Neural Foundations of Preferences

Ernst Fehr

Department of Economics

Laboratory for Social and Neural Systems Research

University of Zurich

Outline of presentation

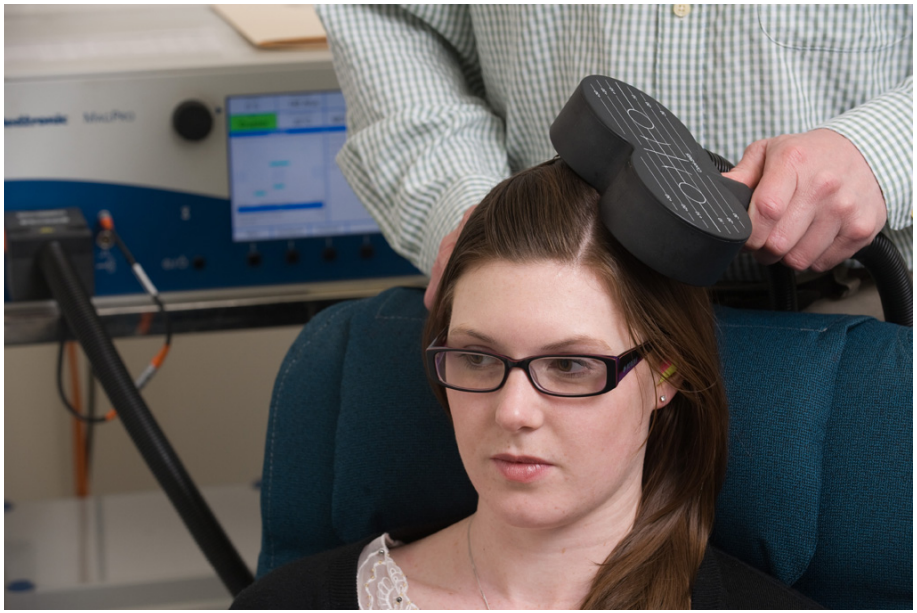
- Neural activity causes behavior
- The neuronal basis of decision values
- Neural model of choice behavior – the drift diffusion model
- Explaining other regarding choices with the drift diffusion model

Neural activity causes our behavior

- Targeted experimental manipulation of brain activity changes specific human behaviors
- Economic behavior is no exemption from this rule
- How can this be proven?

Non-invasive brain stimulation enables targeted changes in neural activations and connectivity

- Transcranial magnetic stimulation



- Transcranial direct current stimulation

Transcranial direct current stimulation in our laboratory

- Cathodal stimulation reduces neuronal excitability
- Anodal stimulation increases neuronal excitability

Neuronal activity has causal effects on important economic behaviors

- Reduction of neuronal activity in lateral PFC
  - Changes WTP for simple consumer goods (Camus et al. 2008)
  - Renders people more impatient (Figner et al. 2010)
  - Reduces rejections of low offers in UG (Knoch et al. 2006)
  - Reduces sanction-induced norm compliance (Knoch et al. 2009, Ruff et al. 2011)
  - But increases voluntary norm compliance (Ruff et al. 2011)
  - Renders people more selfish (Knoch et al. 2006)

- A thorough understanding of economic behaviors also requires knowledge about the neural mechanisms and computations that drive economic behaviors

- Economic and non-economic events that change the relevant neural activity also change the associated behaviors

Why did economists shy away from examining the neural foundations of choice?

- Technically infeasible
- Theory of revealed preferences
  - allowed doing economics without knowledge about the psychological and biological Foundations of behavior

Theory of revealed preferences

- If human choice satisfies certain plausible axioms we can represent it by the maximization of some utility function
- People behave **as if** they maximize a utility function
- Utility functions only **represent** behavior
- The underlying neuronal and psychological processes are completely irrelevant
- **It is a «reduced form theory» and not a «structural» model of choice**

Implications of revealed preference theory

(1) By assumption the best alternative is chosen – there are **no decision errors**

(2) Preferences or utilities, resp., are not a **cause** for our behaviors

- An individual does not choose a risky alternative because she has a preference for risk – she has such preferences because she chose the risky alternative
- Preferences and utilities are merely representations of behavior
- **They have no existence independently of behavior**

(3) The subjective value of a good only reveals itself through choice

- Only behavioral data have relevance for welfare judgements
- There is no subjective value of goods outside the behavioral domain
- In principle, subjective value is a superfluous category

Neuroeconomic approach is fundamentally different

- Replacement of the **«as if»** approach by the **«as is»** aproach
  - Implies the study of the **actual** processes that are the basis of choice
- The hypothesis that humans behave as if they maximize a utility function can only systematically predict behavior correctly if the underlying processes indeed maximize a utility function
- However, like in revealed preference theory no conscious maximization is needed

Friedman's Example

- The leaves of trees «behave» **as if** they maximize the receipt of solar radiation
  - Provides plausible predictions for the behavior of leaves

- This hypothesis can only provide systematically correct predictions if the underlying biological processes indeed imply that the receipt of solar radiation is maximized

Hypotheses and insights from neuroeconomic research

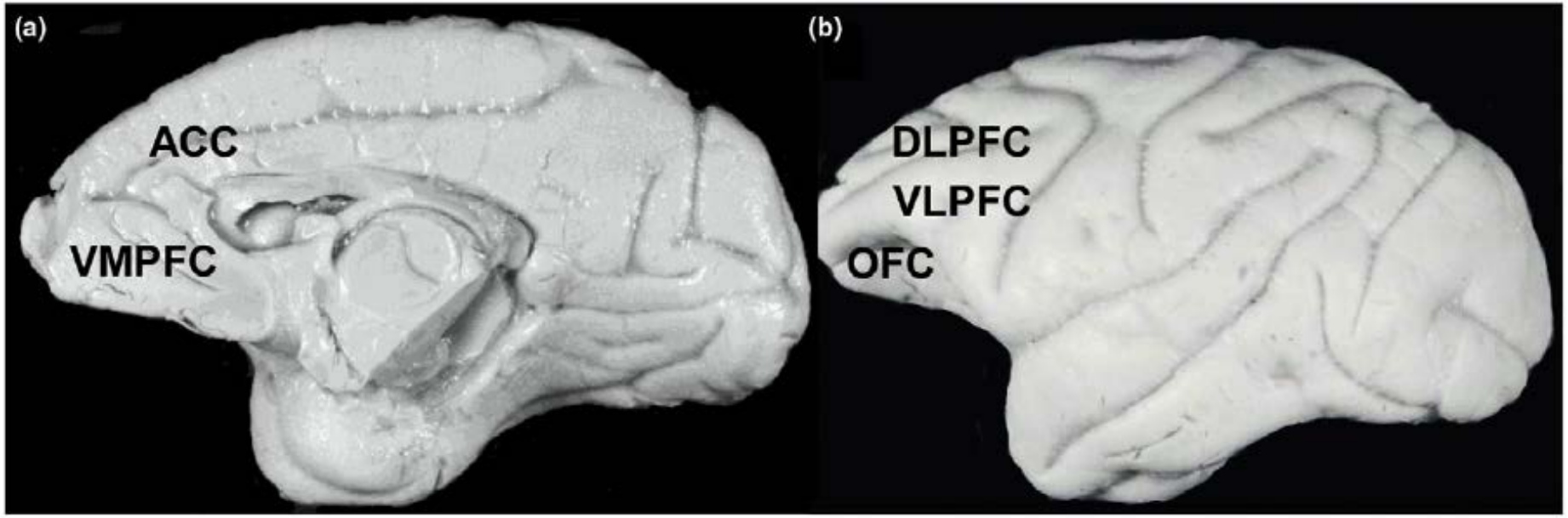
(1) The brain computes
   (1) a subjectiv «decision» value signal for each perceived alternative at the time of choice
   (2) an «experiences value signal» at the time of consumption
(2) Choices are based on the comparison of decision values
   (1) The comparison process can be captured by drift diffusion models

(3) Decision values are based on the integration of information about the attributes and the attractiveness of the perceived alternatives

(4) The computation and the comparison of decision values is shaped by attentional processes

The encoding of decision values in the primate brain

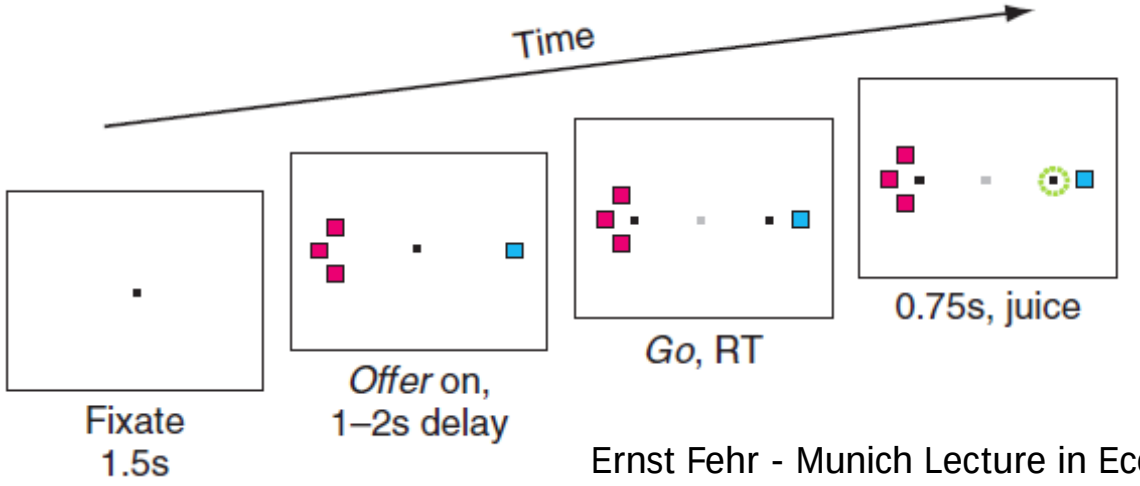
Important brain regions implicated in economic choice



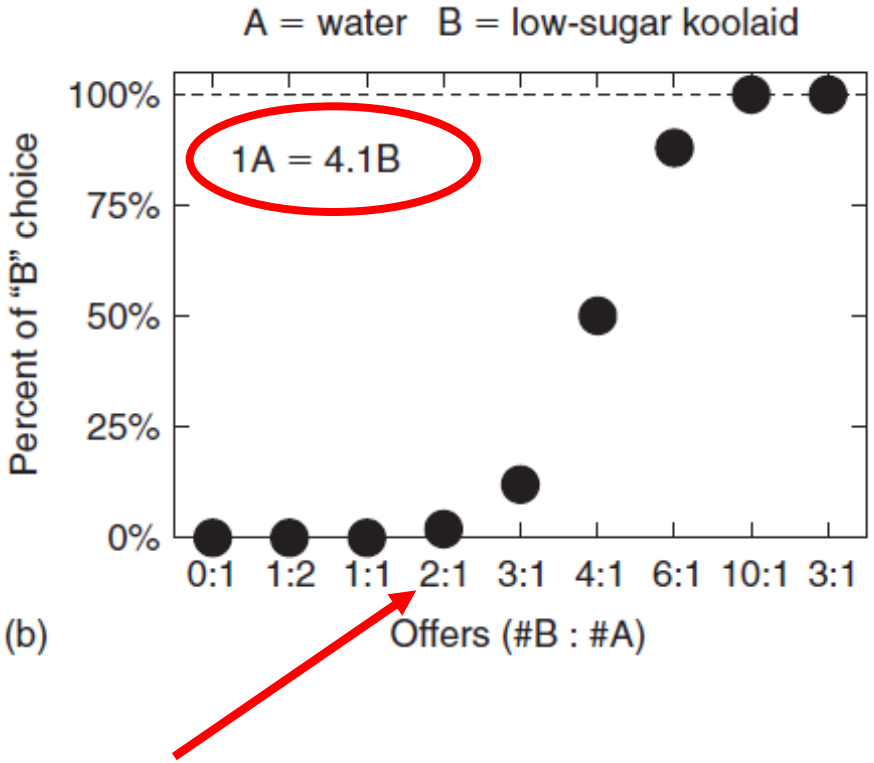
Are there neuronal measures of subjective economic value?

(Padoa-Schioppa & Assad 2006)

- Monkeys can choose between different units of good A (the preferred good) and good B
- Color of squares indicates food type
- Number of squares indicates food amount



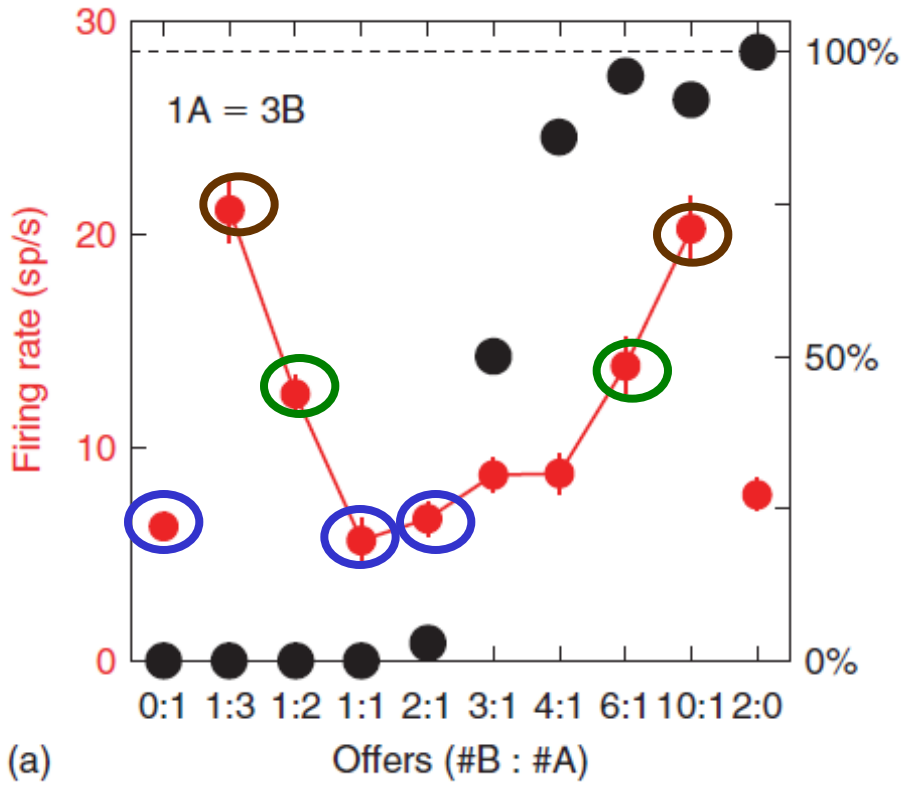
Computing economic value



- V(1A) = V(4.1B)
- V(1A) = 4.1V(B)
- Determines the offered value for each good
  - Search for offer value neurons
- Determines the value of the chosen good
  - Search for chosen value neuron

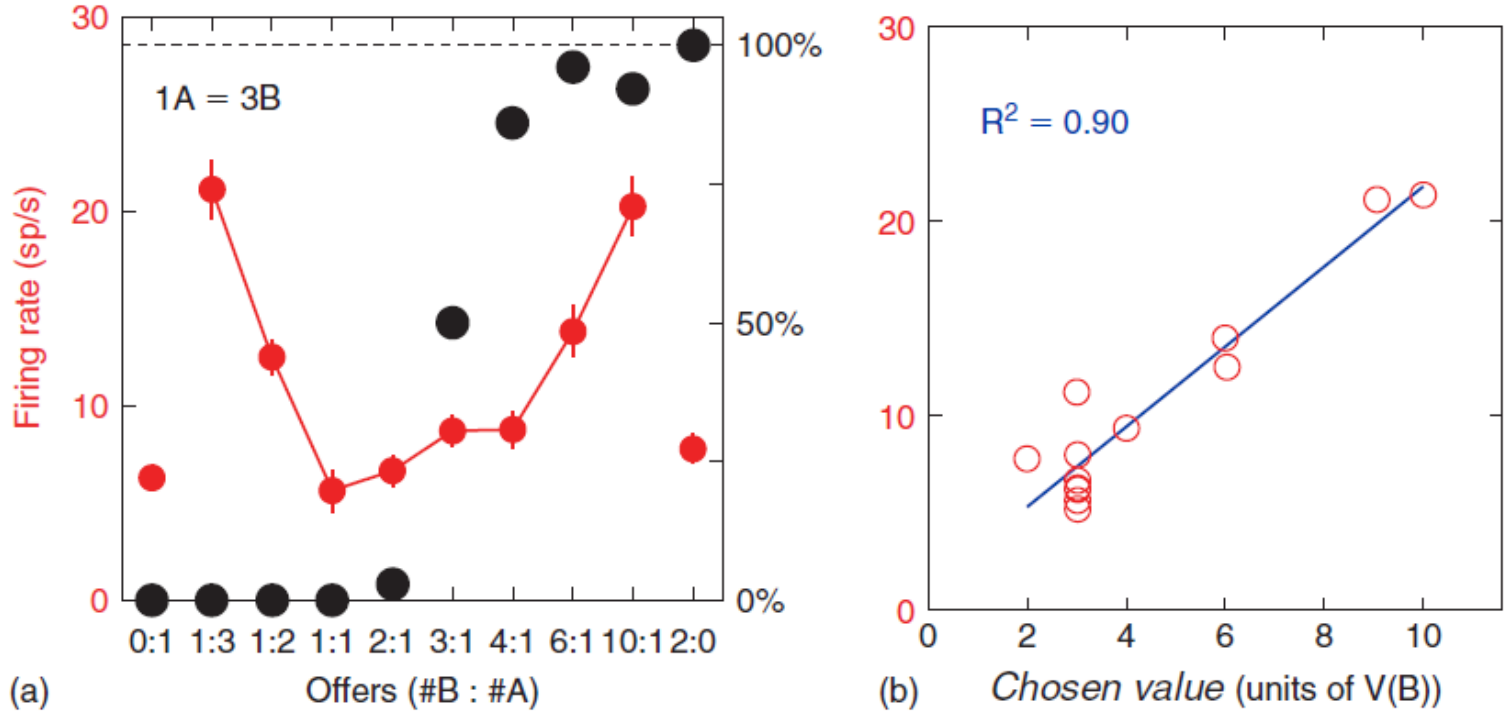
Offered value for B = 2B
Offered value for A = 4.1B
Chosen value = 4.1B

Do neurons in the orbitofrontal cortex encode subjective economic value?



- Why OFC?
  - OFC lesions associated with pathological choice
- Blue circles: one unit of A chosen
  - Chosen value = 3
- Green circles:
  - Chosen value = 6
- Brown circles:
  - Chosen value = 9-10

Relationship between firing of OFC neurons and subjective economic value



- 18% of 931 OFC neuron encoded chosen value

Other findings

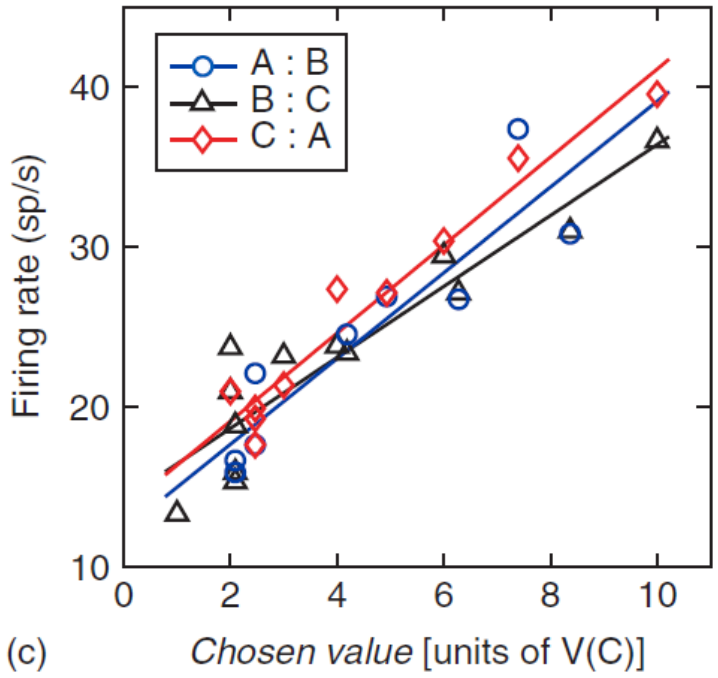
- Almost all neuronal responses (95%) did not depend on spatial configuration of the visual stimuli

- Some neurons in OFC encode the value of only one of the goods offered in the choice menu

Does neuronal firing encoding chosen values depend on the choice menu?

(Padoa-Schioppa & Assad 2008)

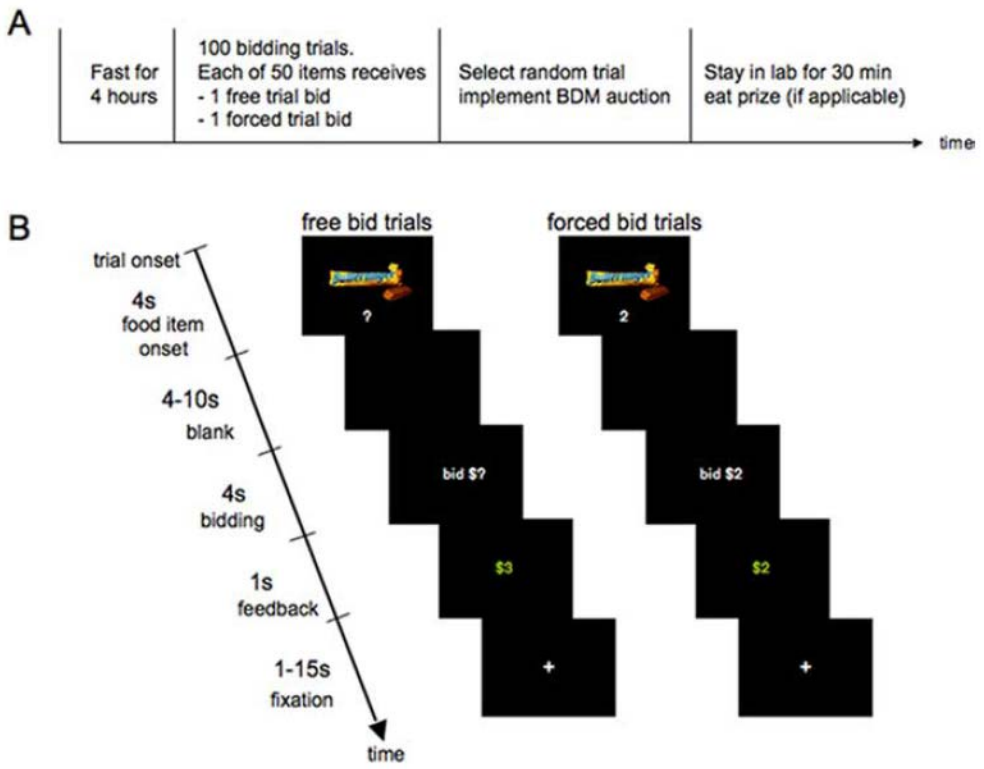
- 3 goods A, B, C with decreasing preference
- Is the „neuronal value“ of the chosen good different depending on the offered pair of choices?
- No, the neuronal value of the chosen option is the same regardless of the choice pair
- **Neuronal firing seems to provide a cardinal measure of chosen value**
- Also, activity of neurons encoding **offer value** does not depend on the other good.
- Caveat: adaptive coding



Does OFC encode subjective economic value in humans?
(Plassman, O'Doherty, Rangel 2007)

- Subjects have to bid (reveal their maximum willingness to pay) for the right to eat a familiar junk food item at the end of the experiment
- 0, 1, 2 or 3 $ could be bid
- Bidding on 50 different food items in two conditions. One trial randomly selected at the end.
- Random device selects a price p for that item
  - If b > p, subject must buy at price p
  - If b < p, subject does not buy the item
- Free bid trials: subjects need to compute a willingness to pay
- Forced bid trials: subjects must bid a randomly determined amount

Subjects are scanned with fMRI while bidding



- Brain activation in free-forced trials is of interest because it encodes the computation of subjective value

In which brain region does neuronal activity in (free-forced) trials correlate with WTP?

- OFC and DLPFC activity correlates with WTP
- OFC does NOT correlate during forced bid trials with WTP
- OFC does NOT correlate during forced bid trials with forced bid

Comparison of decision values
The Drift-Diffusion Model
(Ratcliff 1978, Ratcliff & McKoon 2008)

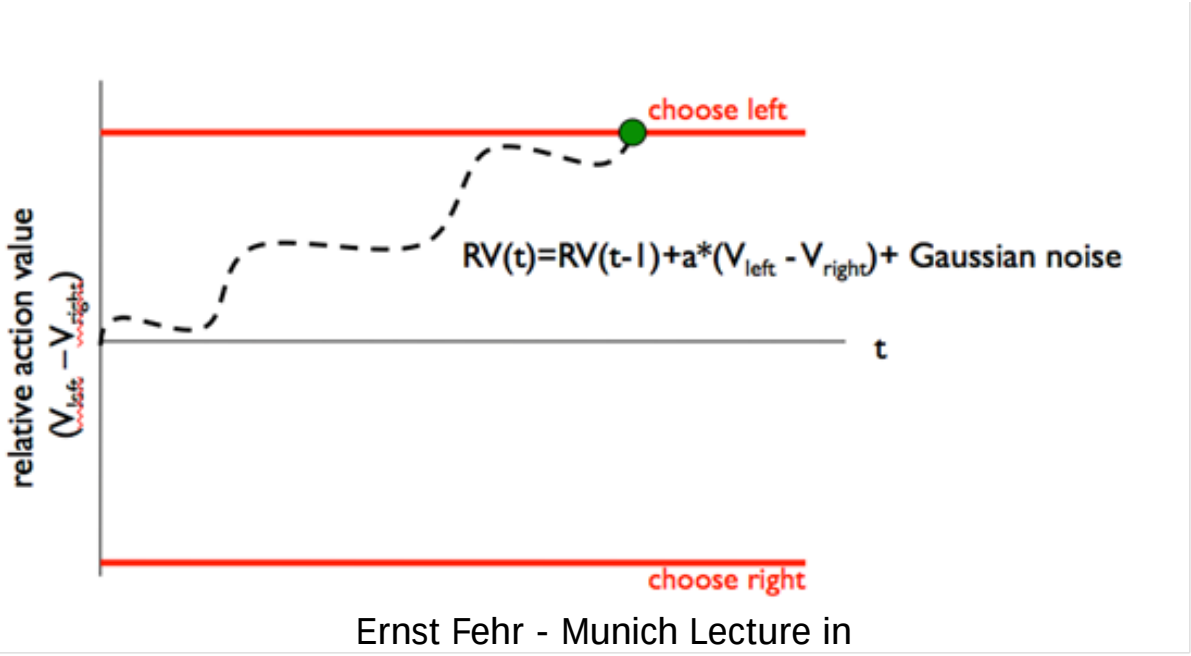
- Binary choice between alternatives X and Y
- Latent values of x and y measured in terms of neural activations
- Relative decision value signal RV reflects the value difference x – y
- RV starts at 0 and evolves over time according to

$$RV_t = RV_{t-1} + a[x - y] + \varepsilon_t$$

- a measures how quickly latent decision value signals are integrated into RV
- ε_t random influences («neurons fire stochastically»)

The Drift-Diffusion Model
Basic Idea & Choice Mechanism

- The brain computes the RV_t – it's not just given
- If RV_t passes a threshold a decision occurs



Evidence for DDM

- Initial model developed to explain reactions times and errors for choices that occur within a few seconds
  - Which stimulus is brighter
  - Model received good empirical support

- Application of the DDM to decision values by Rangel and his team
  - Milosavljevic et al. (2010)
  - Basten et al. (2010)
  - Hare et al. (2010)

Properties of the DDM

(1) Because neurons fire stochastically the RV's and the resulting choices are also stochastic

(2) Probability of choosing x is a logistic fucntion of the decision value difference (x – y)

(3) Model implies a clear definition of decision errors and the probability of such errors is always positive

Properties of the DDM

(4) If the stochastic influence is iid and normally distributed the brain implements the optimal statistical decision on the basis of a likelihood ratio test

- Intuition: $RV_t > 0$ can be interpreted as the accumulated evidence for x > y; the larger RV_t the larger is the evidence

(5) The DDM provides a microfoundation for stochastic utility models, i.e., these models can be viewed as a reduced form DDM (Webb 2013)

The computation and the comparison of decision values is affected by attentional processes

$$RV_t = RV_{t-1} + a[\theta x - y] + \varepsilon_t$$

- If attention is directed towards X then RV_t increases
  - Attending to X produces more neuronal evidence in favor of X

$$RV_t = RV_{t-1} + a[x - \theta y] + \varepsilon_t$$

- If attention is directed towards Y then RV_t decreases
  - Attending to Y produces more neuronal evidence in favor of Y

Predictions of attentional DDM

- Higher attention for X increases the probability of choosing X
  - Choice can be influenced by affecting attention
- The good that is attended to immediately before the decision will be chosen
- Empirical tests in choice experiments with measurement of eye movements with eye trackers
  - Armel, Beaumel und Rangel (2008)
  - Krajbich, Armel und Rangel (2010)
- Estimates yielded $\theta = 0.3$ across a number of studies involving food choices

The Social Drift Diffusion Model
(Krajbich, Hare, Bartling Morishima, Fehr 2013)

- Look at X: $RV_t = RV_{t-1} + a[\theta x - y] + \varepsilon$
- Look at Y: $RV_t = RV_{t-1} + a[x - \theta y] + \varepsilon$

- With other-regarding preferences, X and Y are vectors of payoffs (x_i, x_j) and (y_i, y_j)

- We consider the relative decision value of individual i:

- **$RV_t = RV_{t-1} + a\,[(x_i - y_i) + \theta(x_j - y_j)] + \varepsilon$**

Other-regarding preferences and attention towards others payoff

- Why should the weight on others' payoff be related to the attention parameter in the food choice study?
  - Fiedler et al. (2013): subject with stronger SVO's allocate more attention to other's payoff
  - Hypothesis: on average, if one faces anonymous strangers, subjects primarily focus on their own payoff such that others payoff is discounted with θ

Social DDM – Parameter Choices

- $RV_t = RV_{t-1} + a\,[(x_i - y_i) + \theta(x_j - y_j)] + \varepsilon$

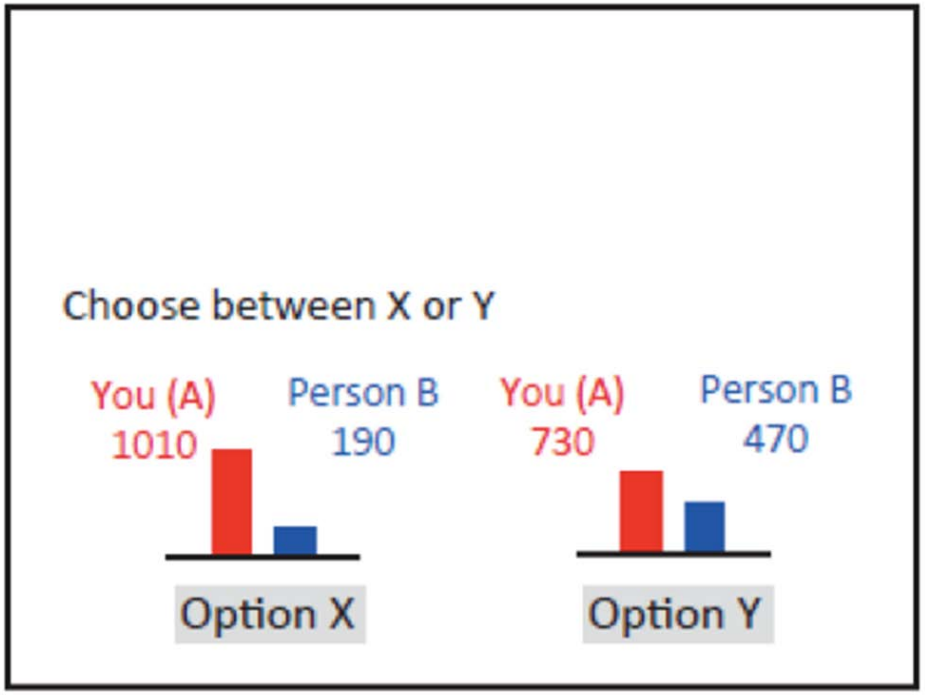
- Can we take the values of «a», the variance ε (denoted by σ) and the attentional parameter θ from the food choice studies by Krajbich et al. and still predict social choices well?

Parameter choices for predicting behavior in social preference tasks

	Payoffs' order of magnitude $\bar{v}$	d	$d \cdot \bar{v}$	σ	θ
Food task	10	0.0002	**0.002**	**0.02**	**0.3**
Task 1	1000	0.000002	**0.002**	**0.02**	**0.3**
Task 2	100	0.00002	**0.002**	**0.02**	**0.3**
Task 3	10	0.0002	**0.002**	**0.02**	**-0.3**

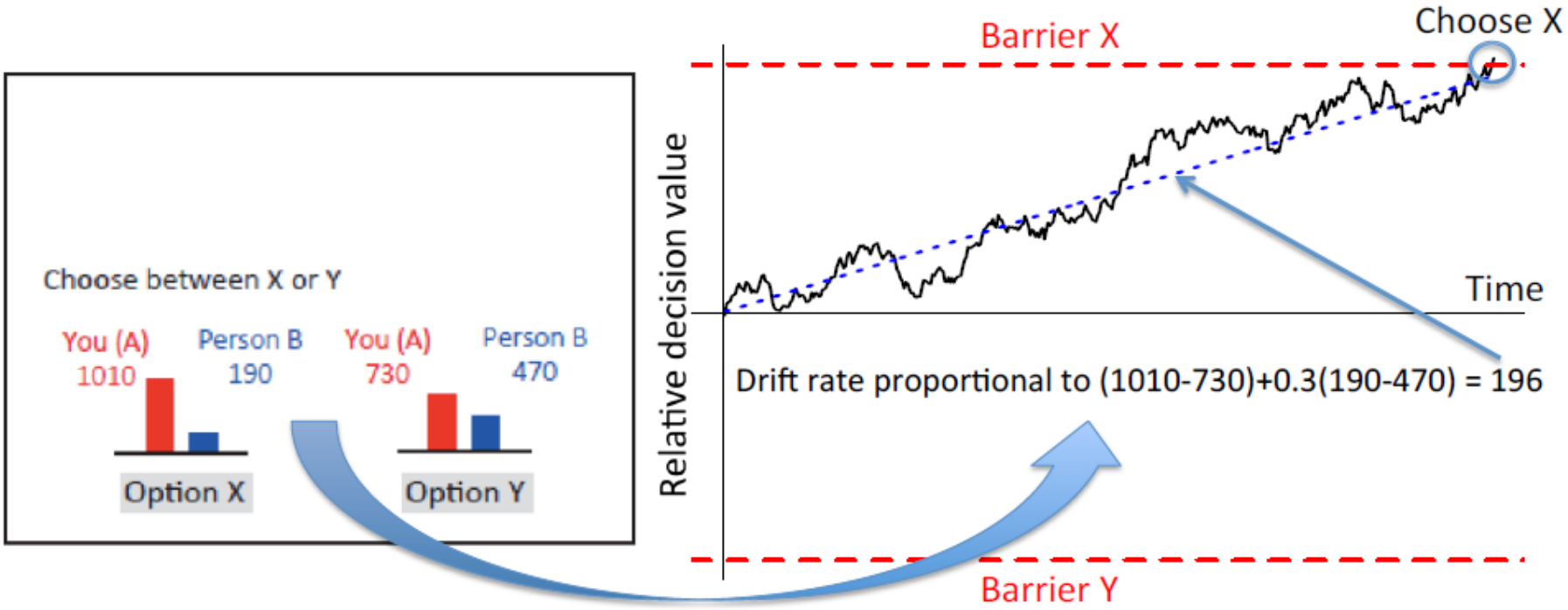
Table S1: Value ranges and parameters used in the original food study and the social-preference Tasks 1-3 of the current study.

Binary dictator games

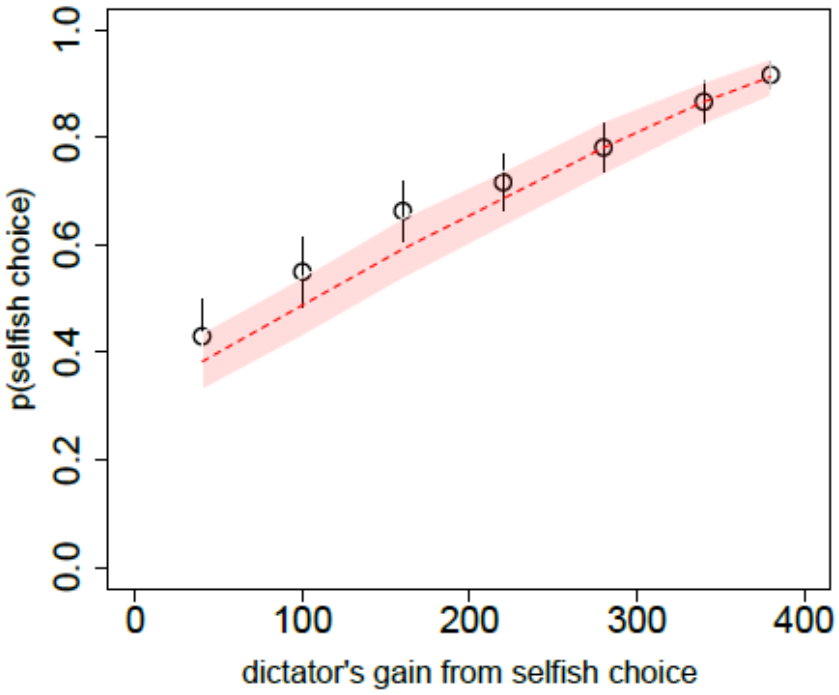


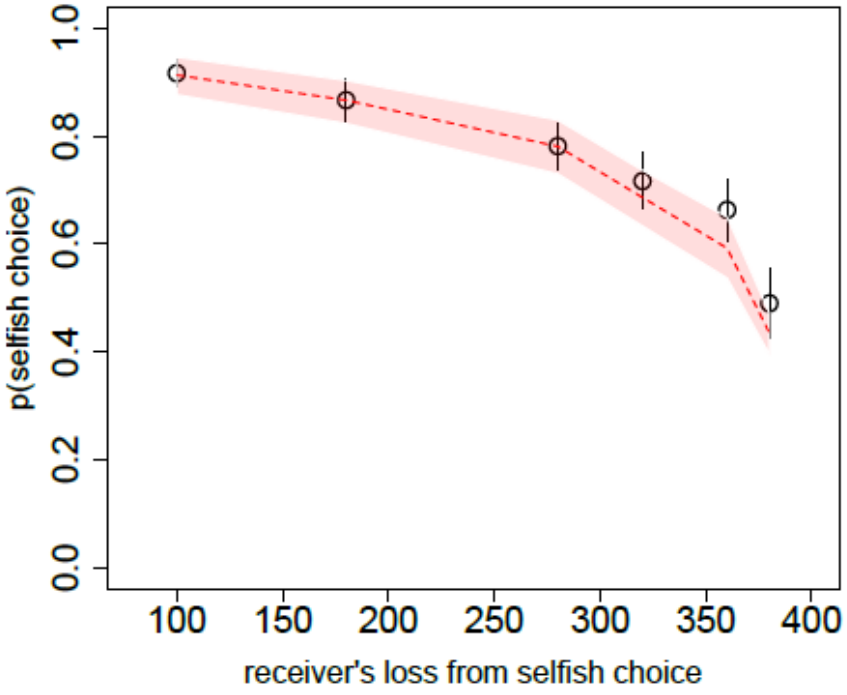
- 30 subjects
- 70 binary decisions each

Mechanics of Social DDM

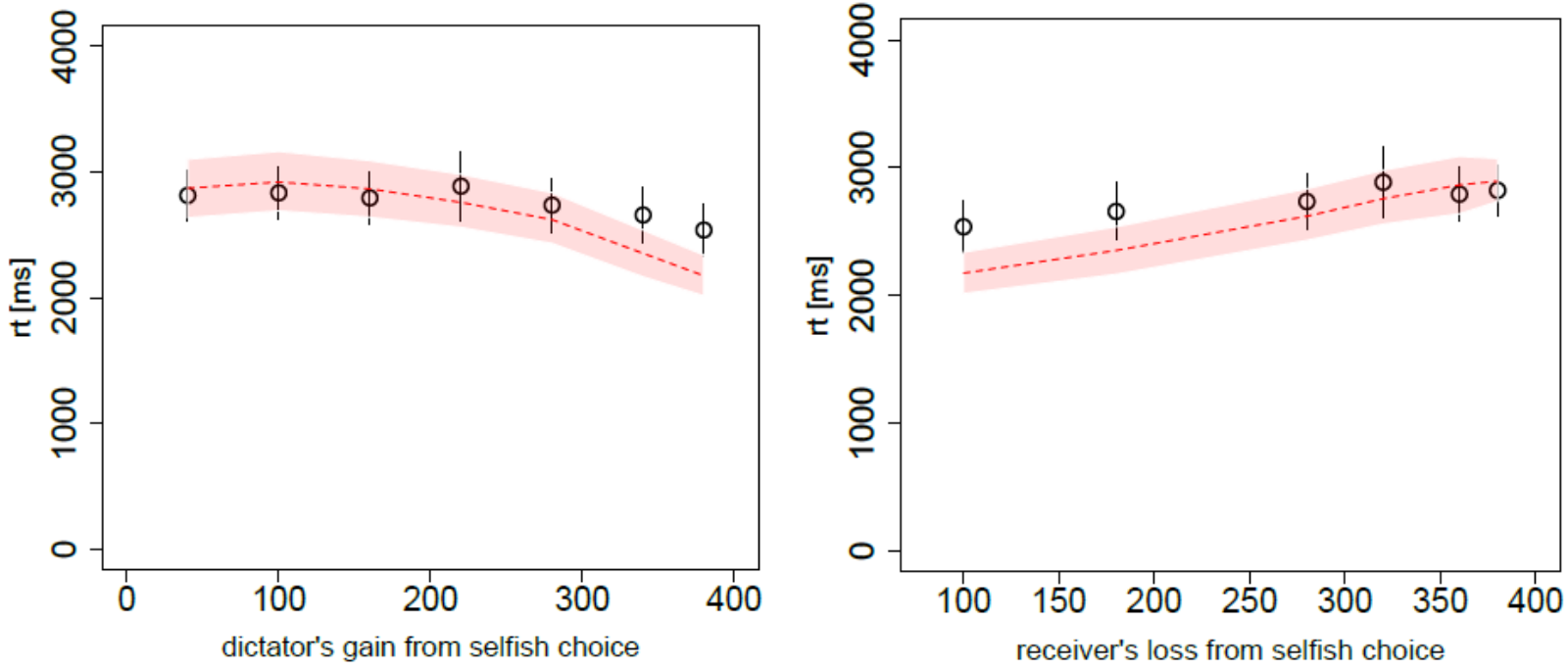


Do we predict average behavior?





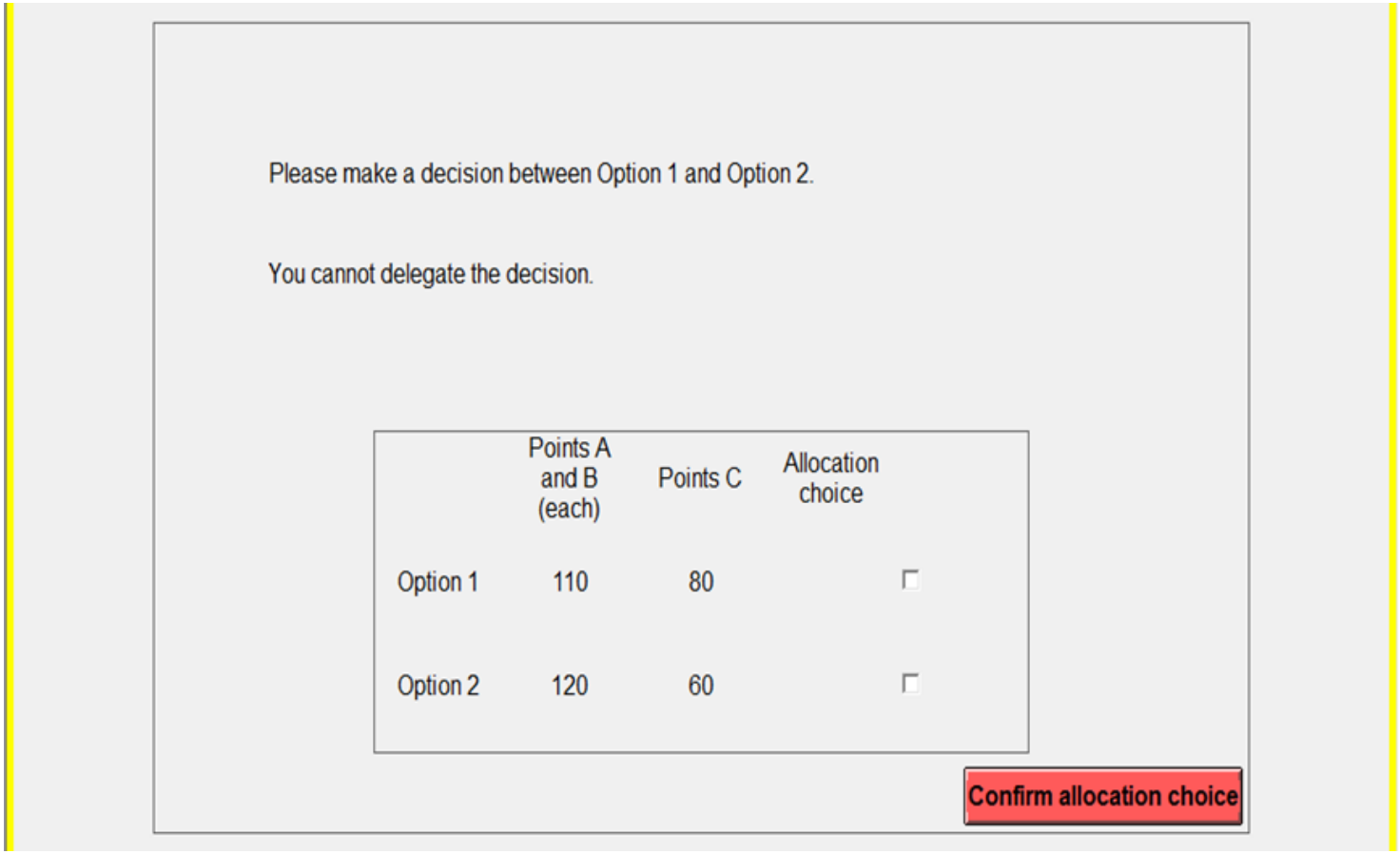
Do we predict reaction times?



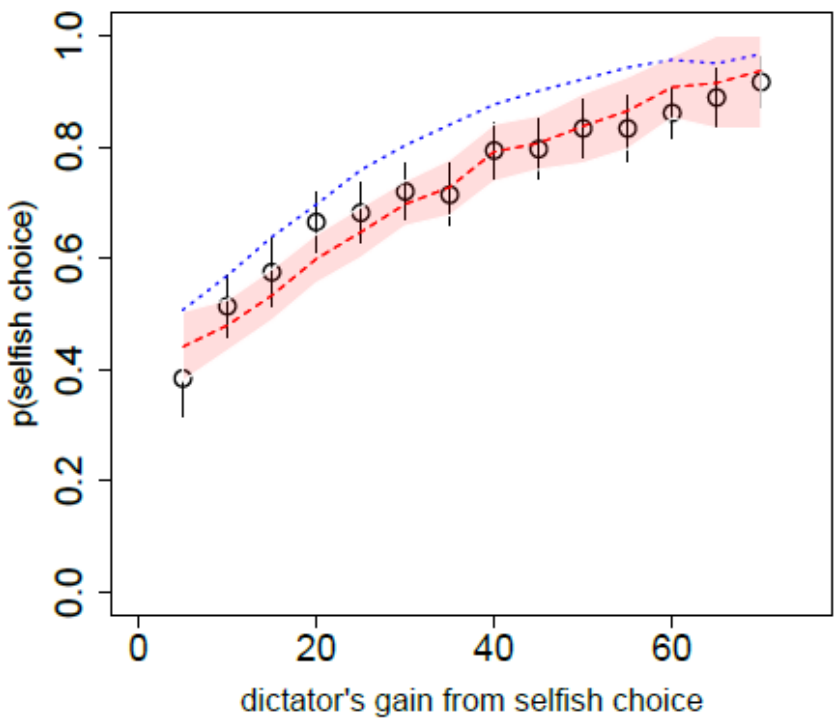
A second dictator game (with two recipients)

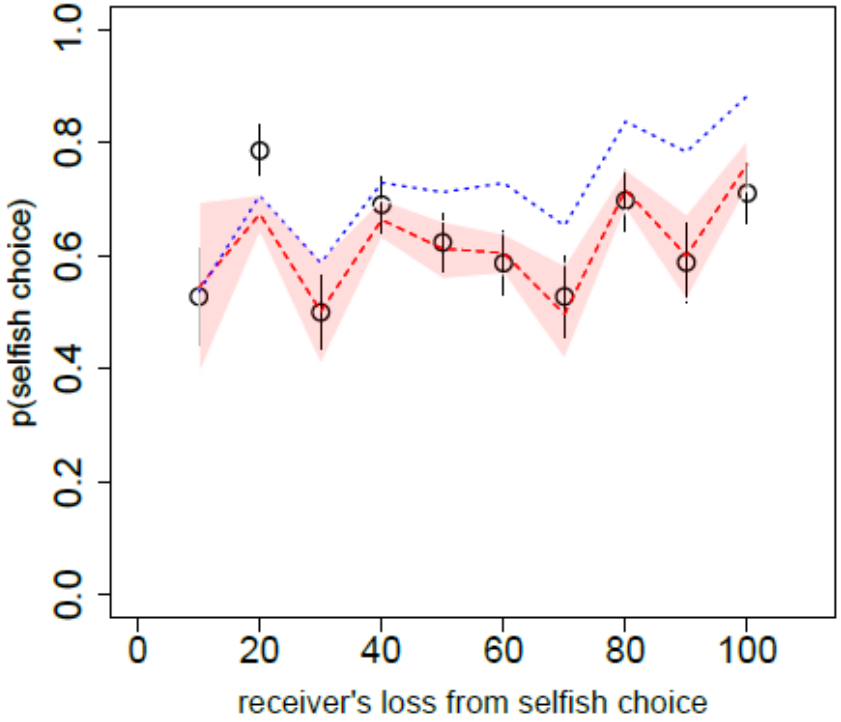
- 36 dictators, 100 binary choice problems
- One partner always received the same amount as the dictator
- Fehr-Schmidt predicts that in this 3-person DG the weight for the other's payoff is reduced from 1 to 1/(n-1) = ½ relative to the 2-person DG
- Social DDM predicts that no change, i.e. the same θ-value applies because they need not attend to the third player's payoff

Payoff Screen in 3-player DG

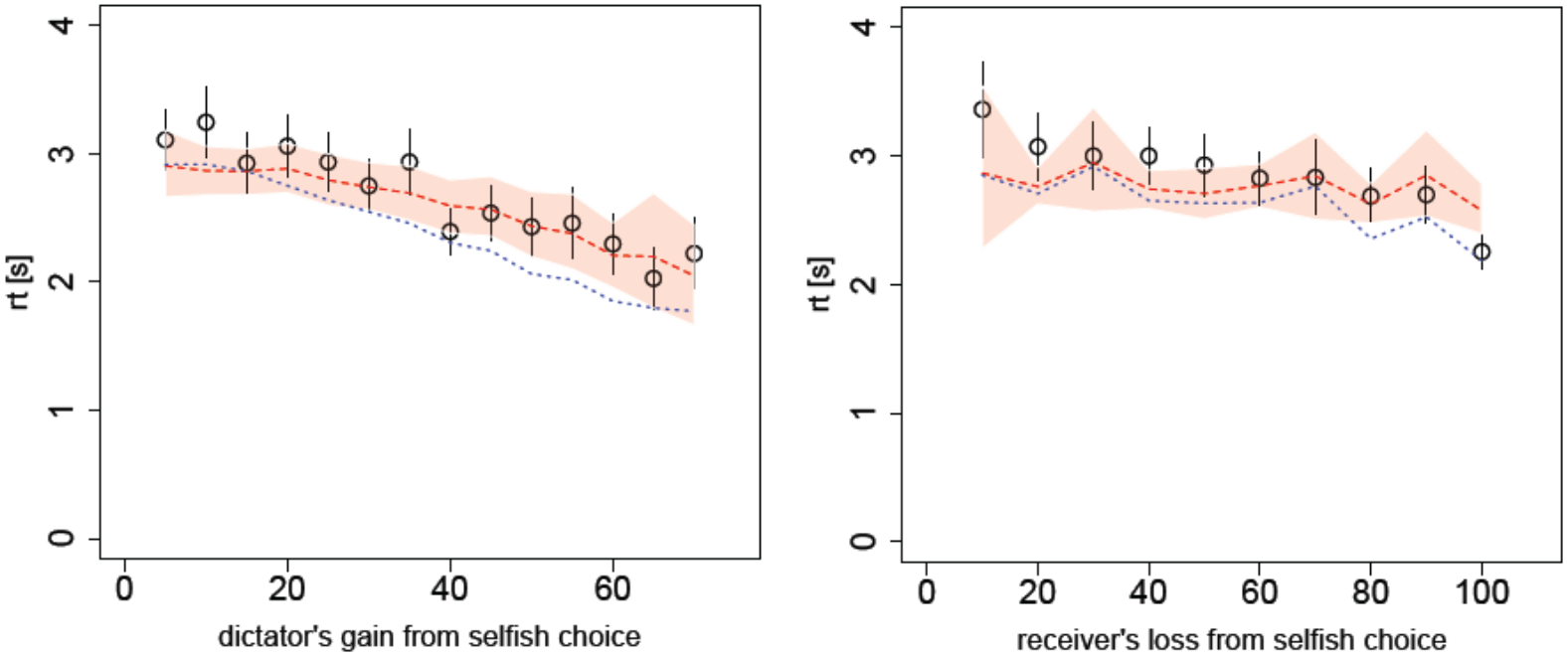


Predicting Behavior (3-player DG)

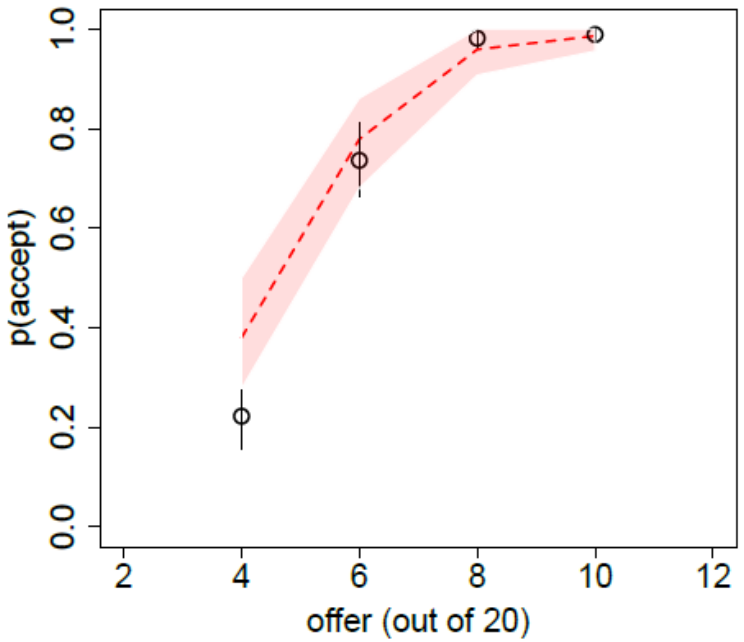


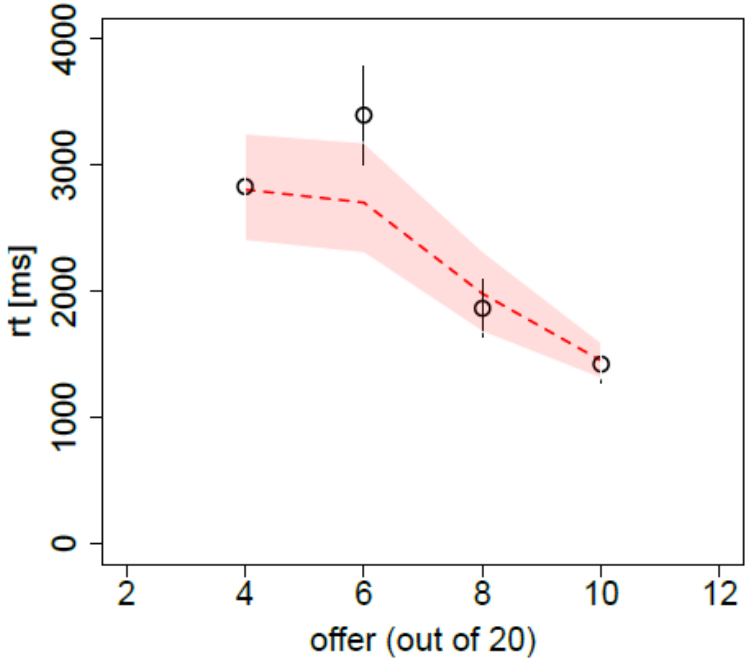


Predicting Reaction Times (3-player DG)



Ultimatum Game
Rejection Behavior and Reaction times





Summary I

- Neural activity is causal for our behaviors
- A full understanding of human behavior involves understanding the underlying neural ciruitry
- The economic approach to behavior is a reduced form approach that lacks a microfoundation in terms of the underlying neural circuitry

Summary II

- Decision values are encoded in ventromedial prefrontal cortex
- The DDM is a promising – neuronally informed – model of choice processes
- The DDM (or similar models) provide a microfoundation of stochastic utility models
- The social DDM provides a remarkable good quantitative fit for other-regarding behaviors