Safe models for risky decisions

Steingröver, H.M.

Appendix B

Data from 617 Healthy Participants Performing the Iowa Gambling Task: A "Many Labs" Collaboration

This chapter has been published as:
Helen Steingroever, Hasker Davis, Daniel J. Fridberg, Annette Horstmann, Kimberly L. Kjome, Veena Kumari, Scott D. Lane, Tiago V. Maia, James L. McClelland, Thorsten Pachur, Preethi Premkumar, Julie Stout, Ruud Wetzels, Stacey Wood, Darrell A. Worth, and Eric-Jan Wagenmakers (2015).
Data from 617 healthy participants performing the Iowa gambling task: A "many labs" collaboration.
Journal of Open Psychology Data, 3:e5.[1]

Abstract

This data pool (N = 617) comes from 10 studies assessing performance of healthy participants (i.e., no known neurological impairments) on the Iowa gambling task (IGT)—a task measuring decision making under uncertainty in an experimental context. Participants completed a computerized version of the IGT consisting of 95 – 150 trials. The data consist of the choices of each participant on each trial, and the resulting rewards and losses. The data are stored as .rdata, .csv, and .txt files, and can be reused to (1) analyze IGT performance of healthy participants; (2) create a "super control group"; or (3) facilitate model-comparison efforts.

[1]The final publication is available at `http://openpsychologydata.metajnl.com/article/10.5334/jopd.ak/`.

B.1 Overview

Context

Collection Date(s)

2000 - 2013

Background

This data pool comes from eight independent published studies (Fridberg et al., 2010; Horstmann; Kjome et al., 2010; Maia & McClelland, 2004; Premkumar et al., 2008; Wetzels, Vandekerckhove, et al., 2010; Wood et al., 2005; Worthy, Pang, & Byrne, 2013), one submitted study (Steingroever et al., submitted), and one unpublished study (Steingroever, Wetzels, & Wagenmakers, 2011). These studies report the performance of a total of 617 healthy participants on the Iowa gambling task (IGT; Bechara et al., 1994). The IGT is arguably the most popular neuropsychological paradigm to measure decision-making deficits in an experimental context. Part of the data was already reanalyzed elsewhere (i.e., Steingroever et al., submitted; Steingroever, Wetzels, Horstmann, et al., 2013; Steingroever, Wetzels, & Wagenmakers, 2013a, 2013b; Steingroever et al., 2014, 2016) in order to assess basic assumptions underlying the performance of healthy participants on the IGT, and to compare reinforcement-learning models that try to disentangle psychological processes underlying performance on the IGT.

B.2 Methods

Sample

Table B.1 describes the data pool. All included studies used (a variant of) the traditional IGT payoff scheme (Bechara et al., 1994) or the payoff scheme introduced by Bechara and Damasio (2002). A detailed description of the payoff schemes can be found in the Supporting Text 1.

In the traditional payoff scheme, the net outcome of 10 cards from the bad decks (i.e., decks A and B) is -250, and +250 in the case of the good decks (i.e., decks C and D). In addition, there are two decks with frequent losses (decks A and C), and two decks with infrequent losses (decks B and D). In the traditional payoff scheme, there is a variable loss in deck C (i.e., either -25, -50, or -75; classified here as payoff scheme 1). However, some of the included studies used a variant of this payoff scheme in which the loss in deck C was held constant (i.e., -50; classified here as payoff scheme 2). A second difference between payoff scheme 1 and 2 is that payoff scheme 1 uses a fixed sequence of rewards and losses, whereas payoff scheme 2 uses a randomly shuffled sequence.

The payoff scheme introduced by Bechara and Damasio, 2002; classified here as payoff scheme 3) also consists of two good decks (decks C and D), and two bad decks (decks A and B), and two decks with frequent losses (decks A and C) and two decks with infrequent losses (decks B and D). However, in contrast to payoff schemes 1 and 2, the schedules of rewards and losses in payoff scheme 3 are structured in such a way that the discrepancy between rewards and losses in the bad decks (decks A and B) changes such that the net outcome decreases by 150 every block of 10 cards (i.e., in the first block, the net outcome is -250, but in the sixth block, it is -1000). By contrast, the net outcome of the good decks (decks C and D) increases by 25 every block of 10 cards (i.e., in the first block, the net outcome is 250, but in the sixth block, it is 375). Thus, the good decks become gradually better, whereas the bad decks become gradually worse. In addition, in contrast

Table B.1: *Overview of the studies included in the data pool. See text for a description of the different payoff schemes.*

Study	Number of participants	Number of trials	Payoff	Demographics[a]
Fridberg et al. (2010)	15	95	1	M = 29.6 years (SD = 7.6)
Horstmann[b]	162	100	2	M = 25.6 years (SD = 4.9), 82 female
Kjome et al. (2010)	19	100	3	M = 33.9 years (SD = 11.2), 6 female
Maia and McClelland (2004)	40	100	1	Undergraduate students
Premkumar et al. (2008)	25	100	3	M = 35.4 years (SD = 11.9), 9 female
Steingroever et al. (submitted)	70	100	2	M = 24.9 years (SD = 5.8), 49 female
Steingroever et al. (2011)	57	150	2	M = 19.9 years (SD = 2.7), 42 female
Wetzels, Vandekerckhove, et al. (2010)[c]	41	150	2	Students
Wood et al. (2005)	153	100	3	M = 45.25 years (SD = 27.21)[d]
Worthy, Pang, and Byrne (2013)	35	100	1	Undergraduate students, 22 female

[a] Information that was provided in the original articles. This information consists of the mean age and the standard deviation in brackets, or alternatively the occupation of the participants. In addition, the number of female participants is provided for most datasets.
[b] Data collected by Annette Horstmann. These data were first published in Steingroever, Wetzels, Horstmann, et al., 2013. A subset of this dataset is published in Horstmann, Villringer, and Neumann (2012).
[c] Data of the standard condition. Data of three other conditions can be downloaded here: `http://www.ruudwetzels.com/data/EV$_$data.zip`.
[d] The first 90 participants of this dataset are between 18-40 years old (M = 23.04, SD = 5.88), and participants 91-153 are between 61 and 88 years old (M = 76.98, SD = 5.20).

to payoff schemes 1 and 2, the wins differ within each deck in payoff scheme 3. Just as payoff scheme 1, payoff scheme 3 uses a fixed sequence of wins and losses.

B.3 Dataset Description

Object name

IGTdataSteingroever2014.zip. This zip archive contains the following files:

- IGTdata.rdata
- choice_95.csv, choice_100.csv, choice_150.csv, wi_95.csv, wi_100.csv, wi_150.csv, lo_95.csv, lo_100.csv, lo_150.csv, index_95.csv, index_100.csv, index_150.csv.
- choice_95.txt, choice_100.txt, choice_150.txt, wi_95.txt, wi_100.txt, wi_150.txt, lo_95.txt, lo_100.txt, lo_150.txt, index_95.txt, index_100.txt, index_150.txt.

Data Type

Processed data

Format names and versions

The data are provided in three different formats: .rdata (R), .csv (Excel), and .txt. The .rdata file is called “IGTdata.rdata” and it contains the following 12 matrices:

- choice_95, choice_100, and choice_150: These matrices contain the choices of all studies that used a 95-trial, 100-trial, and 150-trial IGT, respectively. The dimension of each matrix corresponds to the number of subjects x number of trials. For example, choice_95 (Figure B.1) is a 15 x 95 matrix, and the entry of the third row and fifth column corresponds to the choice that the third participant made on the fifth trial. The entries of the three choice matrices are either 1, 2, 3, or 4, where 1, 2, 3, and 4 stand for deck A, B, C, and D, respectively.

- wi_95, wi_100, and wi_150: These matrices contain the rewards of all studies that used a 95-trial, 100-trial, and 150-trial IGT, respectively. The dimension of each matrix corresponds to the number of subjects x number of trials. For example, wi_100 is a 504 x 100 matrix, and the entry of the third row and fifth column corresponds to the reward that the third participant received on the fifth trial. The entries of the three reward matrices vary between 40 and 170.

- lo_95, lo_100, and lo_150: These matrices contain the losses of all studies that used a 95-trial, 100-trial, and 150-trial IGT, respectively. The dimension of each matrix corresponds to the number of subjects x number of trials. For example, lo_150 is a 98 x 150 matrix, and the entry of the third row and fifth column corresponds to the loss that the third participant received on the fifth trial. The entries of the three loss matrices vary between -2500 and 0. Thus, the losses are saved as negative numbers.

- index_95, index_100, and index_150: These matrices contain the name of the first author of the study that reports the data of the corresponding participant. For example, the third entry of index_95 (Figure B.2) can be used to identify who collected the choices saved in the third row of choice_95, wi_95, and lo_95.

These 12 matrices are saved altogether in the “IGTdata.rdata” file. In addition, we saved the 12 matrices as separate .csv and.txt files. For example, the matrix choice_95 (Figure B.1) is contained in the “IGTdata.rdata” file, but also in “choice_95.csv” and “choice_95.txt”. The file names of the .csv and .txt files indicate which matrix they contain.

Data Collectors

Language

N/A

License

Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

```
> choice_95
        Choice_1 Choice_2 Choice_3 Choice_4 Choice_5 Choice_6
Subj_1         2        2        2        2        2        2
Subj_2         1        2        3        2        2        2
Subj_3         3        4        3        2        2        1
Subj_4         4        3        1        1        1        2
Subj_5         1        2        3        4        3        1
Subj_6         1        2        1        2        1        2
Subj_7         1        2        3        4        4        3
Subj_8         4        2        1        3        2        3
Subj_9         1        2        1        2        3        4
Subj_10        4        2        2        1        3        1
```

Figure B.1: Screenshot of a subset of the choice_95 matrix. Each row contains the sequence of choices from a specific participant. For example, the entry of the third row and fifth column corresponds to the choice that the third participant made on the fifth trial (i.e., “2” – deck B). To determine who collected the data of this particular participant, one needs to refer to the third row of index_95 (cf. Figure B.2).

```
> index_95
      Subj Study
 [1,] "1"  "Fridberg"
 [2,] "2"  "Fridberg"
 [3,] "3"  "Fridberg"
 [4,] "4"  "Fridberg"
 [5,] "5"  "Fridberg"
 [6,] "6"  "Fridberg"
 [7,] "7"  "Fridberg"
 [8,] "8"  "Fridberg"
 [9,] "9"  "Fridberg"
[10,] "10" "Fridberg"
```

Figure B.2: Screenshot of a subset of the index_95 matrix. Each row can be used to identify who collected the data of a specific participant. The screenshot shows that the data of subjects $1 - 10$ who completed a 95-trial IGT were collected by Fridberg et al. (2010). The data of these subjects are saved in the corresponding rows of the choice_95, reward_95, and loss_95 matrices.

Embargo

N/A

Repository location

`https://osf.io/8t7rm`

Publication date

05/11/2014

B.4 Reuse Potential

This data pool has several reuse potentials: First, it could be used to more thoroughly investigate healthy participants' performance on the IGT. Second, it could be reused as a "super control group". This means that performance of an experimental group can be assessed relative to the performance of healthy participants included in this data pool. Third, the data pool could be reused to compare computational models for the IGT. However, it should be noted that the 10 datasets were collected in different environments, and that the performance of the participants on the IGT may possibly be affected by factors that varied across the included studies (e.g., the use and type of incentives, questions about the IGT during the performance to asses participants' awareness, randomly shuffled payoff or fixed payoff sequence, the type of task instruction).

Acknowledgements

We would like to thank Kala Battistelli, Andreas Below, Katie Chamberlain, Courtni East, Noémie J. Eichhorn, Lindsey Ferris, Monica Gamboa, Kaitlynn Goldman, Karolin Gohlke, Karla Gomez, Alexis Gregg, Jordan Hall, Christy Ho, Jonas Klinkenberg, Lauren Laserna, Katja Macher, Samantha Mallec, Megan McDermott, Ramona Menger, Gerard Moeller, Michael Pang, Anthony Schmidt, Candice Tharp, Lucas Weatherall, Michael Wesley, and Christopher Whitlow for their help in collecting the data.

Funding Statement

This work was supported by BMBF [IFB Adiposity Diseases (FKZ:01EO1001, ADI-K50F, ADI-K753, ADI-K754) to Annette Horstmann], the DFG (Mind and Brain; CRC 1052 "Obesity Mechanisms" A5 to Annette Horstmann), the Einstein-Stiftung (Mind and Brain Institute), a fellowship from the Foundation for Science and Technology of Portugal (to Tiago V. Maia) and by National Institute of Mental Health Grant P50-MH64445 (to James L. McClelland), funds from the Wellcome Trust, UK (067427/z/02/z), the National Institute on Drug Abuse (NIDA; PI F.G. Moeller, contract/grant number: K02DA00403, P50DA009262), and CCTS/CRU grant number UL1 RR024148.

B.5 Supporting Text 1

This document provides details on the three different payoff schemes employed by the included studies. In addition, it explains how the data can be read in R, and presents an example of an analysis.

Payoff Schemes

Table B.2: *Rewards and losses from 40 choices of each deck as used in the traditional payoff scheme with variable loss in deck C (see Bechara et al., 1994; classified here as payoff scheme 1). Within each deck, the presented payoff sequence is repeated after participants have made 40 choices from the corresponding deck.*

Trial	win A	win B	win C	win D	loss A	loss B	loss C	loss D
1	100	100	50	50	0	0	0	0
2	100	100	50	50	0	0	0	0
3	100	100	50	50	-150	0	-50	0
4	100	100	50	50	0	0	0	0
5	100	100	50	50	-300	0	-50	0
6	100	100	50	50	0	0	0	0
7	100	100	50	50	-200	0	-50	0
8	100	100	50	50	0	0	0	0
9	100	100	50	50	-250	-1250	-50	0
10	100	100	50	50	-350	0	-50	-250
11	100	100	50	50	0	0	0	0
12	100	100	50	50	-350	0	-25	0
13	100	100	50	50	0	0	-75	0
14	100	100	50	50	-250	-1250	0	0
15	100	100	50	50	-200	0	0	0
16	100	100	50	50	0	0	0	0
17	100	100	50	50	-300	0	-25	0
18	100	100	50	50	-150	0	-75	0
19	100	100	50	50	0	0	0	0
20	100	100	50	50	0	0	-50	-250
21	100	100	50	50	0	-1250	0	0
22	100	100	50	50	-300	0	0	0
23	100	100	50	50	0	0	0	0
24	100	100	50	50	-350	0	-50	0
25	100	100	50	50	0	0	-25	0
26	100	100	50	50	-200	0	-50	0
27	100	100	50	50	-250	0	0	0
28	100	100	50	50	-150	0	0	0
29	100	100	50	50	0	0	-75	-250
30	100	100	50	50	0	0	-50	0

Continued on next page

Table B.2 – *Continued from previous page*

Trial	**win A**	**win B**	**win C**	**win D**	**loss A**	**loss B**	**loss C**	**loss D**
31	100	100	50	50	-350	0	0	0
32	100	100	50	50	-200	-1250	0	0
33	100	100	50	50	-250	0	0	0
34	100	100	50	50	0	0	-25	0
35	100	100	50	50	0	0	-25	-250
36	100	100	50	50	0	0	0	0
37	100	100	50	50	-150	0	-75	0
38	100	100	50	50	-300	0	0	0
39	100	100	50	50	0	0	-50	0
40	100	100	50	50	0	0	-75	0

Table B.3: *A possible sequence of rewards and losses from 10 choices of each deck based on the traditional payoff scheme with constant loss in deck C (see Bechara et al., 1994; classified here as payoff scheme 2). A payoff sequence with the presented characteristics is randomly generated for each block of 10 trials.*

Trial	win A	win B	win C	win D	loss A	loss B	loss C	loss D
1	100	100	50	50	0	0	0	0
2	100	100	50	50	-300	0	0	0
3	100	100	50	50	-150	0	-50	0
4	100	100	50	50	0	0	0	0
5	100	100	50	50	-350	0	-50	0
6	100	100	50	50	0	-1250	0	0
7	100	100	50	50	0	0	-50	0
8	100	100	50	50	-250	0	0	0
9	100	100	50	50	0	0	-50	-250
10	100	100	50	50	-200	0	-50	0

Tables B.2 – B.4 illustrate the payoff schemes classified as payoff scheme 1, 2, and 3, respectively. The difference between payoff scheme 1 and 2 is that the first one uses variable losses in deck C and has a fixed sequence of wins and losses (Table B.2), whereas the second one uses a constant loss in deck C and, within each deck, a random repetition of the payoffs presented in Table B.3. The third payoff scheme changes rewards and losses in such a way that the net outcome of the bad decks becomes increasingly negative every 10 trials, whereas the net outcome of the good decks becomes increasingly positive. Thus, the difference in the net outcomes of the good and bad decks increases every 10 trials. In addition, in contrast to payoff schemes 1 and 2, the wins differ within each deck in payoff scheme 3. Just as payoff scheme 1, payoff scheme 3 uses a fixed sequence of wins and losses.

Table B.4: *Rewards and losses from 60 choices of each deck as used in the payoff scheme introduced by Bechara and Damasio, 2002; classified here as payoff scheme 3).*

Trial	**win A**	**win B**	**win C**	**win D**	**loss A**	**loss B**	**loss C**	**loss D**
1	100	100	50	50	0	0	0	0

Continued on next page

Table B.4 – *Continued from previous page*

Trial	win A	win B	win C	win D	loss A	loss B	loss C	loss D
2	120	80	60	40	0	0	0	0
3	80	110	40	45	-150		-50	0
4	90	120	55	45	0	0	0	0
5	110	90	55	55	-300	0	-50	0
6	100	100	45	60	0	0	0	0
7	80	90	50	40	-200	0	-50	0
8	120	120	45	55	0	0	0	0
9	110	110	60	50	-250	-1250	-50	0
10	90	80	40	60	-350	0	-50	-250
11	110	110	55	55	0	0	0	0
12	130	100	55	40	-350	0	-25	0
13	90	90	65	60	0	0	-75	0
14	100	130	45	40	-250	-1500	0	0
15	120	120	70	45	-200	0	-25	0
16	110	130	40	55	0	0	0	0
17	90	110	50	65	-300	0	-25	0
18	130	90	60	70	-150	0	-75	0
19	120	100	70	50	-250	0	0	0
20	100	120	40	70	0	0	-50	-275
21	120	120	60	60	-250	-1750	0	0
22	140	110	65	55	-300	0	-25	0
23	110	140	55	65	0	0	0	0
24	110	130	80	80	-350	0	-50	0
25	100	100	40	40	0	0	-25	0
26	120	110	60	80	-200	0	-50	0
27	130	120	55	40	-250	0	0	0
28	110	120	65	65	-150	0	-25	0
29	140	140	40	55	-250	0	-75	-300
30	120	110	80	60	0	0	-50	0
31	130	130	65	65	-350	0	-25	0
32	120	140	75	75	-200	-2000	0	0
33	140	120	55	60	-250	0	-25	0
34	130	110	60	65	-250	0	-25	0
35	110	130	70	75	-150	0	-25	-325
36	150	150	65	85	0	0	0	0
37	140	110	55	45	-150	0	-75	0
38	120	150	75	55	-300	0	-25	0
39	150	120	45	70	-350	0	-50	0
40	110	140	85	55	0	0	-75	0
41	140	140	70	70	-350	0	-25	0
42	130	150	80	80	-200	0	0	0

Continued on next page

Table B.4 – *Continued from previous page*

Trial	win A	win B	win C	win D	loss A	loss B	loss C	loss D
43	150	130	60	65	-250	0	-25	0
44	140	120	65	70	-250	0	-25	0
45	120	140	75	80	-150	0	-25	-350
46	160	160	70	90	0	-2250	-25	0
47	150	120	60	50	-150	0	-75	0
48	130	160	80	60	-300	0	-25	0
49	160	130	50	75	-350	0	-50	0
50	120	150	90	60	-250	0	-75	0
51	150	150	75	75	-350	0	-25	0
52	140	160	85	85	-200	0	-25	0
53	160	140	65	70	-250	0	-25	0
54	150	130	70	75	-250	0	-25	0
55	130	150	80	85	-150	0	-25	0
56	170	170	75	95	-250	0	-25	0
57	160	130	65	55	-150	0	-75	0
58	140	170	85	65	-300	-2500	-25	-375
59	170	140	55	80	-350	0	-50	0
60	130	160	95	65	-250	0	-75	0