Modelling the effects of information gathering on social decision making

Damon Hayhurst

Project Report Submitted as

partial fulfilment for the

Degree of Cognition and Computation,

Birkbeck, University of London

July 30, 2024

I confirm that this write up is my own work and does not involve

plagiarism as defined in the module information provided

Signed and dated

This is the abstract

Introduction

The aims of this study is to gain an insight into how the effects of information gathering influence our decision making in a social context. Using process tracing, we hope to observe and characterise strategies that are used in order to arrive at these decisions. Such an approach hopes to shed light further on the processes that underlie decision making in human behaviour.

Aims/Objectives

- Ask whether information gathering strategy can be used a predictor in social decision making
- Understand whether perceived monetary outcomes to self and another have an affect on the way we gather information
- Determine if information gathering strategy in social decision making reflects an individual preference

Bounded Rationality

Herbert A Simon first outlined the idea that human behaviour in decision making was not to be defined by the decisions taken but that of the information surrounding the decision. This concept of 'Bounded Rationality' suggested an approach that took consideration of the contextual information regarding the decision as well as the computational capacities available to the decision maker. The concept sought to defy prior notions that human behaviour in decision making could be understood through a global notion of homo economius; Where decisions taken by agents were always perfectly rational, seeking to maximise their own self interests, and, instead, suggested that a human's capability to be perfectly rational was constrained by internal and external constraints.

The approach proposed by Simon places as much of an emphasis on the processing of information during the decision making as the decision itself. Pivotal to the concept of 'Bounded Rationality', Simon suggested that human's mental capacity to make choices is subject to heuristics or mental simplifications that contend with it's ability to be a perfectly rational agent but aid it's ability to make decisions in the face of increasing complexity (Payne, 1976). Through the use of process tracing in the information gathering stage of decision making, one can shed light upon the influence of environmental factors and the computational limitations attributed to the mind that prevent humans from acting in a rational manner.

Process Tracing

Process tracing

Early research into the heuristics that govern our decision making has uncovered the necessity for capturing process data through process tracing methods. The Priority Heuristic is one such model that was created without the use of process data. Using an ordered list of reasons, the Priority Heuristic attempts to explain choices within the context of simple monetary gambles. While the mechanism was shown to be significant at the aggregate level, the model failed to capture differences at the individual level. Further research into the manifestation of the Priority Heuristic, using process data, uncovered that in actual fact the underlying process did not mimic the priority ordering set out by the heuristic. The data, however, revealed that choice could be predicted significantly by the number of transitions into and out of a choice. The uncovering of such mechanisms highlights the need for processing data in explaining how humans make decisions. Process data can provide the explanatory power for decision making differences and inform the development of heuristics that can be applied at the individual level (Brandstätter et al., 2006).

Process tracing has shown efficacy in the development of other more generalised decision making strategies. Show

It is important to note that it is widely believed that human's decision making strategy is constructed contingent on the task demands and individual preferences. Decision making strategies have been shown to vary widely based on the number of choices at hand and the task complexity (Payne and Bettman, 2004). Human's have a 'toolbox' of strategies which can be utilised in the cwhen it comes to information gathering and the choice reflects task demands and individual preferences. Such strat

Using process tracing one can observe and defined these atomised strategies

MouselabWeb

Intertemporal Choice task

Social Decision Making

Methods

A variation of the Message game was used in the experiment. The game uses honesty to illustrate how decision making can be affected by value judgements that pit our own self interest against our altruistic sensitivities. The game involves two roles. The Sender and the Receiver. The Sender must send a message to the Receiver informing them of a particular box to open from a choice of four. Each of the boxes contains differing amounts of money for the Sender and the Receiver should that box be opened. The Sender can ultimately see what's inside each of the boxes in terms of how much money they will receive versus the amount that the Receiver will. Only the Receiver opens the box, however. A message is sent by the Sender to the Receiver with words to the affect of 'Box X is best for you'. The dishonesty comes from the Sender being able to suggest a box that might not necessarily represent the truth and instead picking a box that contains a larger amount of money for themselves and less money for the Receiver than that of one of the other boxes.With the Receiver unable to see the quantity of money they potentially missed and in this variation, two of the boxes being blank, the Receiver has no choice but to trust the potentially dishonest disclosed by the Sender.

The experiment took place through a web browser. Each participant was provided with a run through of the game from the perspectives of both roles. Showing the nature of choosing a box to suggest as the Sender and opening a suggested box as the Receiver. With an learned understanding of both roles, the rest of the experiment takes place with the participant in the Sender role. Over 80 trials, differing values are placed in to two of the four boxes within the screen's width. The position of the two boxes that contain values varies as do the values contained within each of them. Each box has two values inside containing the listed values received for both the Sender and the Receiver. Each of these values is covered up and in order for the participant (or Sender) to view the value, they must hover over each item in order to see the contained value. Being unable to see the values without hovering over them suggests that when a participant is hovering over a value there is suggestion that they are dwelling on such an idea.

Processing the data

Preprocessing

All analyses was performed on participants that matched the condition wherein a proportion of their responses represented choices that were lying to the receiver. Participants who selected over 95% lie choices were not considered and similarly participants who chose over 95% truthful options were not selected either. Any participant who picked one of the two blank boxes in over 5% trials were also not selected. Any participant who did not finish the 80 trials was also removed from the analyses suggesting a lack of engagement. Similarly, any trials that

Identifying areas of interest

Creating a timeline

Average analysis

Distance analysis

Creating time series sequences

Measuring the distance between trials

Clustering

Results/Findings

Trial and Participant Quantities

The processing of data removed a proportion based on criteria that represented a lack of engagement with the experiment. six participants were removed for not finishing it. 20 were removed for telling the truth over 95% the time and a further 14 participants were removed for recommending the blank boxes to the receiver in over five percent of the trials. Interestingly, no participants were removed for lying over 95% of the time; despite this being the approach favoured in game theory. A further 40 trials were removed from the dataset because they recorded decision making times that were three standard deviations away from the mean. An approach also taken by Reeck et al., suggesting a similar lack of engagement with the task at hand.

Irregularities and missing data that represented a failure in process tracing were also removed from the dataset. 10 participants were removed for having not recorded mouse coordinates in over 85% of their trials and a further 108 trials removed from the remaining valid set because of similar lack of mouse coordinates. Irregularities in some of the coordinate data meant that AOIs could not identified through measurement. 42 trials were removed because of this and one participant also, that failed to record screen width data. If such impairments in the dataset had been included by merely estimating their missing qualities, it would have compromised the overall integrity. Omitting them highlights the essential role of accurate process tracing.

After accounting for engagement criteria and irregularities, the number of valid trials was recorded for each participant and a further two participants were excluded from the analysis. The number of valid trials for each of these participants represented a quantity that was less than three quarters of the intended 80 trials for each participant. Given, the ordering of each trial was randomised, the set of valid trials for each of these participants could represent biases in likelihood to lie relative to the whole data set.

From the original pool of 89 participants, only 36 remained after all removals. An overall data set that contained 2783 trial instances each with a valid process trace.

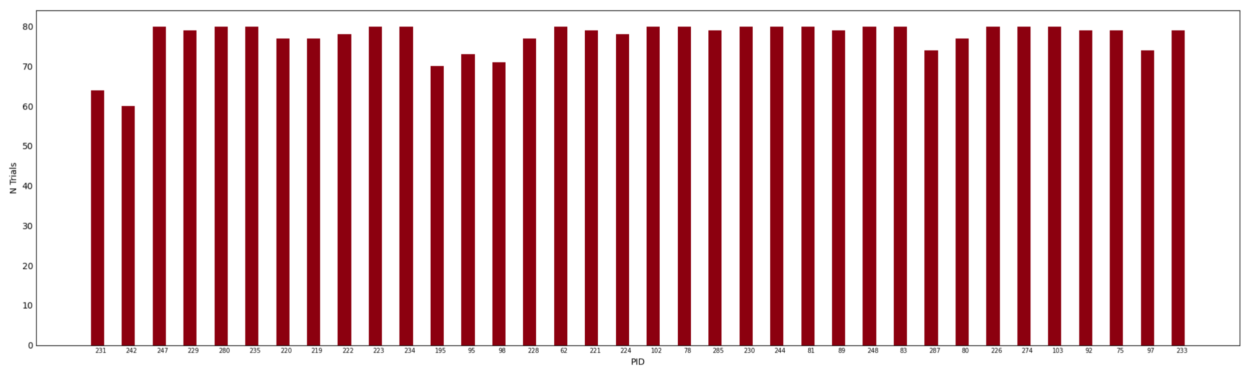


Figure 1

Bar graph representing number of trials in the final valid data set for each participant.

Descriptives

Lie Percentage

Across trials the propensity to lie displayed a high degree of variance. The average likelihood to lie on any given trial was 39% ($SD = 49\%$). Across participants, the propensity to lie was measured between 3%-91% for a given individual. Across instances of unique trial, where unique is defined by the specific quantities pertained to in the boxes, the range was 0-87% for any given trial.

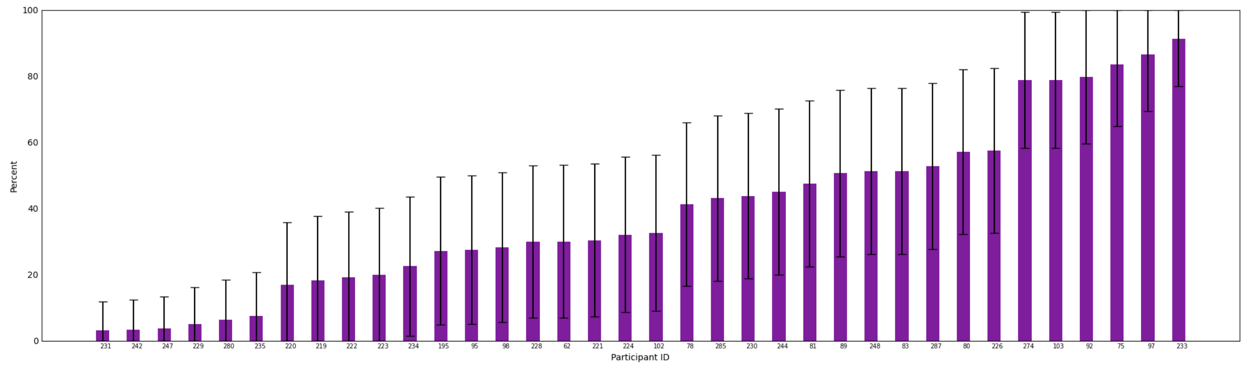


Figure 2

Bar graph representing the percentage amount of lies across participants

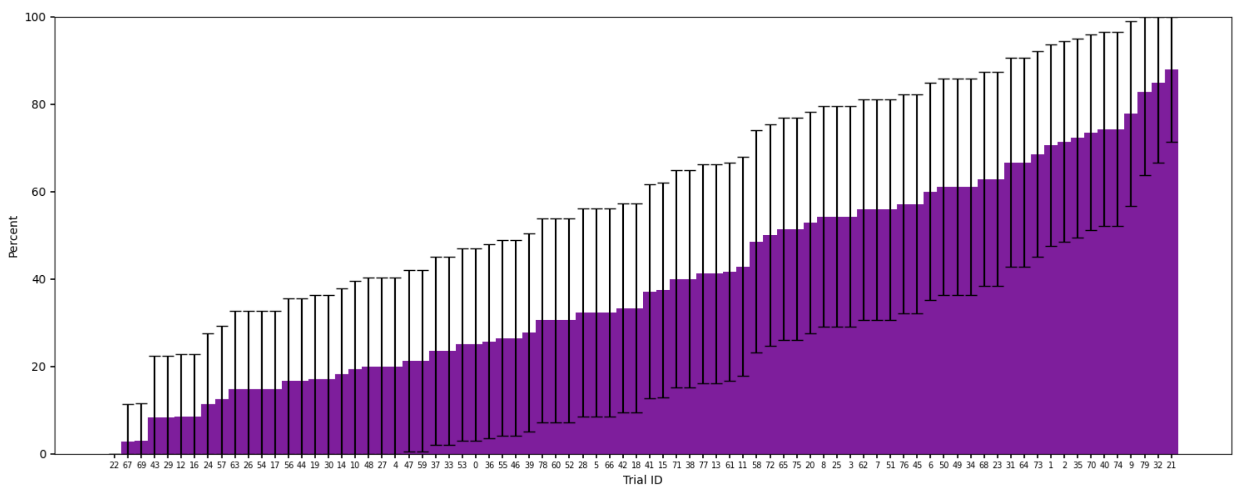


Figure 3

Bar graph showing the percentage of lies for each trial condition.

Dwell Time Distribution

Reaction time (or the time it takes to make a decision) showed a large amount of range. The average reaction time was 6324 ms ($SD = 3059$ms) and ranged from 2114ms - 22333ms,

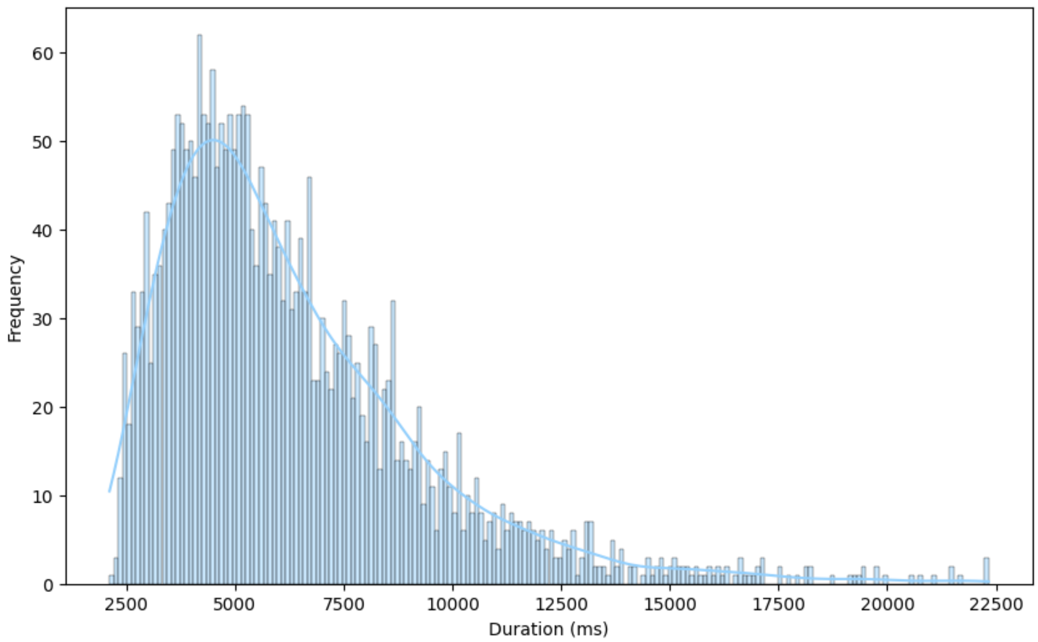


Figure 4

Distribution of reaction time ie. time taken to make a decision.

The amount of spent dwelling on one of the four AOIs on the screen was considered. Values ranged from 4ms - 7082ms. The AOIs, SELF LIE ($M = 381$ms, $SD = 382$ms), SELF LIE ($M = 350$ms, $SD = 328$ms), OTHER LIE ($M = 455$ms, $SD = 411$ms) and OTHER TRUTH ($M = 477$ms, $SD = 416$ms) all recorded means within 100ms of one another. A large proportion of the trials reported no dwell time for at least one of the given AOIs. Both SELF TRUE and SELF LIE were absent from the dwell timelines of 477 and 464 trials respectively, representing an omission rate of almost 20%. OTHER TRUE and OTHER LIE were omitted from 218 and 220 respectively representing just under 10% of all trials.

Figure 5

Distribution of average dwell time for each AOI. The distribution only accounts for average dwell time of each AOI that was under 1000ms. Each AOI also had a proportion of average dwell times that were 0, which were also omitted.

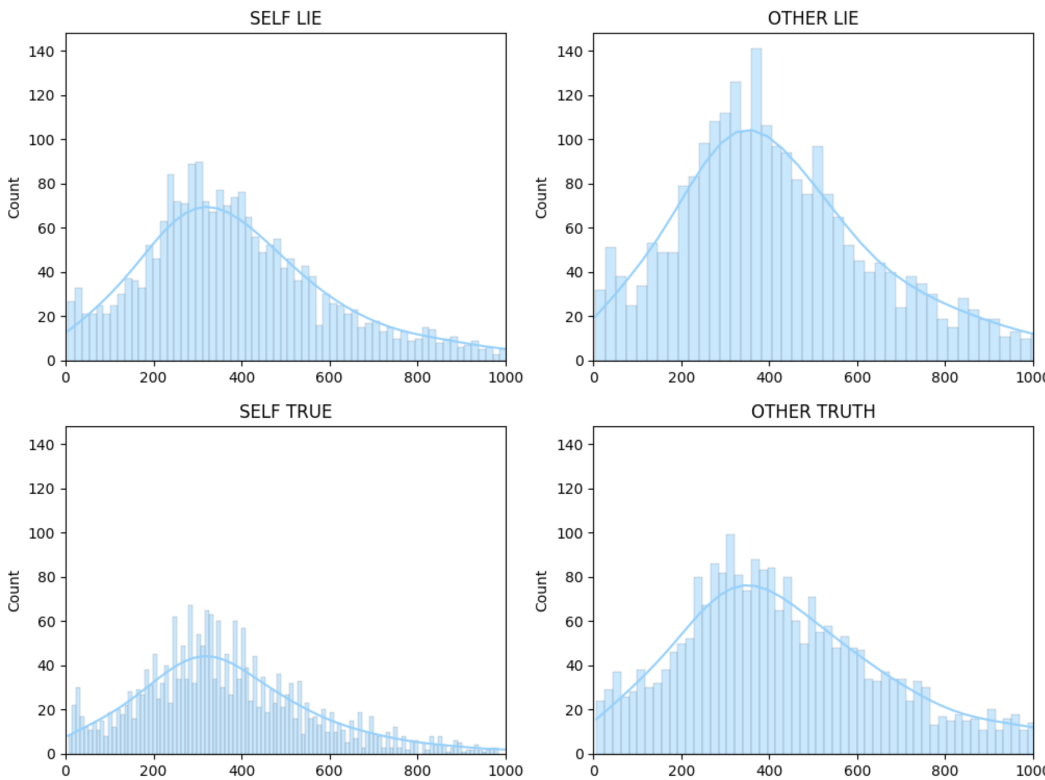
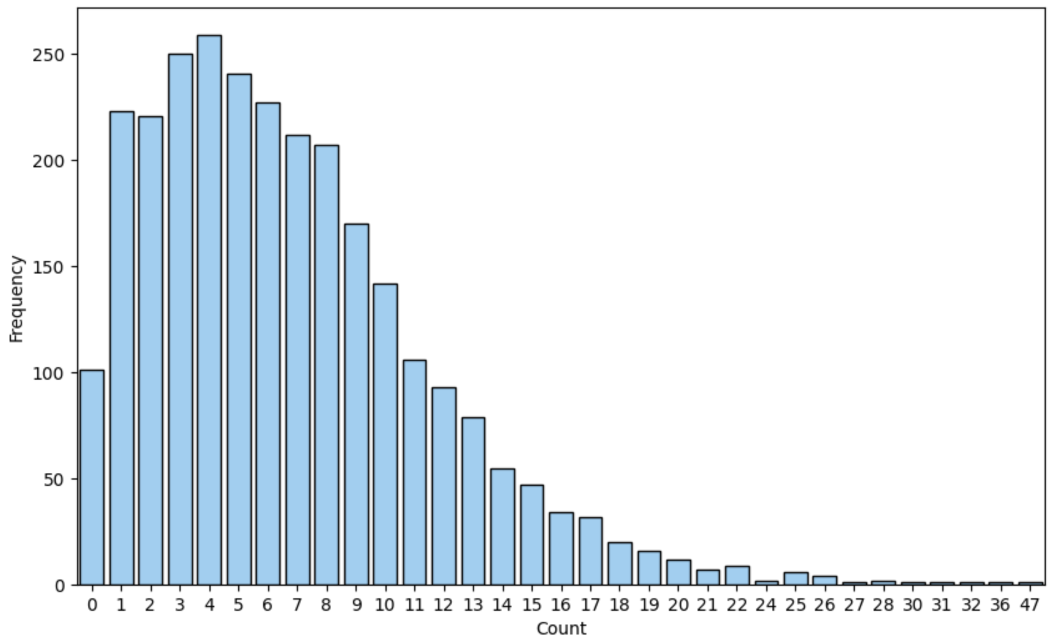


Figure 6

Distribution of number of transitions between AOIs across trials.



The number of transitions where a participant would move the cursor from one AOI to another was also observed. The mean number of transitions between AOIs was ($M = 455$ms, $SD = 411$ms).

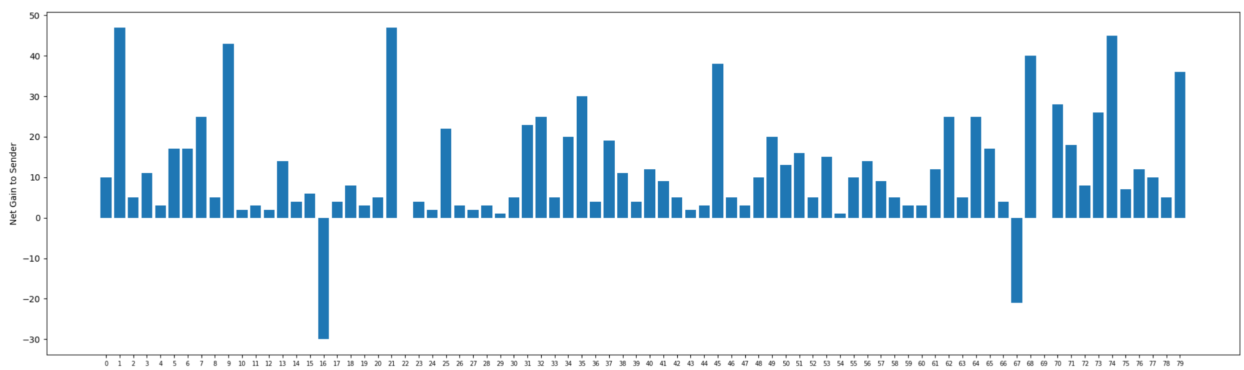
Task Analysis

Lie Percentage

Each of the trials involved giving differing amounts to the sender and the receiver based on whether the sender decided to lie or not. The average net gain to the sender was 12 ($SD = 13$) per trial and the average loss to the receiver was 11 ($SD = 9$).

Figure 7

Bar chart showing the amount gained by the sender should they choose to lie in the corresponding trial.



The percentage of senders that chose to lie in the experiment increased as the net gain to the sender did. Senders lied the least when they incurred a net loss by lying ($M = 6\%$, $SD = 23\%$). From net gains below 10 ($M = 26\%$, $SD = 44\%$) to above 10 ($M = 55\%$, $SD = 50\%$), the percentage who chose to lie increased. When gains were above 20 ($M = 68\%$, $SD = 46\%$) and above 30 ($M = 73\%$, $SD = 44\%$), the amount who decided to lie showed further increase but to a lesser extent.

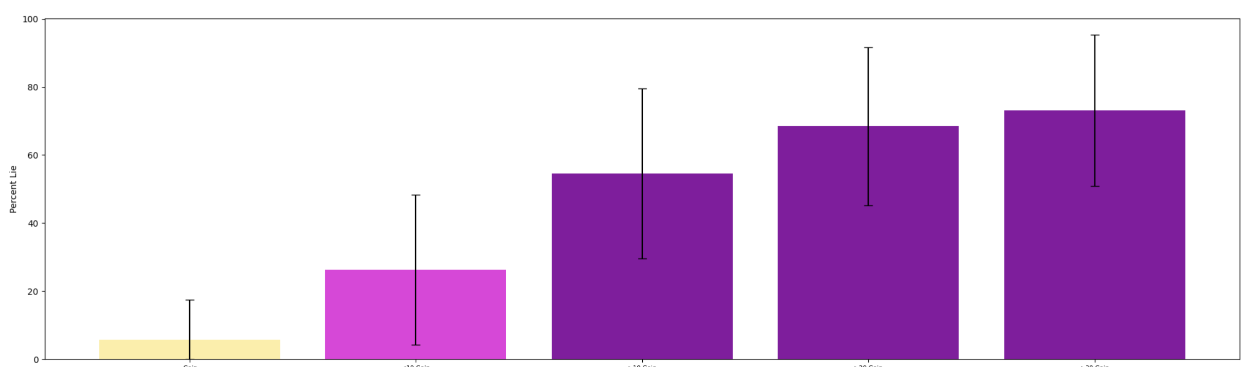


Figure 8

Percentage who chose to lie based on the amount gained by lying.

A significant difference in lie percentage was also observed across trials where the sender gained less than 10 and when they received more than 10, $t(2781) = 15.9$, $p < .001$.

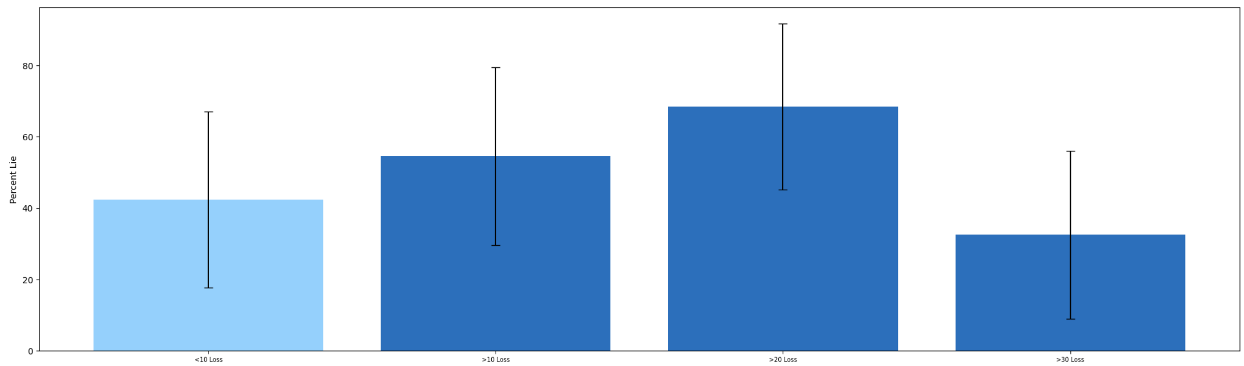


Figure 9

Percentage of senders who chose to lie based on how much the receiver would lose.

Senders were less likely to lie when the losses to the receiver were great. From losses under 10 ($M = 42\%$, $SD = 49\%$), losses greater than 10 ($M = 55\%$, $SD = 50\%$) through to losses greater than 20 ($M = 68\%$, $SD = 47\%$) the percentage of senders who chose to lie showed a notable increase. When losses to the receiver were larger than thirty the amount of senders who lied showed a notable decrease ($M = 32\%$, $SD = 47\%$). The difference in lie percentage between trials where losses to the receiver were over 10 and trials where losses to the receiver were under 10 was shown to be significant $t(2781) = 2.5$, $p = .001$.

Dwell Time

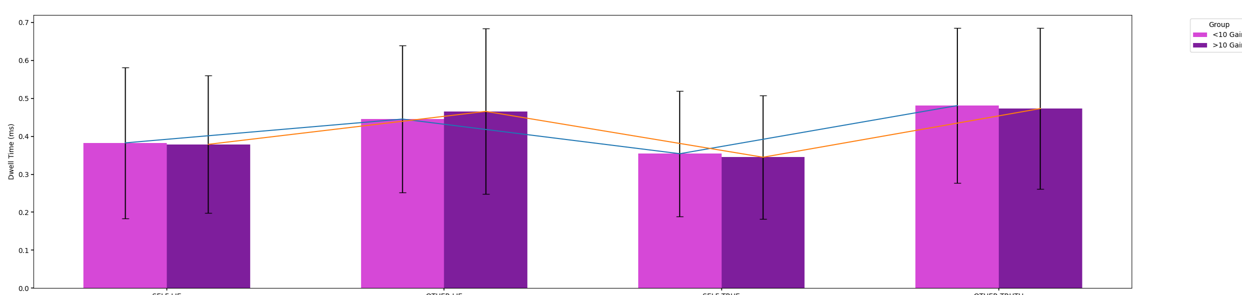


Figure 10

Average Dwell Time for each AOI comparison between trials where net gain to the Sender was less than 10 and trials where the net gain was more than 10

	<10 Gain to Sender		>10 Gain to Sender		T-test result	
	M (ms)	SD (ms)	M (ms)	SD (ms)	$t(2781)$	p
SELF LIE	0.38	0.40	0.38	0.36	-0.26	.793
SELF TRUE	0.35	0.33	0.34	0.33	-0.74	.460
OTHER LIE	0.45	0.39	0.47	0.44	1.3	.390
OTHER TRUTH	0.45	0.39	0.47	0.44	-0.50	.622

Table 1

Comparison of mean and standard deviation of dwell times for each AOI across trials where the net gain to Sender was less than 10 versus trials where the net gain was more than 10. (FDR Corrected)

Table 1 shows that there was an no significant difference between the average dwell time of the four AOIs across trials where the net gain to the Sender was less than 10 versus when it was more than 10.

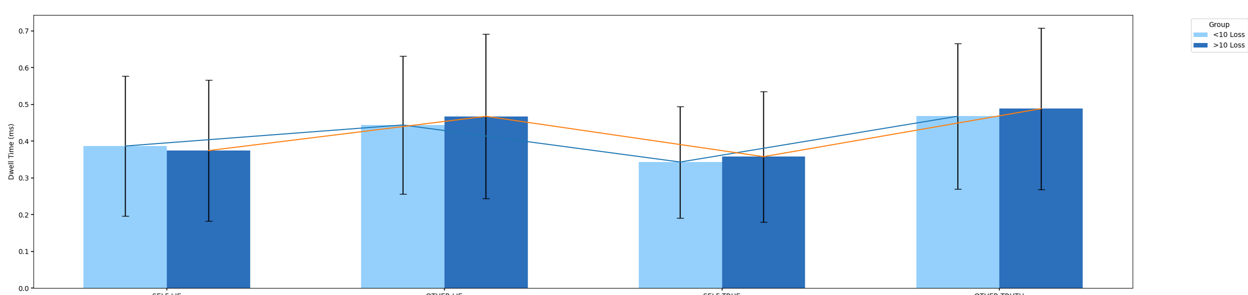


Figure 11

Average Dwell Time for each AOI comparison between trials where net loss to the Receiver was less than 10 and trials where the net loss was more than 10

	<10 Loss to Receiver		>10 Loss to Receiver		T-test result	
	M (ms)	SD (ms)	M (ms)	SD (ms)	$t(2781)$	p
SELF LIE	0.39	0.38	0.38	0.36	-0.86	.784
SELF TRUE	0.34	0.30	0.34	0.39	1.2	.459
OTHER LIE	0.44	0.38	0.47	0.44	1.5	.195
OTHER TRUTH	0.47	0.40	0.47	0.42	1.3	.387

Table 2

Comparison of mean and standard deviation of dwell times for each AOI across trials where the net gain to Sender was less than 10 versus trials where the net gain was more than 10. (FDR Corrected)

Similarly, in Table 2, there was no significant difference between the dwell times in trials where the loss to the Receiver was above 10 versus trials where it was below 10.

Number of Transitions

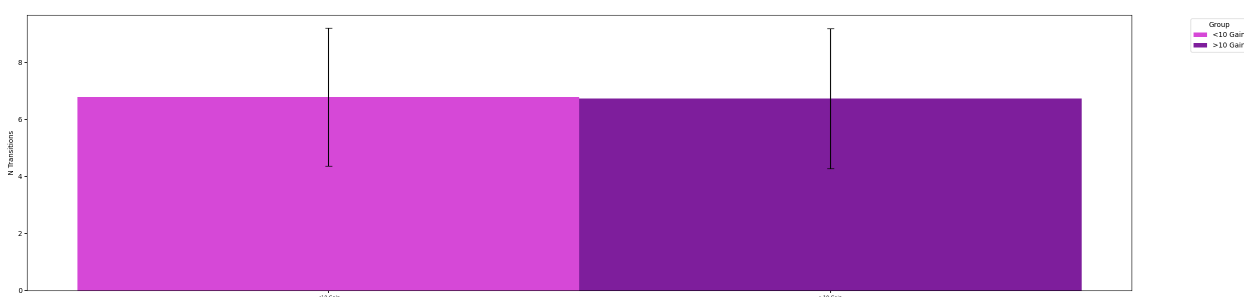


Figure 12

Comparison of average number of transitions between trials where the net gain to the Sender was less than 10 and trials where the net gain was more than 10

The difference in number of transitions proved not to be significant when comparing trials where the net gain to the Sender is above ($M = 42\%$, $SD = 49\%$) or below 10 ($M = 42\%$, $SD = 49\%$), $t(2781) = 2.5$, $p = .001$.

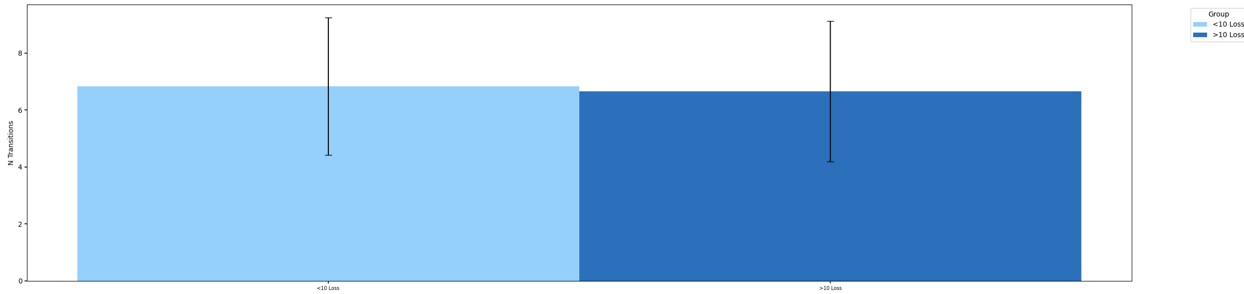


Figure 13

Comparison of average number of transitions between trials where net loss to the Receiver was less than 10 and trials where the net loss was more than 10

Similarly the difference in number of transitions showed to be insignificant when comparing trials where the Receiver made a loss of above ($M = 6.7\%$, $SD = 4.9\%$) or below 10 ($M = 6.8\%$, $SD = 4.8\%$) from lying, $t(2781) = -0.99$, $p = .648$.

Participant Analysis

An analysis was performed on the distance data extracted from the time series analysis at the participant level. The results of which, are taken from averaging the distance between all of the trials of a single participant against another. Through hierarchical clustering and silhouette analysis, the participants were separated in to two clusters.

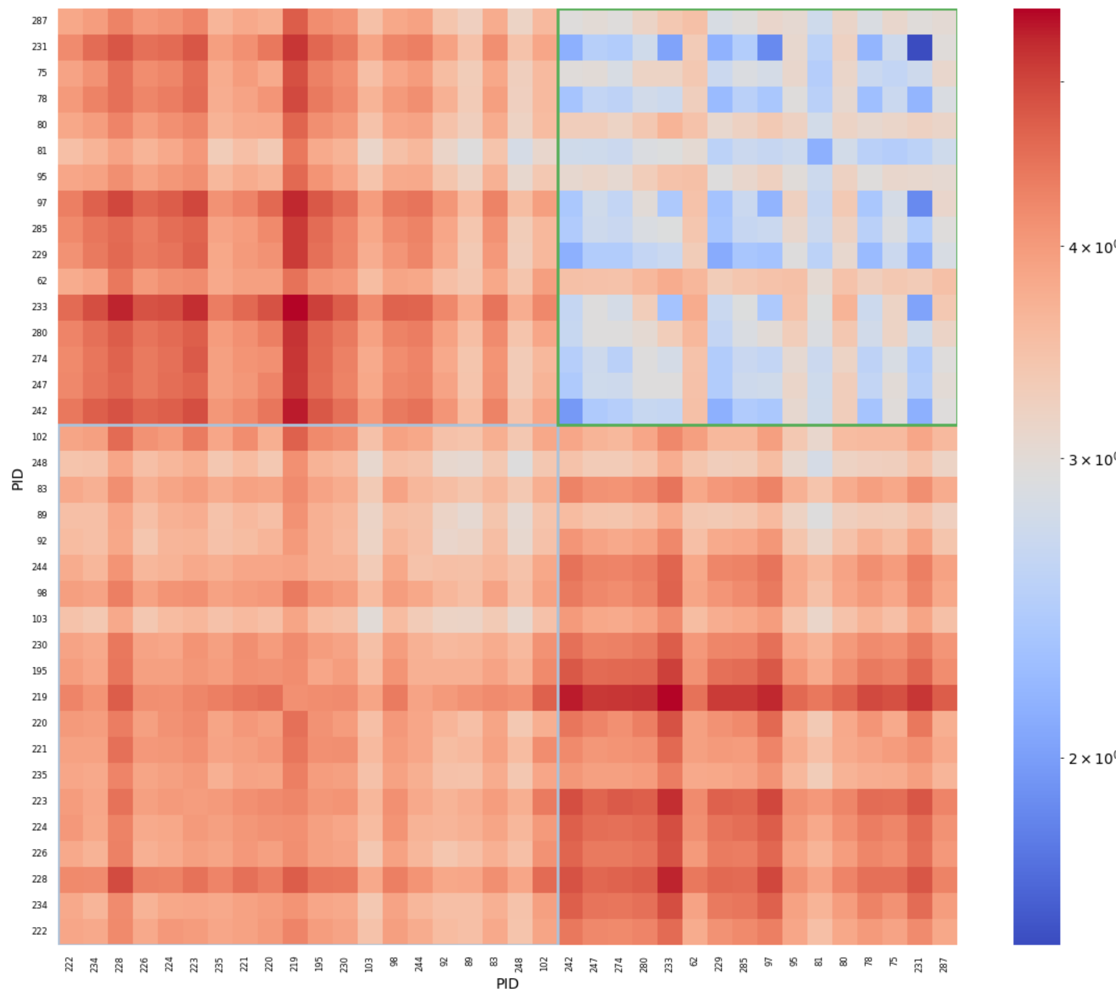


Figure 14

Matrix showing the similarity between participants taken from their mean euclidean distance measure shared with one and another. The matrix is sorted by proximity with the two clusters of participants outlined.

Lie Percentage

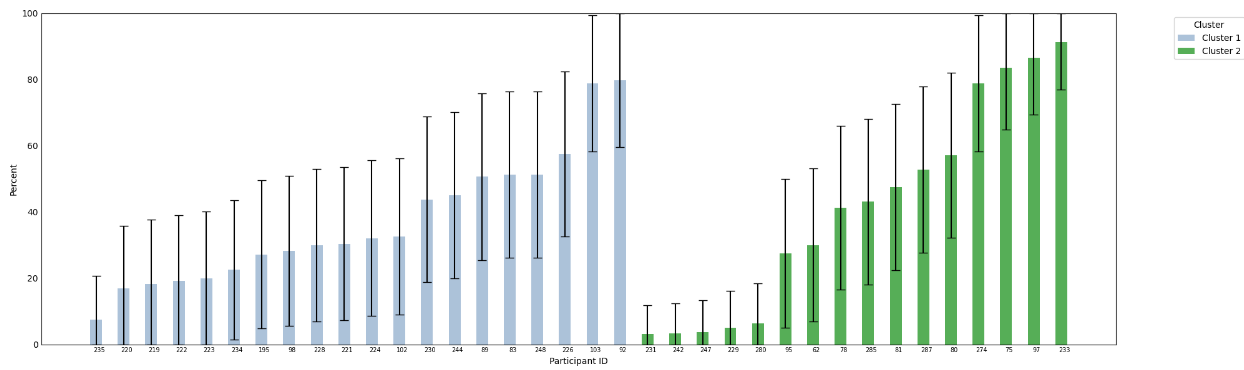


Figure 15

Participants separated by hierarchical distance clustering and ordered by percentage lies.

Using pairwise t-tests and calculating the mean statistic across clusters, it was deemed that the two clusters and their participant memberships showed to have a difference in lie percentage that was not significant $t(34) = 1.6$, $p = .183$ (FDR corrected). Cluster 1 had a mean lie percentage of 37% ($SD = 48\%$) and cluster 2 had a mean of 42% ($SD = 49$). There was no significant difference within clusters either, for Cluster 1 $t(18) = 1.1$, $p = .188$ and Cluster 2 $t(14) = 1.4$, $p = .183$ (FDR corrected).

Dwell Time

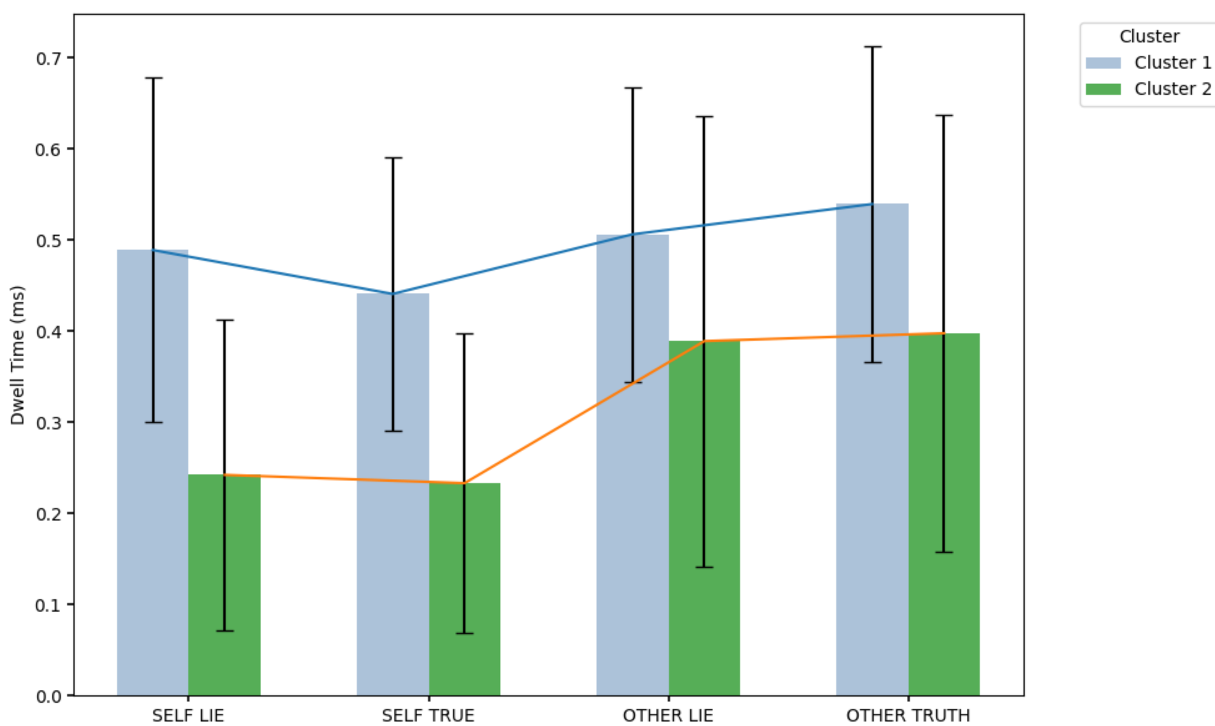


Figure 16

Measures of average dwell time across participant clusters segmented by hierarchical distance clustering.

	Cluster 1				Cluster 2				Between	
	M (ms)	SD (ms)	$t(20)$	p	M (ms)	SD (ms)	$t(18)$	p	$t(34)$	p
SELF LIE	0.49	0.38	0.17	0.215	0.24	0.34	1.5	0.14	1.9	.126
SELF TRUE	0.44	0.30	0.19	0.188	0.23	0.33	1.3	0.204	1.5	.081
OTHER LIE	0.51	0.32	0.12	0.239	0.39	0.49	1.1	0.159	0.19	.211
OTHER TRUTH	0.54	0.35	-0.07	0.184	0.40	0.48	-0.53	0.184	-0.15	.184

Table 3

Mean,. standard deviation and t tests for the average dwell time of each AOI across participant clusters. (FDR corrected)

In Table 3, average dwell times across the four AOIs show no significant differences when compared between and within clusters of participant membership.

Number of Transitions

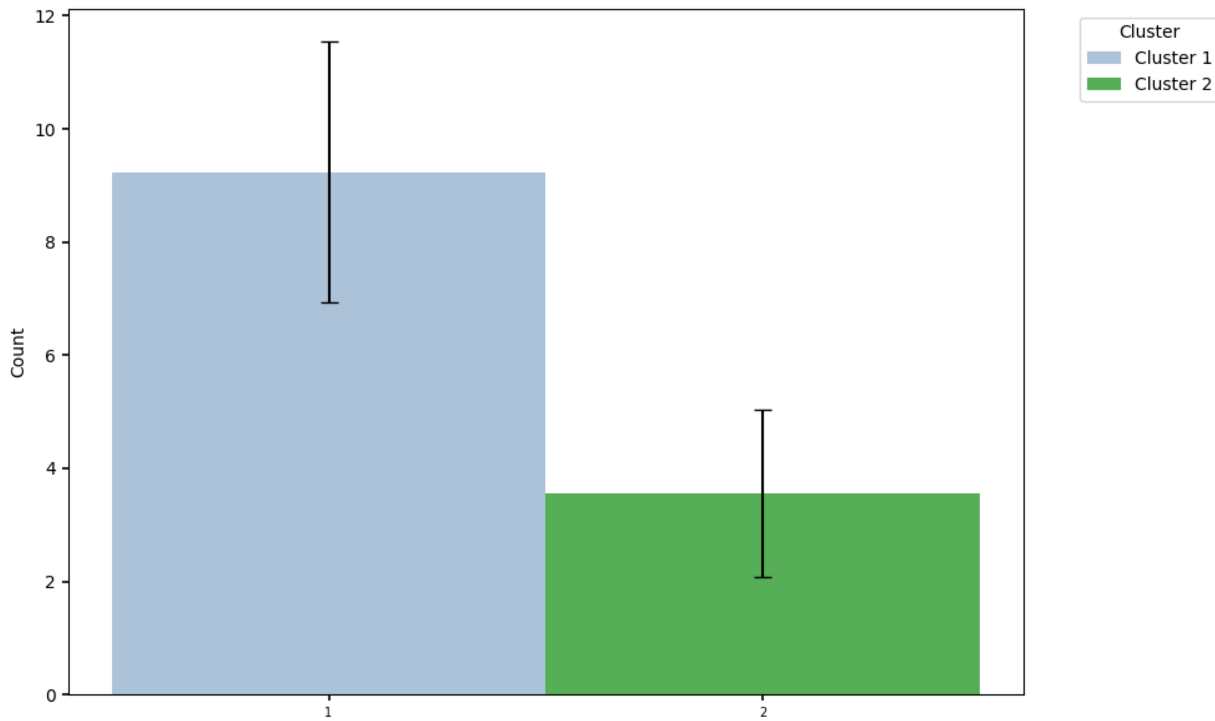


Figure 17

Average number of transitions between clusters of participants determined from hierarchical clustering.

The number of transitions averaged by the participants in Cluster 1 ($M = 9.2$, $SD = 4.6$), showed to be significantly different to that of the participants in Cluster 2 ($M = 3.6\%$, $SD = 2.9\%$), $t(34) = 2.0$, $p = .042$. Comparatively, there was no significant difference in number of transitions within both Cluster 1 $t(34) = -0.75$, $p = .188$ and Cluster 2 $t(34) = 1.5$, $p = .137$ (FDR corrected).

Cross Analysis

Lie Percentage

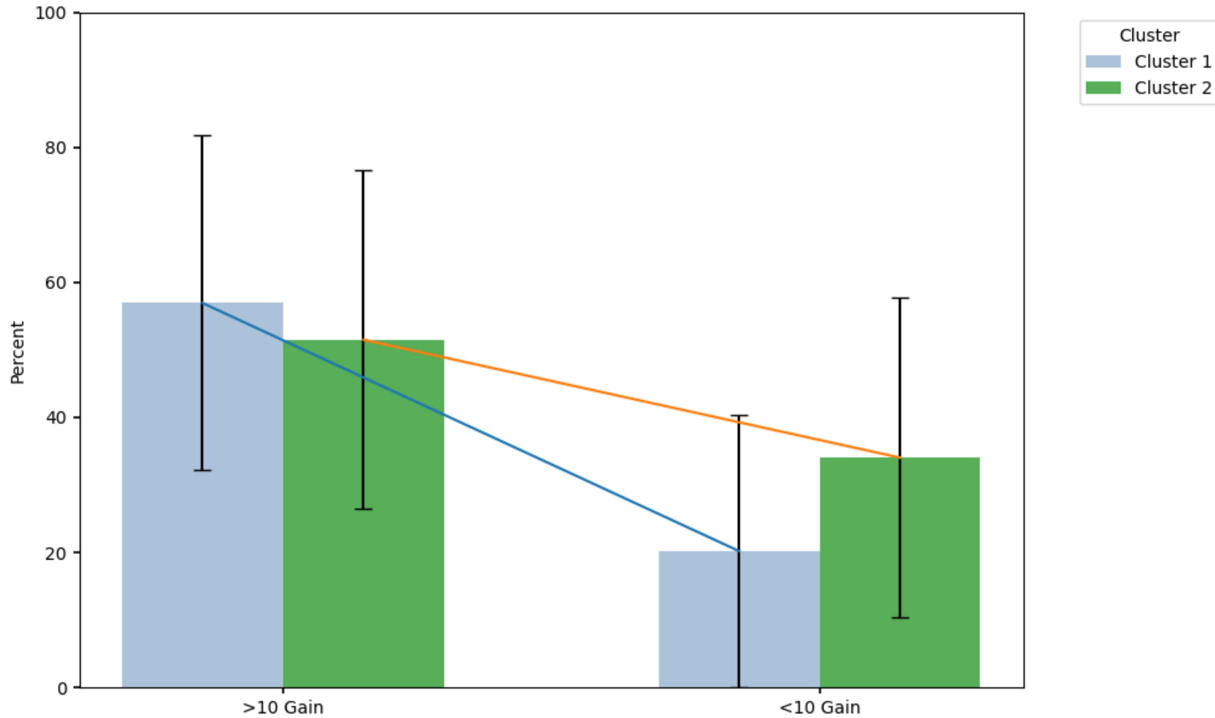


Figure 18

Average lie percentage for each cluster of participants across trial conditions dependent on net gain to sender.

	<10 Gain to Sender		>10 Gain to Sender		T-test result	
	M (%)	SD (%)	M (%)	SD (%)	$t(66)$	p
Cluster 1	20	40	56	49	-0.26	.793
Cluster 2	34	47	52	50	-0.74	.460

Table 4

Comparison of mean and standard deviation of average percent lie for each cluster across trial conditions (FDR Corrected)

Dwell Time

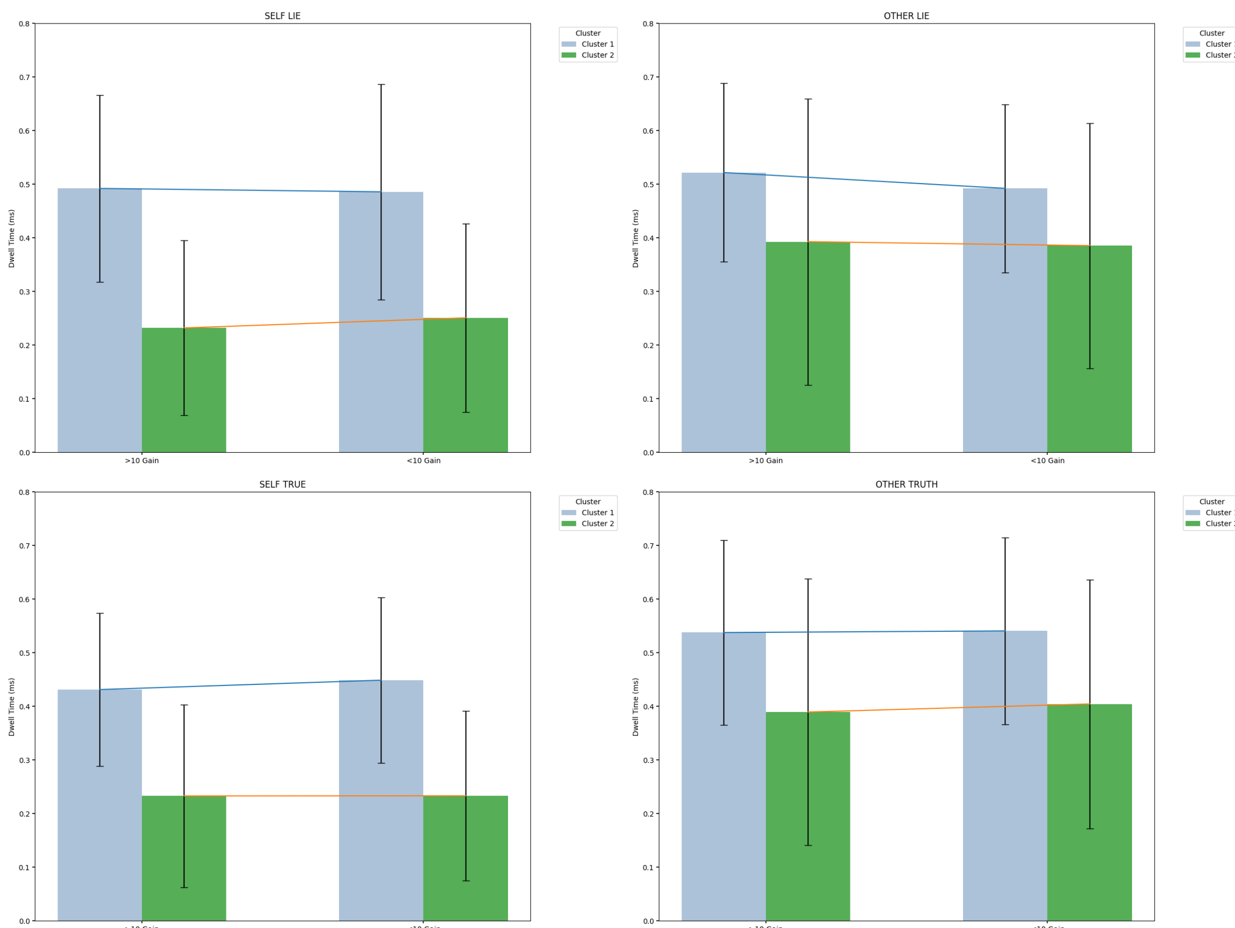


Figure 19

Measures of average dwell time across participant clusters segmented by hierarchical distance clustering.

	Gain to Sender < 10				Gain to Sender > 10					
	Cluster 1		Cluster 2		Cluster 1		Cluster 2		T-test result	
	M (ms)	*SD* (ms)	*M* (ms)	*SD* (ms)	*M* (ms)	*SD* (ms)	*M* (ms)	*SD* (ms)	$t(34)$	p
SELF LIE	0.49	0.38	0.24	0.34	0.49	0.38	0.24	0.34	1.9	.126
SELF TRUE	0.44	0.30	0.23	0.33	0.49	0.38	0.24	0.34	1.5	.081
OTHER LIE	0.51	0.32	0.39	0.49	0.49	0.38	0.24	0.34	0.19	.211
OTHER TRUTH	0.54	0.35	0.40	0.48	0.49	0.38	0.24	0.34	- 0.15	.184

Table 5

Mean and standard deviation of the dwell time of each AOI across participant clusters.

Number of Transitions

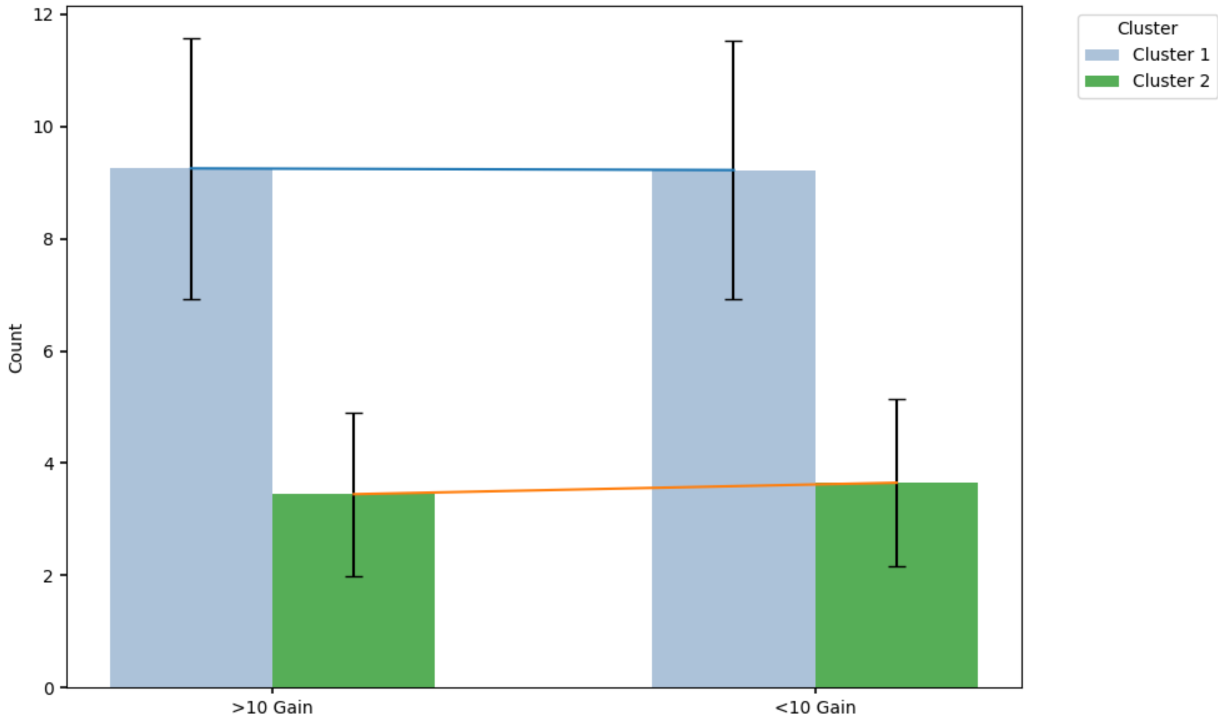


Figure 20

Measures of average dwell time across participant clusters segmented by hierarchical distance clustering.

Discussion

References

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without tradeoffs. Psychological Review, 113, 409-432. *Psychological review*, *113*, 409–32.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, *16*(2), 366–387.

Payne, J. W., & Bettman, J. R. (2004, January). Walking with the Scarecrow: The Information-Processing Approach to Decision Research. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (1st ed., pp. 110–132). Wiley.

Reeck, C., Wall, D., & Johnson, E. J. (2017). Search predicts and changes patience in intertemporal choice. *Proceedings of the National Academy of Sciences*, *114*(45), 11890–11895.