

**Replicability Instructions for Data Analysis:**  
**"Fatal Attraction: Salience, Naiveté, and Sophistication in Experimental 'Hide-and-Seek' Games"**  
**by Vincent P. Crawford and Nagore Iriberry**

This document provides the data from the three groups of experiments analyzed in the paper, as well as instructions for replicating the analysis. First, we provide the data sets and then we provide Matlab files for replicating the maximum likelihood estimates.

**A. Data Sets**

The paper uses three data sets: Rubinstein, Tversky and Heller's (RTH), O'Neill's, and Rapoport and Boebel's. RTH's data are used to estimate the alternative models proposed in the paper. O'Neill's and Rapoport and Boebel's data are used only to test for portability, hence only to calculate MSEs for different models given the parameters estimated using RTH's data.

O'Neill's and Rapoport and Boebel's data, which have not been previously published (though they have been made freely available to researchers on request), are published here with their permission. We gratefully acknowledge their generosity.

RTH's aggregate choice frequency data, which is all that is used in our paper, were published in their cited papers. As explained in the web appendix, we made minor adjustments to their published data to reconcile reported frequencies and sample sizes.

**RTH's Data:**

Their data are provided in the paper, disaggregated for all six treatments in Table 1, and aggregated over treatments in the first two lines of Table 3.

**O'Neill's Data:**

O'Neill's entire dataset is provided in the document O'NeillRawData.doc. The data are presented in a matrix. The columns refer to the matched pairs' IDs (1-25) and the rows refer to periods (1-105). Each entry in the matrix consists of a pair of actions referring to the action of Player 1 and Player 2 respectively.

With one exception, our analysis uses only O'Neill's subjects' initial responses, which are recorded in the first 2 rows in Table 5 in the paper. The exception is our footnote 24, which depends on his subjects' time-aggregated responses as well. The time-aggregated responses are published in O'Neill's paper, but they also can be extracted from his entire dataset.

**Rapoport and Boebel's Data:**

These data are provided in the document RapoportBoebelRawData.doc. The variables in each column starting from left to right refer to:

Experiment: 1 and 2 (Footnote 22 in the paper or Table 2 on page 265 in Rapoport and Boebel).

Sessions: 1-20 refer to Experiment 1 and 21-40 to Experiment 2. Session 1 (1-10) and Session 2 (11-20), Session 1 (21-30) and Session 2 (31-40).

Period: 1-120.

Player X (1 in the paper): 1-40.

Player Y (2 in the paper): 1-40.

Action for Player X: 1-5.

Action for Player Y: 1-5.

Since we are only interested in initial responses we extracted Period 1's data from Session 1. The initial responses are provided in the first 4 rows in Table 6 in the paper.

## **B. Data Analysis:**

We computed maximum likelihood estimations for RTH's data set, both for the aggregated data reported in Table 3 and for each of the six treatments reported in Table A2 in the Appendix. We provide m-files from Matlab that we used for the maximum likelihood estimation.

LiH.m and LiS.m: for  $i=0,1,2,3,4$ . Given input  $p$  and  $q$  (choice probabilities for end points and focally labelled B) the output yielded are the choice probabilities for different levels (L0, L1, L2, L3, L4) and different player roles (H for Hiders and S for Seekers).

LogLikeHS.m: it defines the log likelihood function for the alternative models of Level-k (with symmetric and asymmetric L0) and equilibrium with perturbations. The inputs are the parameters to be estimated for each model and the output is the log likelihood function.

minHS.m: it uses the fmincon routine to find the maximum likelihood estimates.