Analysis protocol

# Learning network complexity through gameplay: Part B. Analysing player actions

2021

*Authors:*

Chris Moreno-Stokoe, Claire Haworth & Oliver Davis

Contents

## Background

## Introduction

## Aims

## Methods

## Data collection

## Design

## Data analysis plan

## Data cleaning

## Scoring

## Data linkage

## Planned analyses

# Background

# Introduction

Public health exists as a complex network of factors. Factors in public health share many complex and poorly understood relationships with each other. For an example of a public health network see Figure 1.

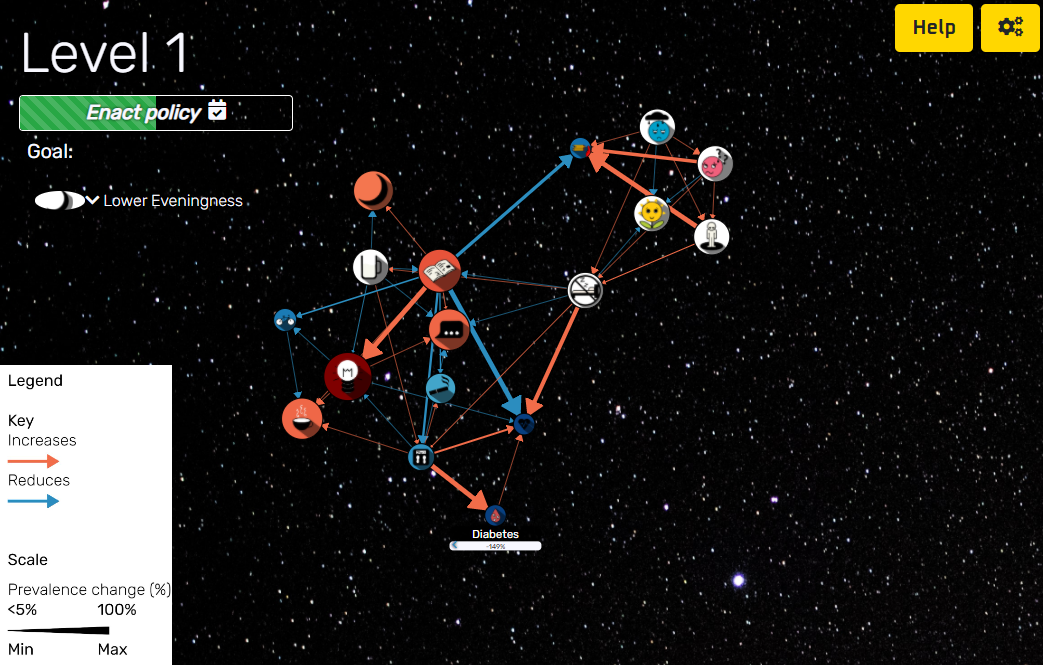


Figure 1. Screenshot of a public health network from the simulation game Mendel: <https://www.morenostok.io/mendel/game.html>

The network properties of public health complicates the design of public health intervention and policy making. The effects of interventions propagate through the public health network, for example an intervention on insomnia may have many side-effects on related traits which in turn cause side-effects of their own. This aspect of mathematical complexity makes it difficult to predict the precise effects of interventions. This network complexity reflects the real-life public health policy making environment where problems may have many possible courses of action but no single correct answer (Schrier, 2016).

Offering solutions to problems in public health requires both mathematical understanding and critical reasoning. In contrast to humans, computers excel at solving complex mathematical problems. However, while humans make decisions by reasoning computers cannot so offer impractical solutions in real-life. Computer-made solutions can be viewed as unbiased whereas humans can be viewed as biased by reasoning and individual perspectives. Effective and practical public health policies require solutions based on both critical reasoning and maths (Schrier, 2016).

Humans and computers can work together to offer solutions in *knowledge games*. A knowledge game entirely integrates a problem into a game and can allow players to offer real-life solutions (Schrier, 2016). For example, in simulation games players interact with a realistic model of a real-life phenomenon (Bilson, Bekebrede & Mayer, 2010; Guerts et al., 2007).

Investigating the solutions offered by players of a public health policy simulation game could help better understand the policy making abilities of humans and computers. *Mendel* (Chapter 4) modelled a network dataset of public health. Within this network players could explore a network dataset and encountered problems (see Fig 1). Problems were framed as objectives which players had to achieve by designing public health interventions which increase or reduce the prevalence of a factor. For example a player asked to “increase wellbeing” may design an intervention which increases wellbeing by reducing insomnia. Players were given free choice within some boundaries intended to direct players to explore network effects and offer useful solutions: players could not intervene on the objective directly (e.g., wellbeing could not be directly increased) and players could not make interventions detrimental to public health (e.g., they could not increase the prevalence of heart disease).

In *Mendel* players both played an intervention simulation game to learn about a public health network dataset and offered solutions to problems as play. Chapter 5a analysed the player experience and learning outcomes. Players’ actions were also recorded during this experiment. The present chapter analyse the recorded player data to investigate the solutions which players offered.

Furthermore, an a-priori hypothesis from Chapter 5a will be tested. Chapter 5a did not reveal significant learning improvements for players compared with an interactive visualisation control. However, an outstanding issue is whether the multiple-choice questionnaire assessment captured over-time learning. Players of the game may have learned to better design interventions through the course of their gameplay. This is due to the concrete experiences and active experimentation which knowledge games afford (Ricardi & DePaulis, 2014). Players who successfully apply information to suggest effective intervention solutions can be said to have achieved a higher understanding of public health (e.g., Bloom’s Taxonomy of Learning: Anderson & Bloom, 2001). This can similarly be investigated by analysing how the interventions that players suggested changed during the course of gameplay.

# Aims

**Research question**

What can be learned from analysing players’ actions in a public health intervention simulation game?

**Objectives**

Player actions will be investigated to test the following hypotheses:

* Players will learn to design more effective interventions over-time
* Humans will make different interventions to computers

Additionally, exploratory analyses are planned to describe:

* The differences between human and computer interventions
* The network properties which contribute to complexity and decision making in the simulation

# Methods

# Data collection

**Human-made intervention suggestions**

In-game player actions recorded as part of Chapter 5a will be analysed. 90 participants were assigned to the game condition and made a total of 2255 interventions throughout gameplay. This data was collected anonymously so participant information is not available for joining. Data collection was automated. On making an intervention in the JavaScript game client an XML HTTP request sent player data to a database. For storing data on the server side a PHP script issued pre-prepared SQL statements. Once stored, data was ready in CSV format.

**Computer-made intervention suggestions**

A Network MR algorithm was developed in Chapter 4 to provide an AI opponent for players to compete against in *Mendel*. The AI calculates the effects of every possible intervention for a given objective and returns the mathematically optimal solution. This model is deterministic meaning that the suggested solution to a given problem is always the same. The “optimal solution” in this context is one which achieves the objective the most, regardless of any other factors. For example, the optimal solution to increases wellbeing is the intervention which most increases the prevalence of wellbeing. This AI can be engineered to provide an exhaustive list of its suggestions for every objective players were presented with in the game. This AI was developed using JavaScript and a local script can generate CSV outputs ready for analysis in the same way as player data.

# Design

The different hypotheses require different designs to answer:

* H1 (Players will learn to better design interventions over-time) will measure players’ intervention making over the course of gameplay using a within-participants design
* H2 (Humans will make different interventions to computers) will be investigated by comparing human and computer-made interventions

Exploratory analyses will also compare human and computer made interventions.

# Analysis plan

# Data cleaning

Pilot data will be removed by removing entries from before the study was launched to participants (16th December). Since player data is anonymous individuals cannot be excluded on an individual basis in the same way as in Chapter 5a.

# Scoring

As in Chapter 5a, players’ interventions received scores expressed as a percentage of the optimal intervention. Interventions received both an objective score and side-effect score.

# Data joining

Data from participants will be enriched with properties and network characteristics. This will provide more information to better describe and analyse players’ actions. For each player action a few categories of information are recorded. These are joined into a single table but are separated below for ease of explanation.

Due to the anonymous nature of data collection this data cannot be linked back to participant characteristics or questionnaire data collected in Chapter 5a.

**Information on the player**

A unique user ID (session ID) is recorded for each intervention so that player responses within the same session can be identified for within-subject investigation. This does not uniquely identify the individual but rather their play session, meaning that if they closed and re-opened the game they would be recorded as two different players. A time stamp for when the intervention was made was also recorded. See Table 1 for a blank table of this information.

**Table 1**. Player information

|  |  |
| --- | --- |
| **Unique User ID** | **Timestamp** |
|  |  |

**Information on the current problem**

Information on the problem the player was trying to solve with an intervention was also recorded. The objective that they player was given to achieve is recorded. For example, the objective “reduce depression” would indicate players were attempting to suggest an intervention with the goal of reducing depression. Furthermore, players were sometimes able to make multiple simultaneous interventions in order to achieve the objective and information was provided to identify interventions where this was the case. See Table 2 for a blank table of this information.

**Table 2**. Problem characteristics

|  |  |  |
| --- | --- | --- |
| **Objective** | **Maximum number of permitted interventions** | **Current number of interventions** |
|  |  |  |

**Information on the trait the player chose to intervene on**

To help better understand why a player may chose to intervene on a specific trait rather than another trait, information on traits was joined. The name of the trait was joined with its valence. Valence denotes whether a trait has a good, bad or neutral effect on health. For example, heart disease was considered bad. This would help understand if players tended to intervene on good, bad or neutral traits more than the computer. Definitions of trait valence is consistent with Chapter 5a. Additionally, players may chose traits which appear most central or best connected within the network. In order to understand this, graph information on the number of neighbours (both predecessors and successors) and centrality (eigenvalue and betweenness) of each node in the network was calculated. See Table 3 for a blank table of this information.

**Table 3**. Trait characteristics

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Trait intervened on** | **Valence** | **N Predecessors** | **N Successors** |  | **Centrality (eigenvalue)** |  |  |
|  |  |  |  |  |  |  |  |

**Information on the effects of the intervention the player chose**

In order to measure the mathematical effectiveness of interventions these were scored in the same way as Chapter 5a. Objective and side scores were calculated for each intervention. In order to understand whether players took into account the raw number of effects an intervention would have, the total number of effects was joined. In order to understand if players considered how close a trait is to the objective, the length of the shortest path between every node in the network was joined. Additionally, the effect of every intervention on every individual trait is available for use in analysis. See Table 4 for a blank table of this information.

**Table 4**. Intervention characteristics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Trait intervened on** | **Score (objective)** | **Score (side-effect)** | **Total n effects** | **Effects (on each traits)** | **Paths to each trait (length)** |
|  |  |  |  |  |  |

# Analysis

**Hypothesis testing**

While the accepted alpha level will be 0.05, this will not be used as a strict requirement, and instead this will be evaluated in the context of the magnitude of effect size as well (this following IEU guidelines on significance thresholds). A hypothesis will be rejected if all its tests return non-significant results by this criteria. The hypotheses will be tested as follows:

1. H1: Players will learn to design more effective interventions over-time
   * Within each participant the time order of interventions will be correlated against the score these interventions received (objective score & side-score). For example, if a player’s 1st intervention scores 20 and their 2nd scores 40, this would indicate learning
2. H2: Humans will make different interventions to computers
   * Human and computer-made interventions will be t-tested for differences in:
     + - Objective score (effect on objective trait)
       - Side-effect score (goodness of side effects)
       - Information on the trait (valence, neighbours, centrality)
       - Information on the intervention’s effects (score, path length to objective)

**Exploratory analysis**

Although exploratory analyses are likely to change and expand through the course of analysis, two analyses are planned so-far.

Human and computer-made interventions will be described and the most popular interventions identified. To this end information for three categories (the current problem, target trait, and intervention’s effects) will be compared between human and computer-made interventions.

Matrix correlation between network properties and intervention scores will identify any network properties implicated in making interventions more effective or which contribute to complexity and player confusion.

# References

Anderson, L. W., Bloom, B. S., & others. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives*. Longman,.

Geurts, J. L. A., Duke, R. D., & Vermeulen, P. A. M. (2007). Policy Gaming for Strategy and Change. *Long Range Planning*, *40*(6), 535–558. https://doi.org/10.1016/j.lrp.2007.07.004

Ricciardi, F., & De Paolis, L. T. (2014). A Comprehensive Review of Serious Games in Health Professions. *International Journal of Computer Games Technology*, *2014*. <https://doi.org/10.1155/2014/787968>

Schrier, K. (2016). *Knowledge games: How playing games can solve problems, create insight, and make change*. JHU Press.

van Bilsen, A., Bekebrede, G., & Mayer, I. (2010). Understanding complex adaptive systems by playing games. *Informatics in Education*, *9*(1), 1–18. https://doi.org/10.15388/infedu.2010.01