



SFB/Transregio 266

ACCOUNTING FOR
TRANSPARENCY

botex: Using LLMs as Experimental Participants in oTree

Joachim Gassen

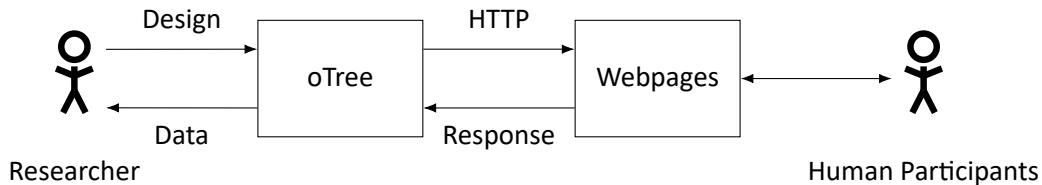
Humboldt-Universität zu Berlin

March 25, 2024

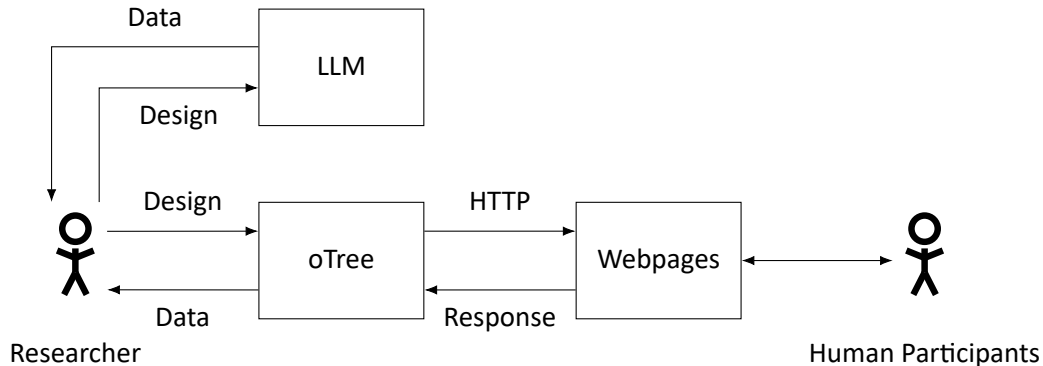
Motivation

- 1 Designing experiments is hard and small differences in experimental materials can have significant impact on experimental findings. Wouldn't it be great to have a "cheap" way to pretest experimental designs without bothering humans?
- 2 Understanding how LLMs act in behavioral experiments - on their own and interacting with humans - is an emergent and rapidly evolving research field. Providing an infrastructure to run such experiments should be beneficial for the profession.
- 3 Applied experimental work often uses context framing in their experimental designs without explicitly hypothesizing and assessing its effect on findings. LLM based experiments might inform priors in that regard.

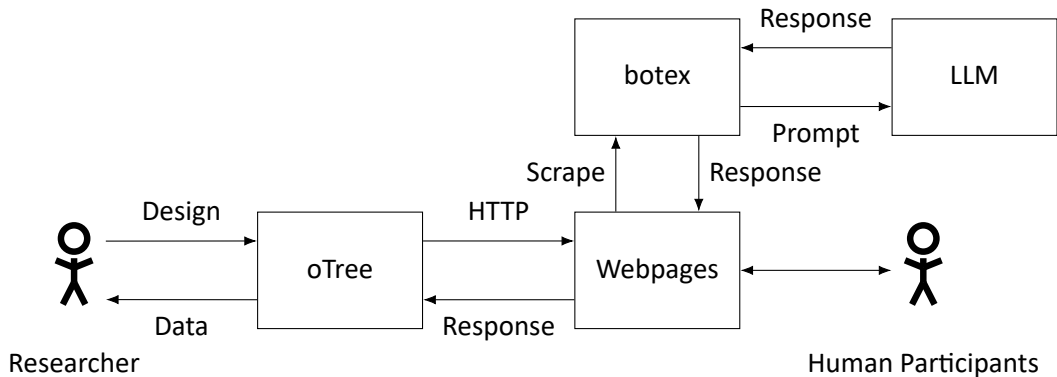
Traditional oTree Setup



oTree + Alter Ego



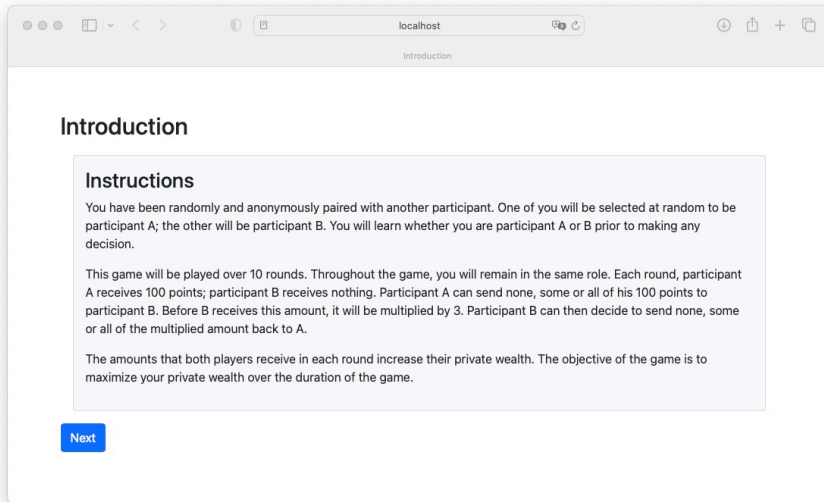
Engel, Grossmann and Ockenfels (2024, SSRN)



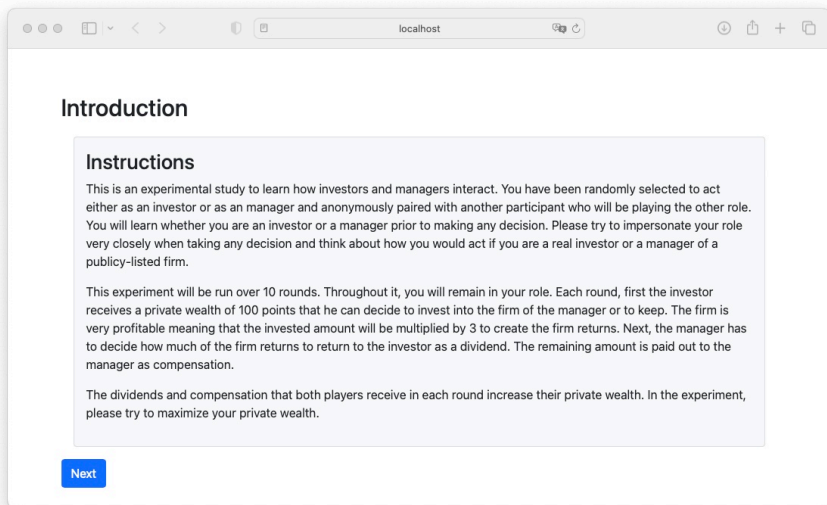
A First Example

- As a first sniff test, I study whether an investor/manager framing in a classic multi-round investment trust game (Berg, Dickhaut and McCabe, Games and Econ Behav 1995) affects the observed outcome.
- Prior work has shown that the investment trust game is subject to context framing effects (e.g., Al-Ubaydli, Houser, Nye, Paganelli, and Pan, PLOS One 2013; Cronk and Wasielewski, J of Evolutionary Psychology 2008)
- Can similar framing effects also be observed for LLMs?

The Neutral Frame



The Investor/Manager Frame



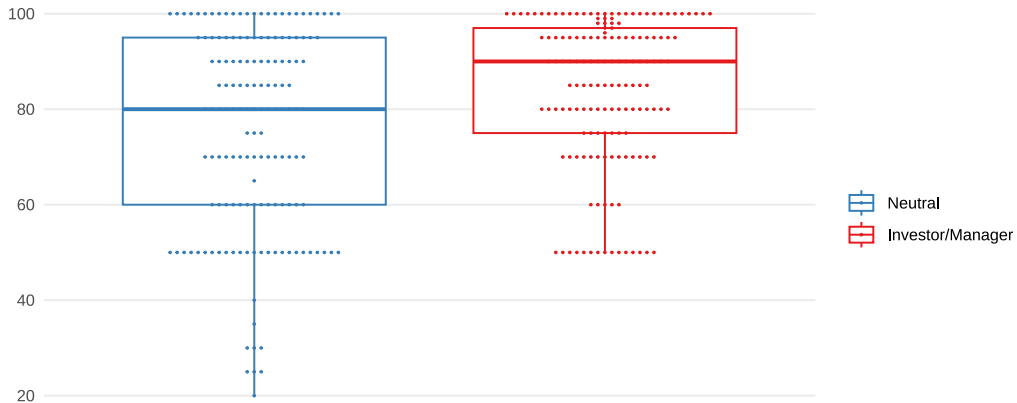
Sample Composition

| | Sender | Receiver |
|-----------------------------|--------|----------|
| Neutral Framing | 15 | 15 |
| Investor/Management Framing | 15 | 15 |
| Total | 30 | 30 |

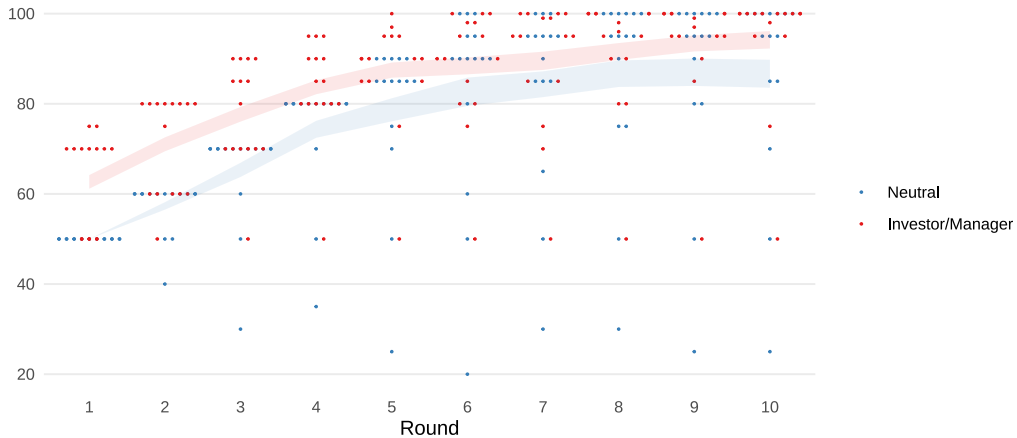
Descriptive Statistics by Framing

| | Neutral Framing | | | | Investor/Management Framing | | | |
|-----------------|-----------------|--------|-------|--------|-----------------------------|--------|-------|--------|
| | N | Mean | S.D. | Median | N | Mean | S.D. | Median |
| Amount sent | 150 | 75.30 | 21.06 | 80.00 | 150 | 83.96 | 15.66 | 90.00 |
| Amount returned | 150 | 114.44 | 29.80 | 120.00 | 150 | 128.98 | 26.61 | 135.00 |
| % returned | 150 | 0.51 | 0.05 | 0.50 | 150 | 0.51 | 0.04 | 0.50 |

Sent Amount by Framing



Sent Amount by Round



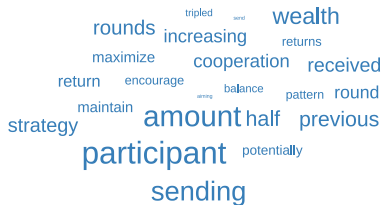
Regression Results

| | (1) |
|-----------------------------|----------------------|
| (Intercept) | 52.689*** (2.693) |
| Round | 4.111*** (0.434) |
| Investor/Management | 13.644*** (3.809) |
| Round × Investor/Management | -0.906 (0.614) |
| Num.Obs. | 300 |
| R2 Adj. | 0.356 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Some Verbal Response

Neutral Framing



Investor/Management Framing



Manipulation Checks ;-)

| | | Neutral | Inv/Manag |
|--|-------|---------|-----------|
| Answers comprehension question correctly | Yes | 30 | 30 |
| | No | 0 | 0 |
| Remembers role | Yes | 29 | 27 |
| | No | 1 | 3 |
| Characterizes as ... | Human | 25 | 21 |
| | Bot | 5 | 9 |

A Very Preliminary Second Example (Gneezy, AER 2005)

| | Neutral Framing | | Accountant/HQ Framing | |
|------------|-----------------|------|-----------------------|------|
| | N | Mean | N | Mean |
| % lied | 60 | 0.62 | 60 | 0.92 |
| % followed | 60 | 1.00 | 60 | 1.00 |

Chi-Square: 13.463***

It works (at least for a very simple albeit repeated game):

- Context framing as investor/manager setting increases willingness of the sender to send money, and leaves strategy of the receiver unaffected
- Costs: ~30 US-\$ for 60 participants, duration: ~4 hours (fully scriptable)

Next steps

- Make framework portable
- Other games?
- Identify seminal accounting study to compare to its “neutral” alternative
- Alternatively: replicate Liberman, Samuels and Ross (2014) study on “Community” vs. “Wallstreet” framing for prisoner’s dilemma (but: likelihood that results are in LLM training material...)
- Compare findings with results for human participants