" of safety: high costs deter the unmotivated, but strong beliefs can overcome economic barriers. While technical and design limitations introduced noise, significant interaction effects suggest that experimental games can serve as valuable proxies for pre-testing public health interventions. Future work involving larger, multi-game deployments will be essential to constructing and refining predictive models for policy application, including agent-based systems driven by autonomous agents trained with real-world behavioral data.

Acknowledgments

The authors wish to thank the American University of Iraq – Baghdad for supporting this project. We would also like extend our gratitude to the AUIB students who volunteered to help during recruitment and logistics of the epigame: Waqas Saad Salih, Mostafa Nabeel Mohammed, Mohammedali Hussein Mihbas, Basmala Mazen Fakihir, Maryam Ahmed Naji, Daniah Ahmed Mawlood, Dur Ali Noaman, and Dimah Nayzak Jamal.

Sources of funding

The project was supported with startup grant funds from AC, compute resources from the Broad Institute, and in-kind contributions from DK.

References

1. Crawford, F.W., et al., *Impact of close interpersonal contact on COVID-19 incidence: Evidence from 1 year of mobile device data.* Sci Adv, 2022. **8**(1): p. eabi5499.
2. Wong, K.L.M., et al., *Social contact patterns during the COVID-19 pandemic in 21 European countries - evidence from a two-year study.* BMC Infect Dis, 2023. **23**(1): p. 268.
3. Yuan, Z., et al., *Post pandemic fatigue: what are effective strategies?* Sci Rep, 2022. **12**(1): p. 9706.
4. Wilke, C.O. and C.T. Bergstrom, *Predicting an epidemic trajectory is difficult.* Proc Natl Acad Sci U S A, 2020. **117**(46): p. 28549-28551.
5. Verelst, F., L. Willem, and P. Beutels, *Behavioural change models for infectious disease transmission: a systematic review (2010-2015).* J R Soc Interface, 2016. **13**(125).
6. Hamilton, A., et al., *Incorporating endogenous human behavior in models of COVID-19 transmission: A systematic scoping review.* Dialogues Health, 2024. **4**: p. 100179.
7. van Dijk, E. and C.K.W. De Dreu, *Experimental Games and Social Decision Making.* Annu Rev Psychol, 2021. **72**: p. 415-438.
8. Cates, J.R., et al., *Developing a Serious Videogame for Preteens to Motivate HPV Vaccination Decision Making: Land of Secret Gardens.* Games Health J, 2018. **7**(1): p. 51-66.
9. Wang, Z., et al., *Communicating sentiment and outlook reverses inaction against collective risks.* Proceedings of the National Academy of Sciences, 2020. **117**(30): p. 17650-17655.

10. Maharaj, S., T. McCaldin, and A. Kleczkowski. *A participatory simulation model for studying attitudes to infection risk*. Vista, CA: Society for Modeling & Simulation International.
11. Kleczkowski, A., et al., *Spontaneous social distancing in response to a simulated epidemic: a virtual experiment.* BMC public health, 2015. **15**(1): p. 973-973.
12. Parsons, D. *Mobile Participatory Simulation of COVID-19 Transmission Using the micro:bit*. IEEE.
13. Woike, J.K., et al., *The transmission game: Testing behavioral interventions in a pandemic-like simulation.* Sci Adv, 2022. **8**(8): p. eabk0428.
14. Nunner, H., et al., *Disease avoidance threatens social cohesion in a large-scale social networking experiment.* Sci Rep, 2023. **13**(1): p. 22586.
15. Chen, D.L., M. Schonger, and C. Wickens, *oTree—An open-source platform for laboratory, online, and field experiments.* Journal of Behavioral and Experimental Finance, 2016. **9**: p. 88-97.
16. Vriens, E., et al., *Assessing the effects of pandemic risk on cooperation and social norms using a before-after Covid-19 comparison in two long-term experiments.* Scientific Reports, 2024. **14**(1): p. 3356.
17. Harrison, G.W., et al., *A case study of an experiment during the COVID-19 pandemic: online elicitation of subjective beliefs and economic preferences.* Journal of the Economic Science Association, 2021. **7**(2): p. 194-209.
18. Lofgren, E.T. and N.H. Fefferman, *The untapped potential of virtual game worlds to shed light on real world epidemics.* The Lancet Infectious Diseases, 2007. **7**(9): p. 625-629.
19. Yoneki, E. *Fluphone study: Virtual disease spread using haggle*. in *Proceedings of the 6th ACM Workshop on Challenged Networks*. 2011.
20. Williams, D., *The Mapping Principle, and a Research Framework for Virtual Worlds.* Communication theory, 2010. **20**(4): p. 451-470.
21. Pisor, A.C., et al., *Preferences and constraints: the value of economic games for studying human behaviour.* Royal Society Open Science, 2020. **7**(6): p. 192090.
22. Baumard, N. and D. Sperber, *Weird people, yes, but also weird experiments.* Behavioral and Brain Sciences, 2010. **33**(2-3): p. 84-85.
23. Hagen, E.H. and P. Hammerstein, *Game theory and human evolution: A critique of some recent interpretations of experimental games.* Theoretical Population Biology, 2006. **69**(3): p. 339-348.
24. Loewenstein, G., *Experimental Economics From the Vantage-point of Behavioural Economics.* The Economic Journal, 2001. **109**(453): p. 25-34.
25. Pisor, A.C. and D.M.T. Fessler, *Importing social preferences across contexts and the pitfall of over-generalization across theories.* Behavioral and Brain Sciences, 2012. **35**(1): p. 34-35.
26. Galizzi, M.M. and D. Navarro-Martinez, *On the external validity of social preference games: a systematic lab-field study.* Management Science, 2019. **65**(3): p. 976-1002.
27. Lightner, A.D., P. Barclay, and E.H. Hagen, *Radical framing effects in the ultimatum game: the impact of explicit culturally transmitted frames on economic decision-making.* Royal Society Open Science, 2017. **4**(12): p. 170543.
28. Jones, L.M. and M. Benthien, *Preparing for a "Big One": the great southern California ShakeOut.* Earthquake Spectra, 2011. **27**(2): p. 575-595.

29. Initiative, H.H. *Humanitarian Response Intensive Course*. 2025; Available from: https://hhi.harvard.edu/humanitarian-response-intensive-course.
30. Naar, N., *Gaming anthropology: The problem of external validity and the challenge of interpreting experimental games.* American Anthropologist, 2020. **122**(4): p. 784-798.
31. Naar, N., M.N. Grote, and M. Borgerhoff Mulder, *What People Say and How People Play: Mixed Methods and Multiple Measures of Fishing Behavior with Implications for External Validity.* Global Perspectives, 2022. **3**(1).
32. Colubri, A. *Epidemica platform*. 2025; Available from: https://epidemica.info.
33. Colubri, A., et al., *Preventing Outbreaks through Interactive, Experiential Real-Life Simulations.* Cell, 2020. **182**(6): p. 1366-1371.
34. Specht, I., et al., *Analyzing the impact of a real-life outbreak simulator on pandemic mitigation: An epidemiological modeling study.* Patterns (N Y), 2022. **3**(8): p. 100572.
35. Musa, S.S., et al., *The effect of Behavioral Factors and Intervention Strategies on Pathogen Transmission: Insights from a Two-Week Epidemic Game at Wenzhou-Kean University in China.* medRxiv, 2024: p. 2024.12. 14.24318955.
36. Foster, E.D. and A. Deardorff, *Open science framework (OSF).* Journal of the Medical Library Association: JMLA, 2017. **105**(2): p. 203.
37. Rosenstock, I.M., *Historical origins of the health belief model.* Health education monographs, 1974. **2**(4): p. 328-335.
38. Rosenstock, I.M., *Why people use health services.* The milbank memorial fund quarterly, 1966. **44**(3): p. 94-127.
39. Alyafei, A. and R. Easton-Carr, *The Health Belief Model of Behavior Change*, in *StatPearls*. 2024, StatPearls Publishing. Copyright © 2024, StatPearls Publishing LLC.: Treasure Island (FL).
40. Atchison, C., et al., *Early perceptions and behavioural responses during the COVID-19 pandemic: a cross-sectional survey of UK adults.* BMJ Open, 2021. **11**(1): p. e043577.
41. Behaviours, S.P.I.G.o., *The impact of financial and other targeted support on rates of self-isolation or quarantine.* 2020.
42. Liu, Y. and J. Rocklöv, *The effective reproductive number of the Omicron variant of SARS-CoV-2 is several times relative to Delta.* J Travel Med, 2022. **29**(3).
43. Taube, J.C., P.B. Miller, and J.M. Drake, *An open-access database of infectious disease transmission trees to explore superspreader epidemiology.* PLoS Biology, 2022. **20**(6): p. e3001685.
44. Donath, J., *Artificial pets: Simple behaviors elicit complex attachments*. 2004, The Encyclopedia of Animal Behavior, Greenwood Press.
45. Dodd, M., A. Fowler, and D. Lottridge, *Purr-ogrammed love: A narrative review of virtual pets.* Entertainment Computing, 2025. **54**: p. 100958.
46. Bowman, N.D. and J. Banks, *Player-Avatar Identification, Relationships, and Interaction: Entertainment Through Asocial, Parasocial, and Fully Social Processes*, in *The Oxford Handbook of Entertainment Theory*, P. Vorderer and C. Klimmt, Editors. 2021, Oxford University Press. p. 0.
47. Whittington, D., W. Adamowicz, and P. Lloyd-Smith, *Asking willingness-to-accept questions in stated preference surveys: a review and research agenda.* Annual Review of Resource Economics, 2017. **9**(1): p. 317-336.
48. Hutchinson, M.K. and M.C. Holtman, *Analysis of count data using poisson regression.* Res Nurs Health, 2005. **28**(5): p. 408-18.

49. Wong, L.P., et al., *The use of the health belief model to assess predictors of intent to receive the COVID-19 vaccine and willingness to pay.* Hum Vaccin Immunother, 2020. **16**(9): p. 2204-2214.
50. Newman, M.E.J., *Mixing patterns in networks.* Phys. Rev. E Stat. Nonlin. Soft Matter Phys., 2003. **67**(2 Pt 2): p. 026126.
51. Galasso, V., et al., *Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries.* Proc Natl Acad Sci U S A, 2020. **117**(44): p. 27285-27291.

Supplementary Materials

Source code: https://github.com/colabobio/auib-epigame-analysis

Pre-, Mid-, and Post-game Surveys

Survey S1

Title: Survey about infectious diseases in real life

Description: Chose the responses that best represent your beliefs. There is no right or wrong answer.

Q1: How likely do you think it is that you will get infected with a respiratory virus (e.g., flu, COVID) in the next three months? 1. Virtually impossible … 6. Almost certain

Q2: If you were to get infected with a respiratory virus (e.g., flu, COVID) in the next three months, how serious do you think the impact would be on your health? 1. No impact at all … 6. Very significant impact

Q3: How confident you are that you would be able to successfully follow home quarantine if there is a new infectious disease with potentially serious health effects (such as COVID-19 at the beginning of the pandemic)? 1. Not confident at all … 6. Completely confident

Q4: How effective you think following home quarantine would be in preventing the spread of infectious diseases? 1. Not effective at all … 6. Extremely effective

Q5: What is your gender? 1. Male, 2. Female, 3. Other, 4. Do not wish to disclose

Q6: What is your school affiliation? 1. Student, 2. Faculty, 3. Staff, 4. Other, 5. Do not wish to disclose

Survey S2

Title: Survey about the virtual epidemic game @ AUIB

Description: Chose the responses that best represent your beliefs. There is no right or wrong answer.

Q1: How likely do you think it is that your avatar will get infected with the virtual pathogen while playing with the Epigames app? 1. Virtually impossible … 6. Almost certain

Q2: If your avatar gets infected with the virtual pathogen, how much do you think this would affect your chances to win a gift card with the point-based lottery at the end of the game? 1. No affected at all … 6. Extremely affected

Q3: How confident you are that you can successfully follow the in-game quarantine? 1. Not confident at all … 6. Completely confident

Q4: How effective you think in-game quarantine is in preventing the spread of the virtual pathogen? 1. Not effective at all … 6. Extremely effective

Survey S3

Title: Survey about your experience during the epidemic game @ AUIB

Description: In this final survey, we are asking you to compare your gameplay to real life. This is a difficult question. Research shows that people often say they would behave one way in a

hypothetical survey, even when their actual behavior is different. Please take a moment to honestly reflect on your real-life priorities and compare them to your in-game choices. Your candid answer is the most valuable data you can provide.

Q1: How well do you think the game represents infectious disease transmission in real life? 1. Very poorly … 6. Extremely accurately

Q2: Were you worried about getting virtually infected during the game? 1. Not worried at all … 6. Very worried

Q3: How closely did your quarantine decisions during the game reflect what you would do in a real-life outbreak? 1. Not at all similar … 6. Exactly what I'd do in real life

Q4: How difficult was it to understand the information you received during the game? 1. Very easy to understand … 6. Extremely hard to understand

Q5: Please select the option that represents your field of study: 1. Healthcare Majors or Biology, 2. Law, International Studies, or English Literature, 3. Education or Psychology, 4. Business, 5. Engineering or Chemistry, 6. Computer Science

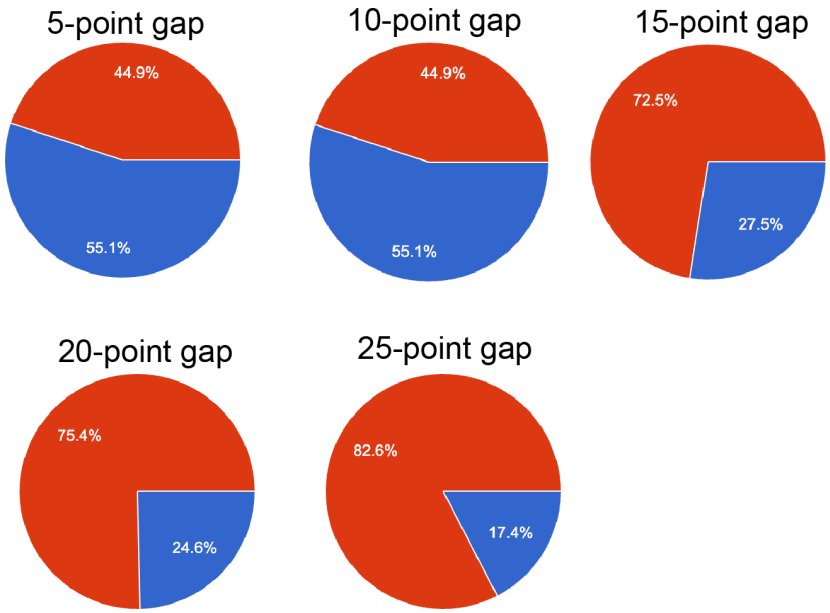
Suppl. Table 1: Poisson Regression Results (quarantine count, S1 beliefs, and group interaction)

Predictor	Coefficient	P-value	Interpretation
Intercept	-0.98	0.073	Baseline quarantine for G1 (Low Cost) with min. motivation.
Group 2 (High Cost)	-0.30	0.671	Not significant.
S2-Self efficacy (Q3)	-0.14	0.144	Not significant.
S2-Self efficacy Group 2	+0.04	0.732	Not significant.
S2-Benefit (Q4)	-0.02	0.873	Not significant.
S2-Benefit Group 2	+0.01	0.942	Not significant.

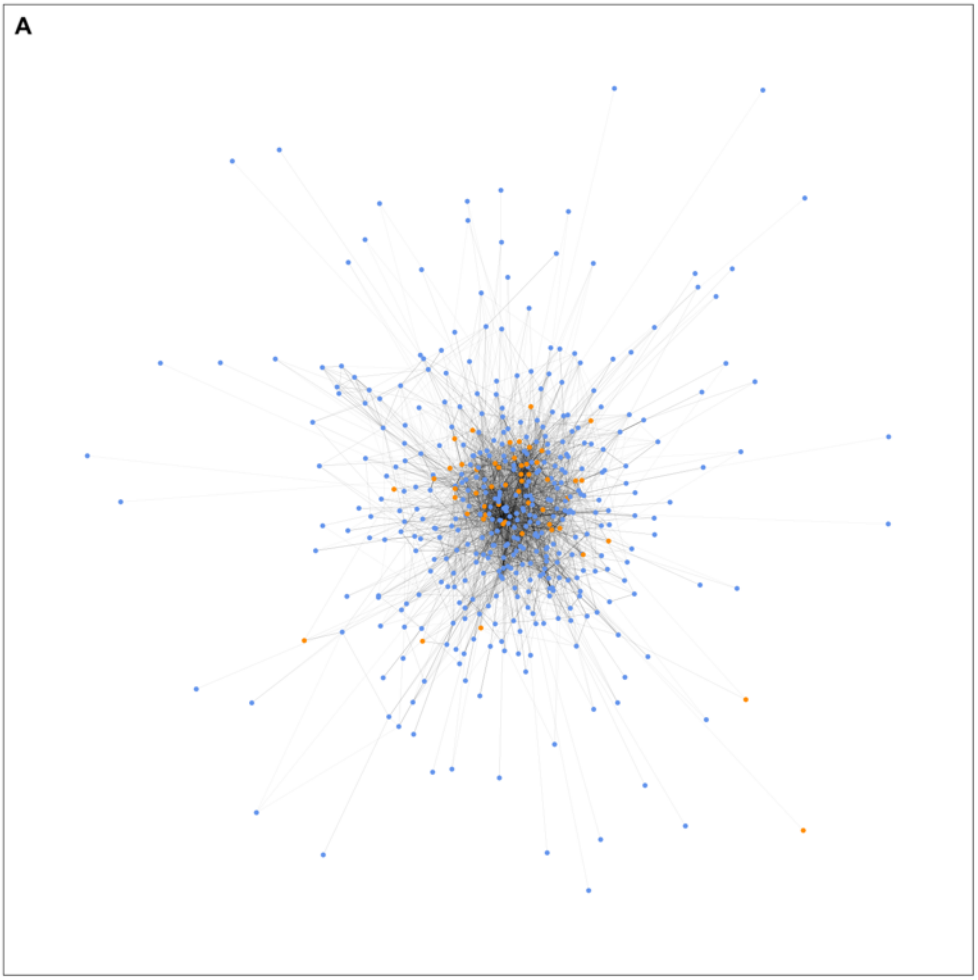
Suppl. Table 2: Health Belief Model parametrized with real-life beliefs

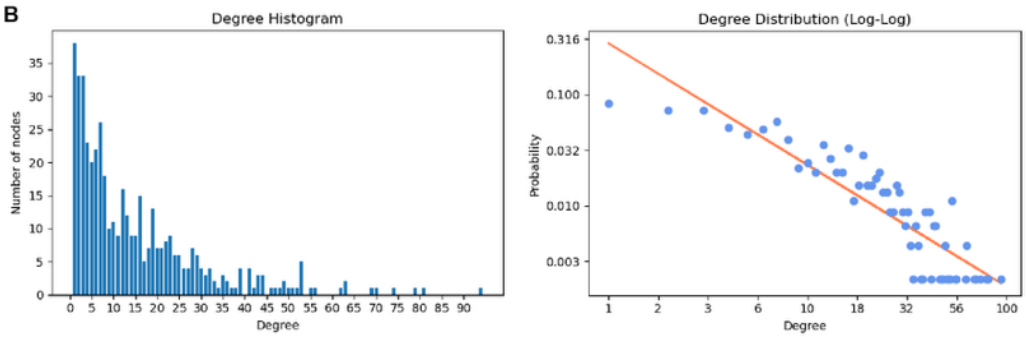
Predictor	Coefficient	Std. Err	P-value	Interpretation
Intercept	**-2.50**	0.407	**< 0.001**	Baseline probability for min. score & low cost.
Susceptibility (S2-Q1)	+0.09	0.073	0.218	Not significant in this additive model.
Severity (S2-Q2)	+0.12	0.071	0.084	Not significant in this additive model.
Self-Efficacy (S2-Q3)	-0.11	0.071	0.108	Not significant in this additive model.
Benefits (S2-Q4)	-0.04	0.096	0.663	Not significant in this additive model.
Gender (Female)	**+0.326**	0.151	**0.031**	Females significantly more likely to quarantine.
Barrier (Cost/Group)	+0.07	0.189	0.710	Not significant in this additive model.

Suppl. Figure 1: WTA results. Each pie chart represents the fraction of participants who picked the no quarantine option (red) versus quarantine (blue) for the given point gap.

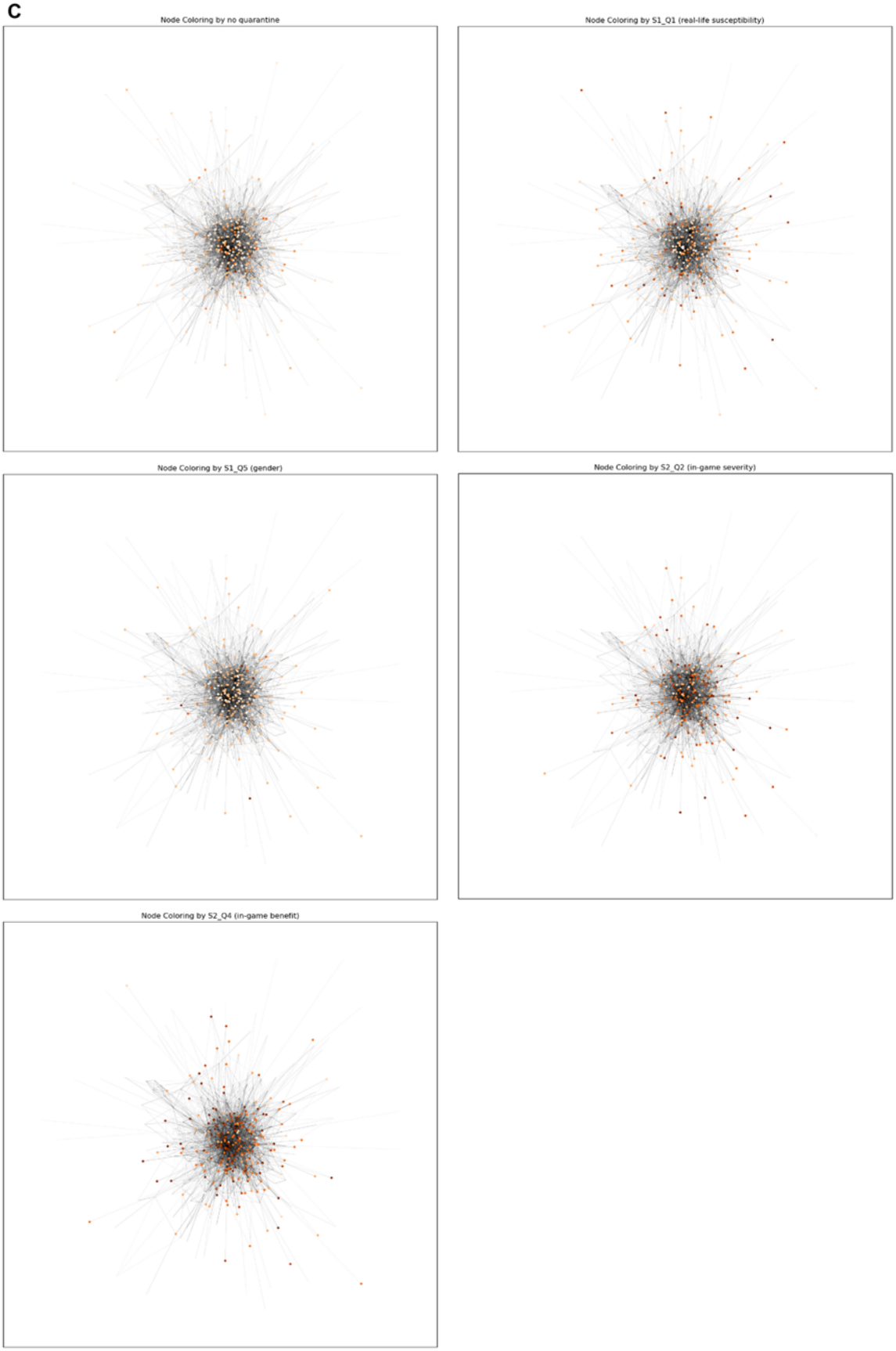


Suppl. Figure 2: Network Analysis. Panel A shows the aggregate contact network (where weighted edges were calculated by adding up the total contact time between each pair of participants). Panel B shows the degree distribution (left is histogram of the node degree, right is the probability plot in log-log scale). Panel C shows the network colored by significantly assortative variables.





C


Node Coloring by no quarantine
Node Coloring by S1_Q1 (real-life susceptibility)
Node Coloring by S1_Q5 (gender)
Node Coloring by S2_Q2 (in-game severity)
Node Coloring by S2_Q4 (in-game benefit)