Two’s company but six is a crowd: emergence of conventions in multiparty communication games

CogSci 2022

Veronica Boyce, Robert Hawkins, Noah Goodman & Mike Frank



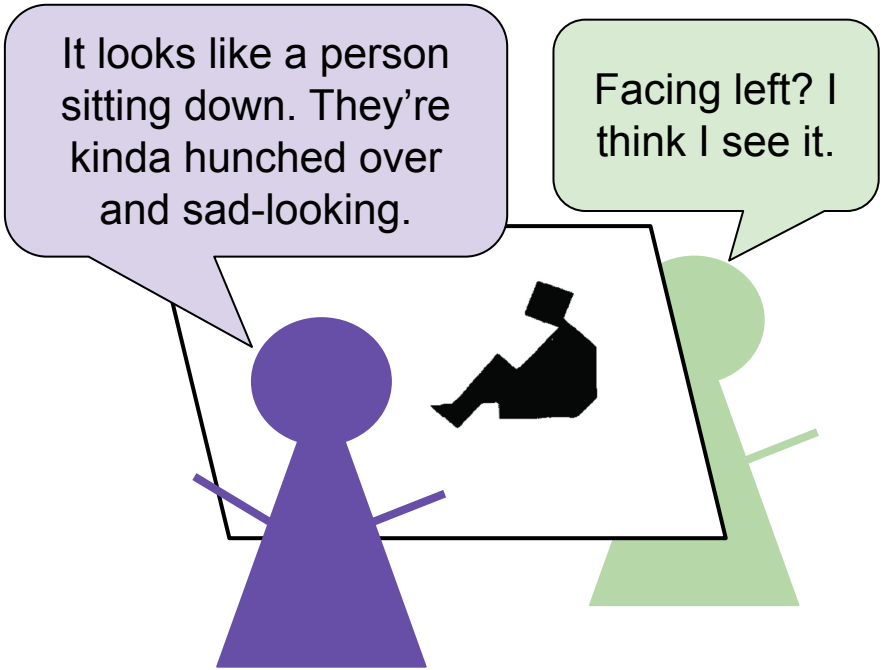

Communication often happens in groups!

We typically study dyadic communication

- Most prior work has used **dyadic** referential communication tasks
  Krauss & Weinheimer 1964, 1966
  Clark & Wilkes-Gibbs 1986
  Hawkins et al. 2020

- 3-5 person groups primarily studied in context of listeners with **different backgrounds**
  Weber & Camerer 2003
  Yoon & Brown-Schmidt 2014, 2018, 2019

- Here we examine adaptive communication in **multi-party tasks**



Key phenomenon: Reduction over repeated reference

Developing a "nickname" for an image over repeated references with the same partner

1st round: I picture **someone wrapped in a blanket.** You cannot make out their legs or arms. Their upper body is up straight, and it looks like their knees are pointed slightly to the right

6th round: **bundle**

1st round: it is the one that doesn't really have anything poking out except at the bottom right is **a little triangle shaped almost like a boot**

6th round: **boot 1**

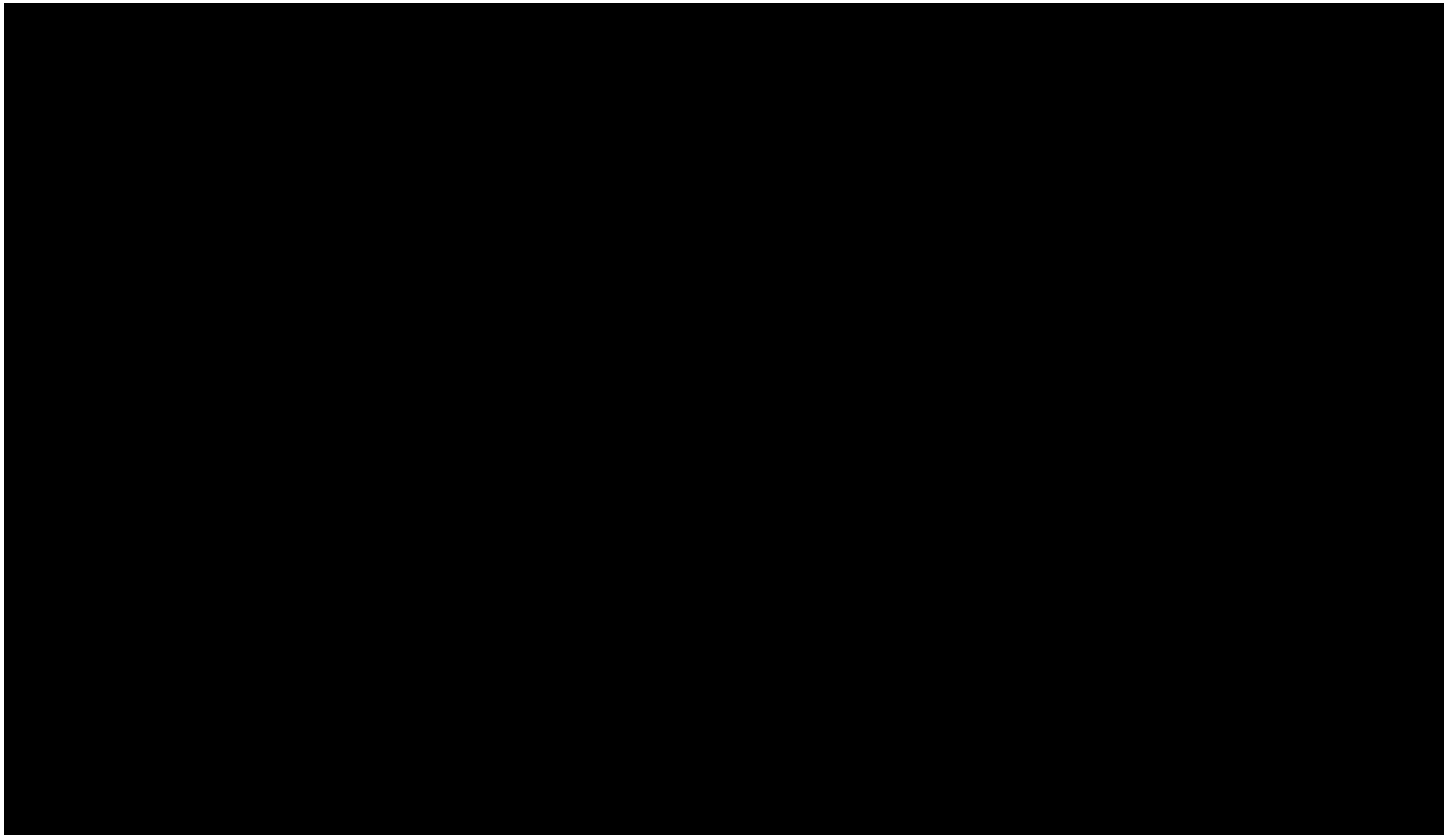
Methods

N=390 Prolific workers

Groups of 2-6 (~15 / size)

6 blocks of 12 targets each (72 trials total)

Rotated speaker between blocks



Empirica platform (Almaatouq et al, 2021)

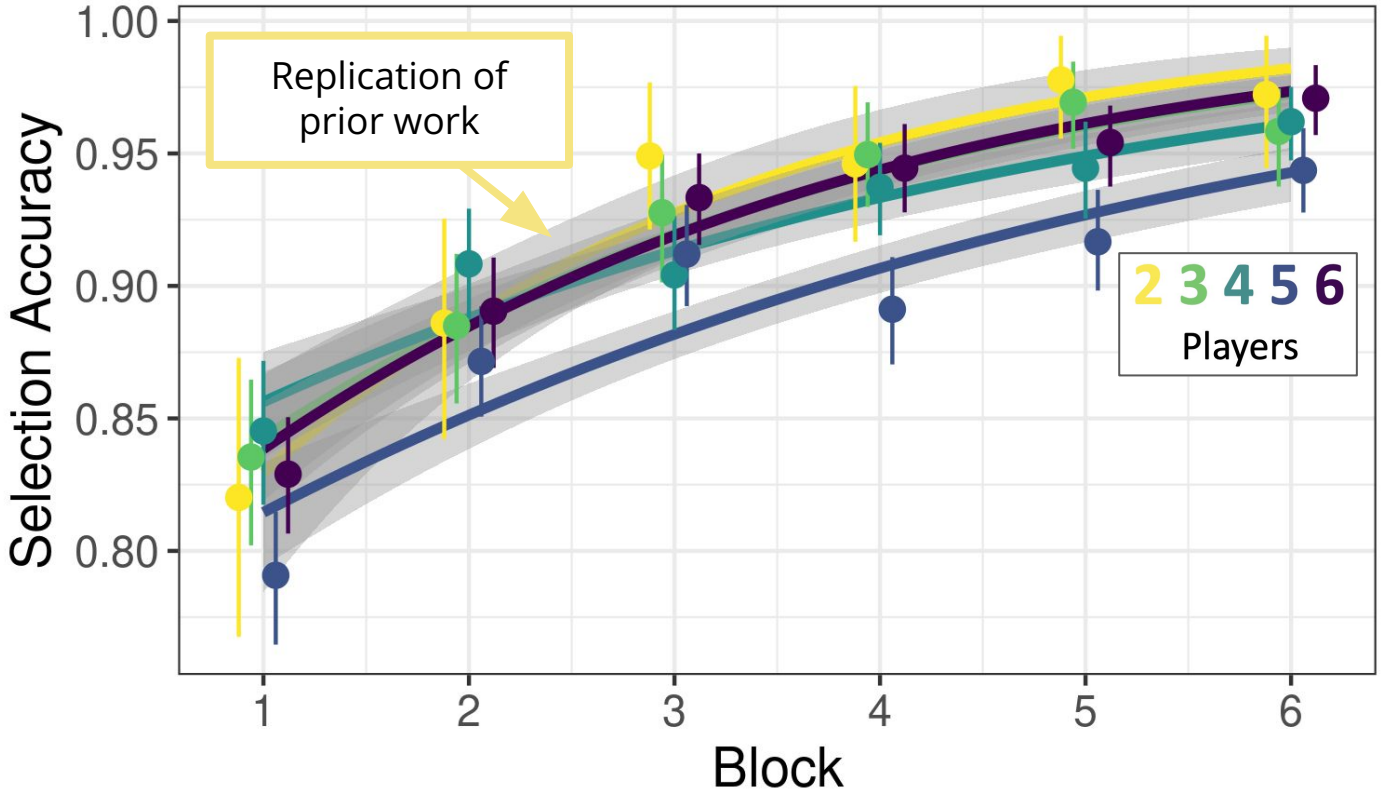
How does group size affect convention formation?

Does accuracy increase?

How does group size affect reduction?

How do conventions emerge in larger groups?

Accuracy is high and increasing



How does group size affect convention formation?

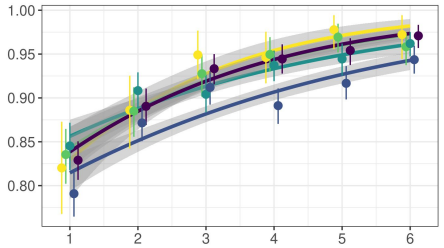
Does accuracy increase?

How does group size affect reduction?

How do conventions emerge in larger groups?

How does group size affect convention formation?

Does accuracy increase? Yes

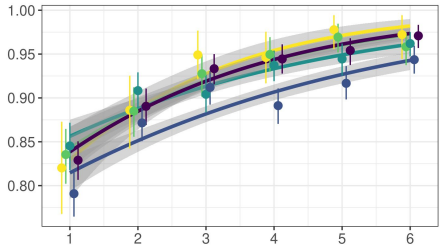


How does group size affect reduction?

How do conventions emerge in larger groups?

How does group size affect convention formation?

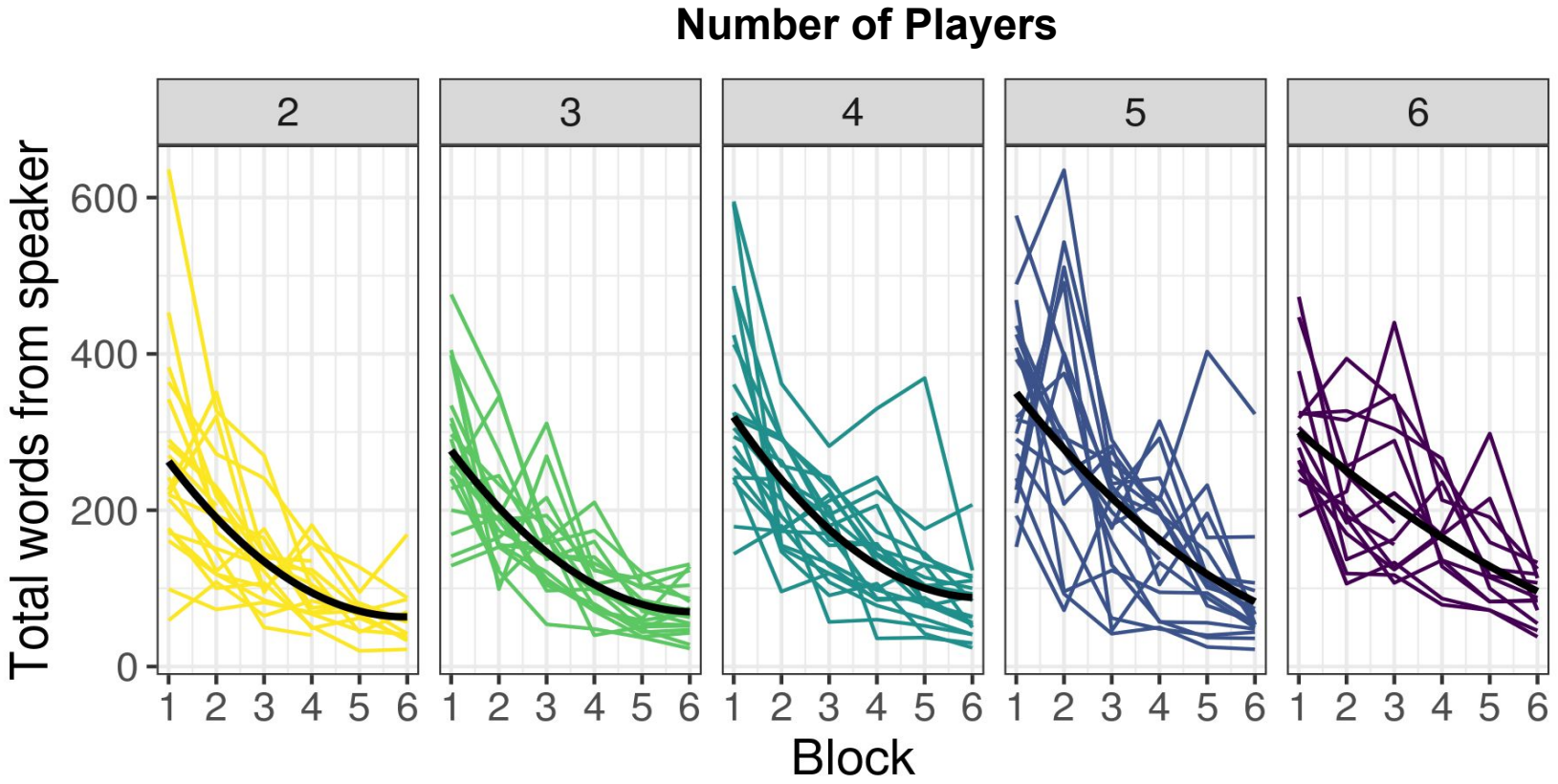
Does accuracy increase? Yes



How does group size affect reduction?

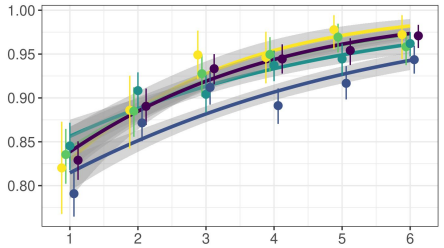
How do conventions emerge in larger groups?

Speakers say less in later blocks



How does group size affect convention formation?

Does accuracy increase? Yes

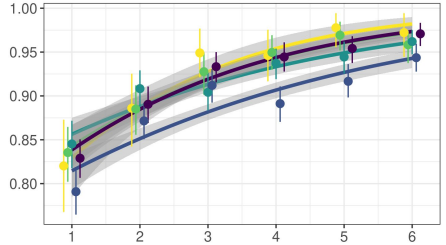


How does group size affect reduction?

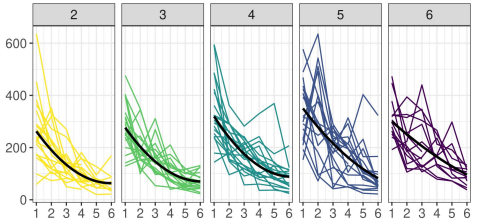
How do conventions emerge in larger groups?

How does group size affect convention formation?

Does accuracy increase? Yes



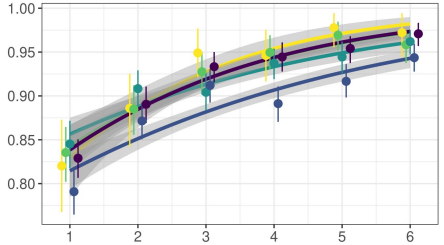
How does group size affect reduction? Bigger groups say a bit more



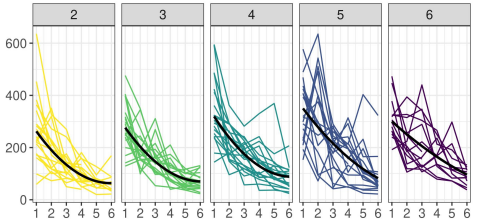
How do conventions emerge in larger groups?

How does group size affect convention formation?

Does accuracy increase? Yes



How does group size affect reduction? Bigger groups say a bit more

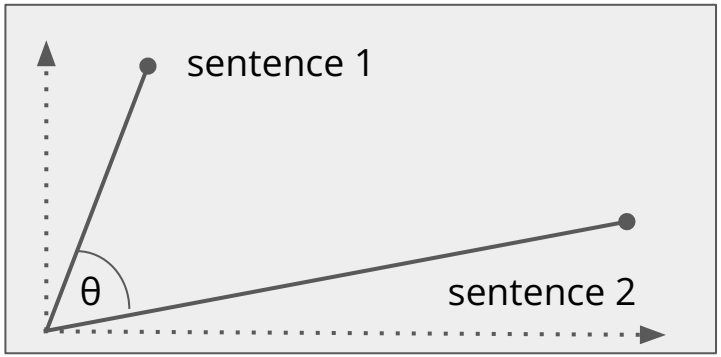


How do conventions emerge in larger groups?

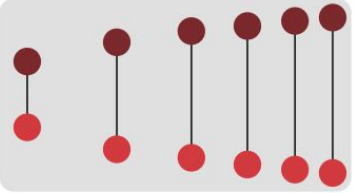
Measuring similarity with sentence embeddings

Turn utterances into vectors with **SBERT** sentence embeddings (Reimers & Gurevych 2019)

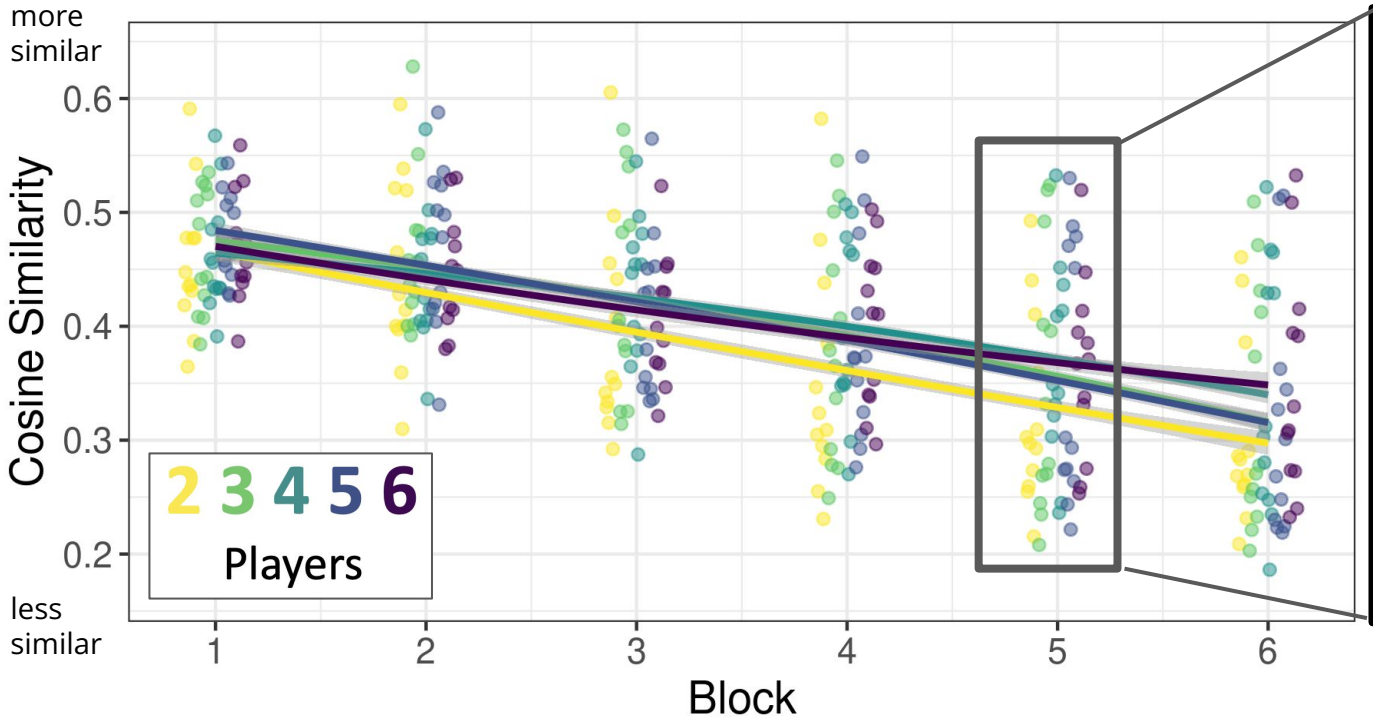
Compute pairwise cosine similarity (smaller angle = bigger cosine)



(4) penguin stomping feet, dinosaur, square head on the right	(6) stomping penguin dinosaur	.83
(4) penguin stomping feet, dinosaur, square head on the right	(5) dinosaur, head tilted backward right	.62
(4) penguin stomping feet, dinosaur, square head on the right	(2) head tilted backward, one leg raised little standing position	.37
(2) head tilted backward, one leg raised little standing position	(6) stomping penguin dinosaur	.14

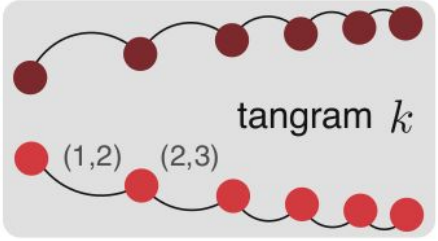


Descriptions diverge across groups

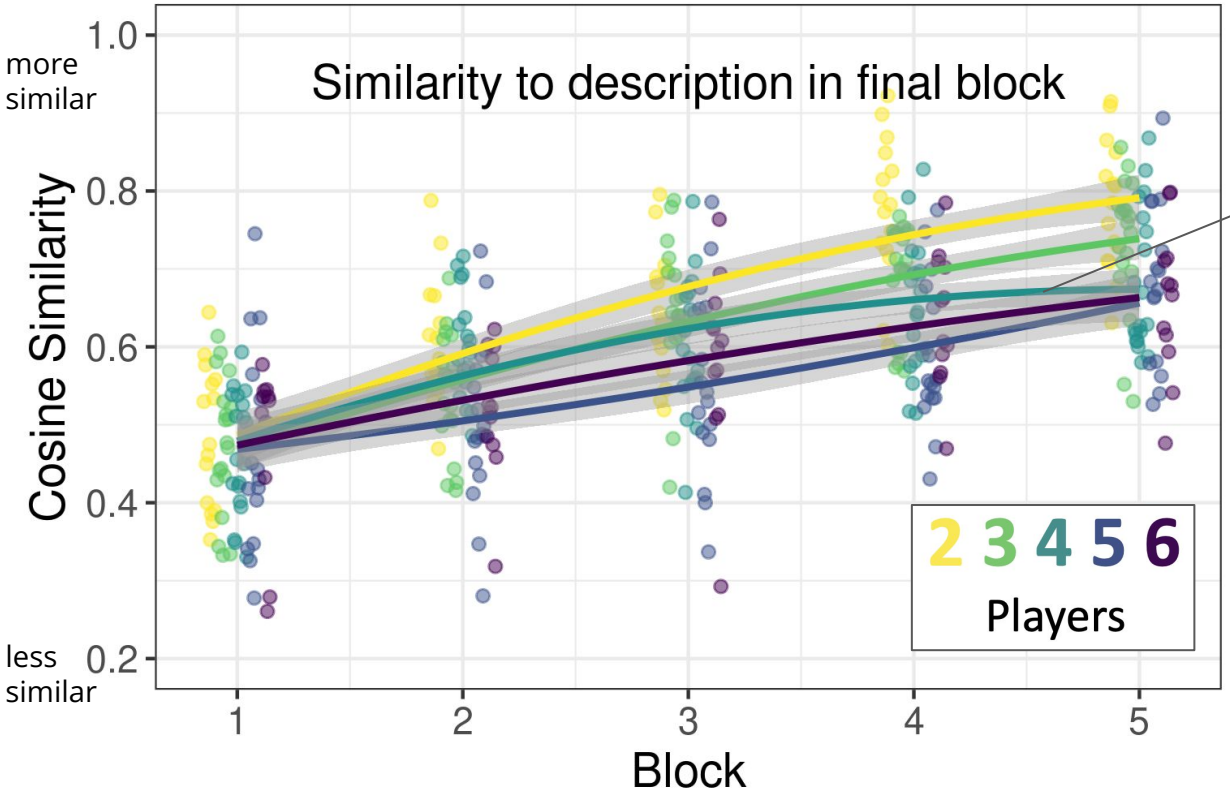


6-player games, block 5:

```
slide tackle
footballer

abstract one

the same zombie one
one arm up one down

The messy one square
and triangle on top.
```



Descriptions converge **slower for bigger groups**



Round 3: `It's the pope guy, [...]`

Round 4: `beggar or pope`

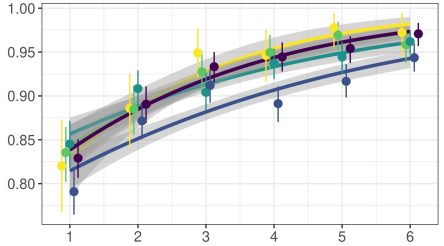
Round 5: `Pope`

Round 6: `Pope`

How does group size affect convention formation?

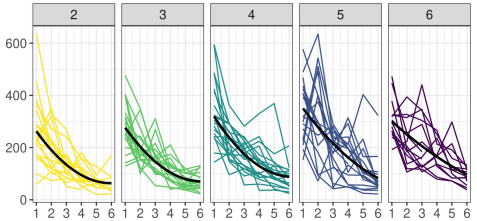
Does accuracy increase?

Yes



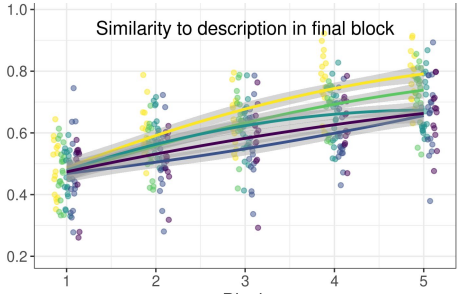
How does group size affect reduction?

Bigger groups say a bit more



How do conventions emerge in larger groups?

Convergence is slightly slower in larger groups



Questions?

More **people**

Differences in **incentives**

Differences in **knowledge**

materials, data, code: github.com/vboyce/FYP

contact: vboyce@stanford.edu

Demographics

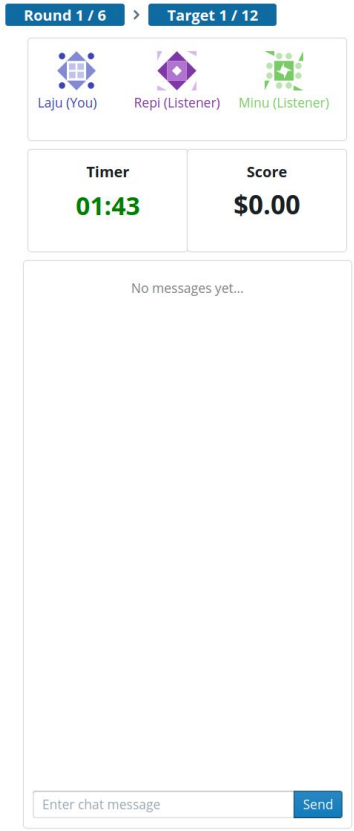
Most thought their partners were human (343 yes ,11 no)

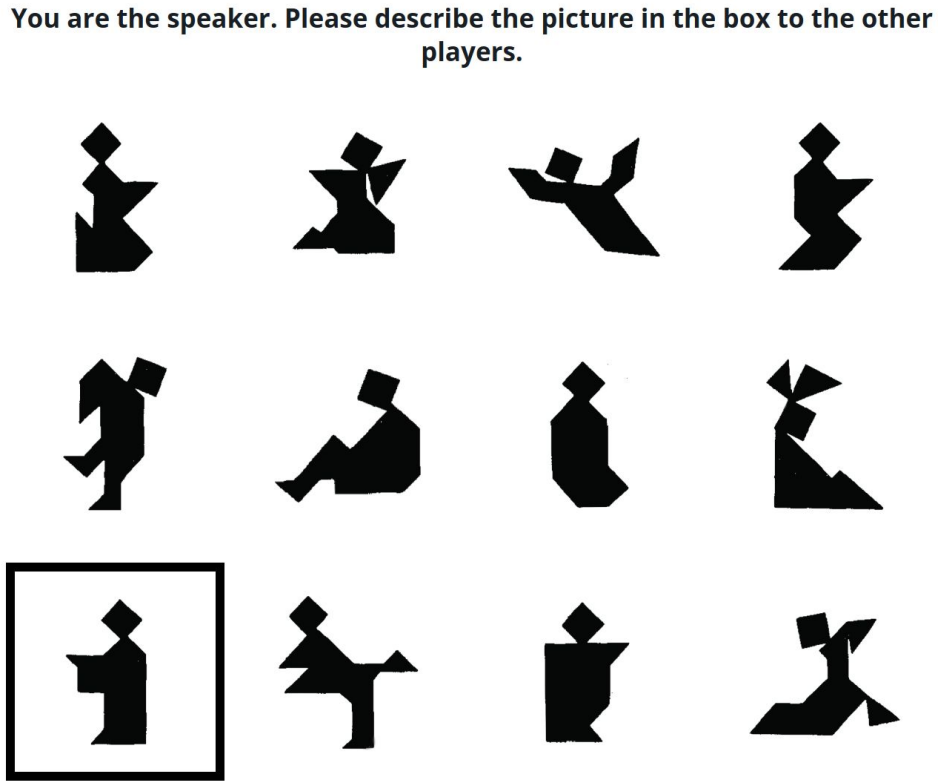
Gender: 164 female, 186 male, 4 non-binary

Age: 18-74, mean of 30, 50% in 24-39 range

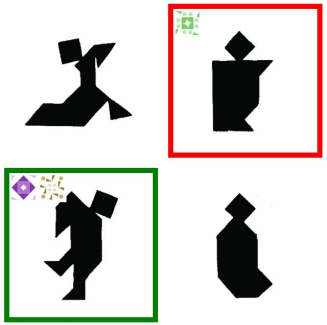
Race/ethnicity: 284 White, 40 Black, 39 Asian, 12 Hispanic, 1 Native American

Methods: Prolific & Empirica

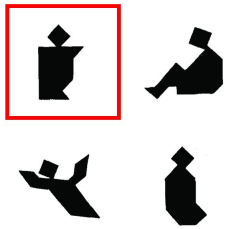




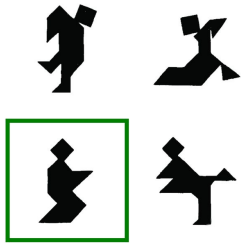
2/3 correct!



Incorrect!



Correct!



Almaatouq et al (2021)

Aimed for 20 games / condition.

Included all complete blocks, even if the game didn’t finish

Players	Complete	Partial
2	15	4
3	18	2
4	19	2
5	17	3
6	12	6

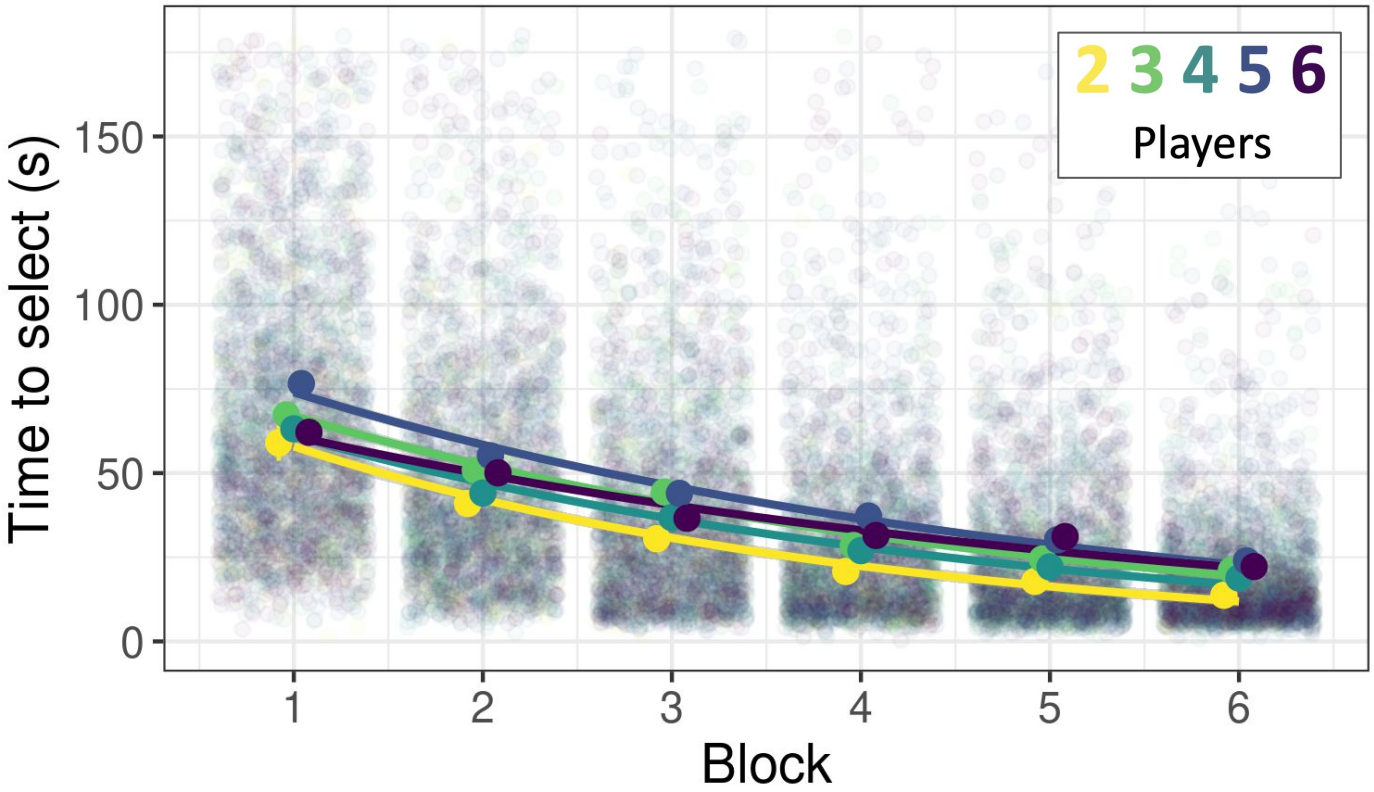
Statistical results:

correct.num ~ block * num.Players
block = .38 [.25, .5]; numPlayers=-.02 [-.08,.03]; block:numPlayers=-.01[-.04,.01]

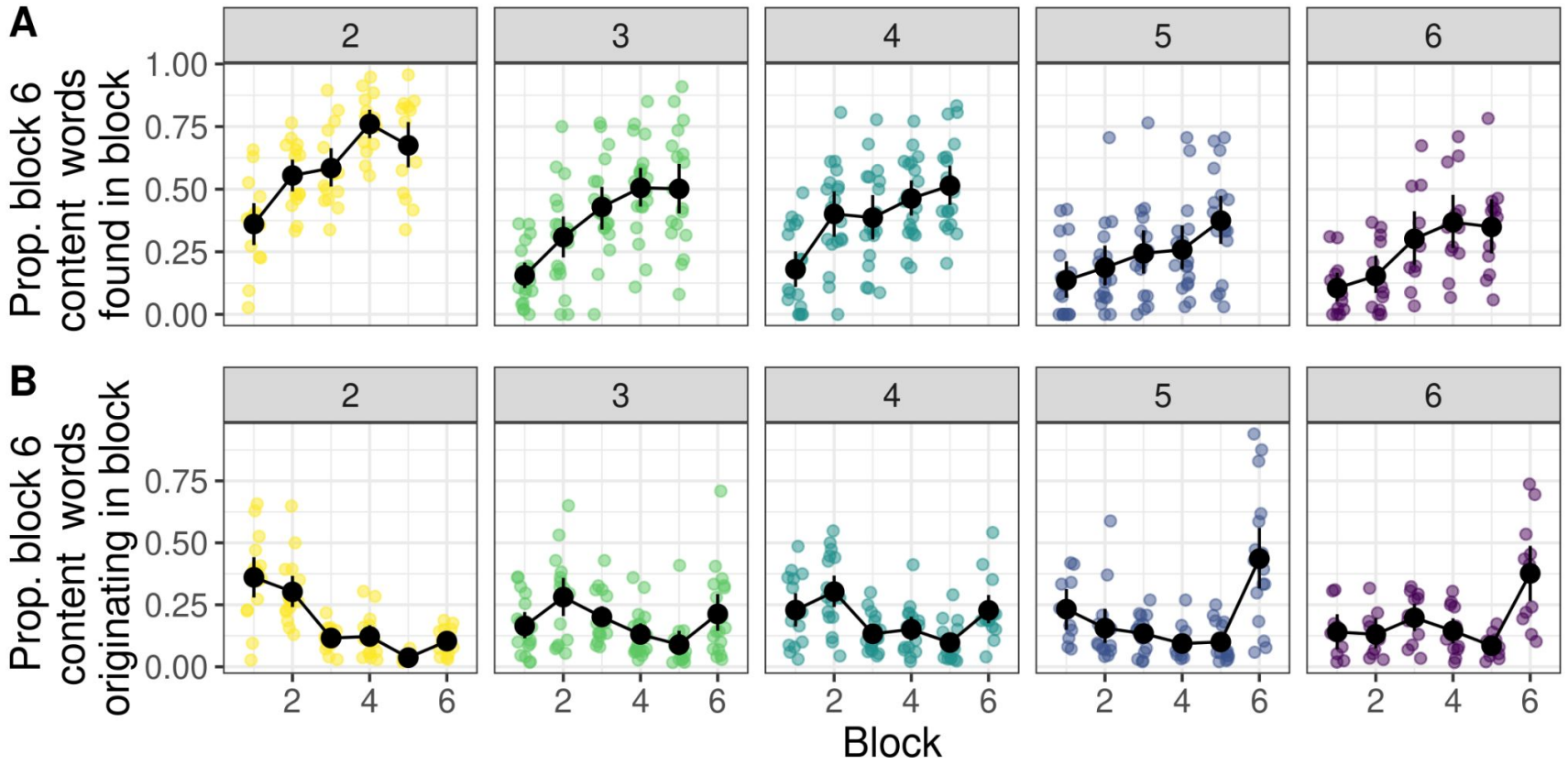
time ~ block * num.Players
block = -10.0 [-11.03, -9.03]; numPlayers=1.03 [.4,1.66]; block:numPlayers=.29 [.07,.5]

words ~ block * numPlayers + (block|tangram) + (1|playerId) + (1|tangram_group) + (block|gameId)
block=-3.35 [-4.58, -2.13]; numPlayers=1.67 [.68, 2.71]; block:numPlayers=-.1 [-.39, .18]

People choose faster in later rounds

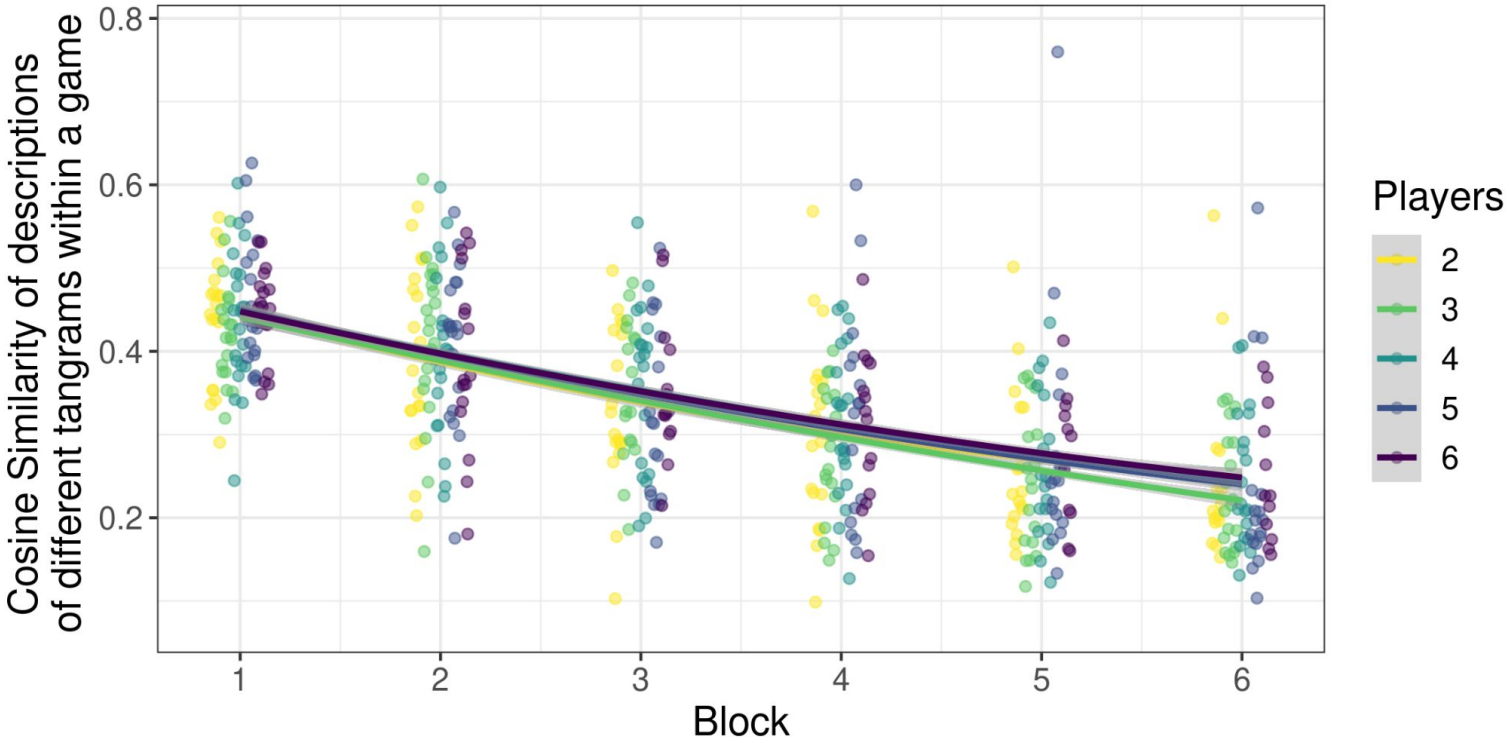


Nicknames originate earlier in smaller groups



overlap ~ block * numPlayers + same_speaker + (1 | gameId) + (1 | target)
block=.1[.08, .12] ; numPlayers=-.07 [-.09, -.04] ;block:numPlayers =0[-.01, 0]; same_speaker=.08 [.05, .1]

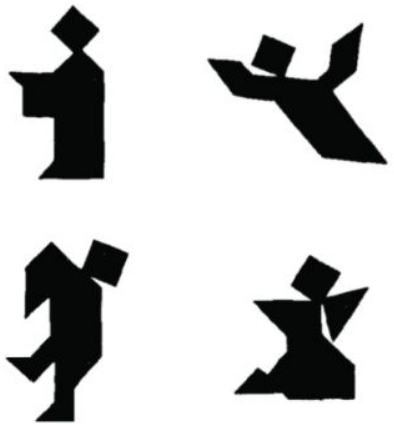
Descriptions diverge across tangrams



Game Overview

Please read these instructions carefully! You will have to pass a quiz on how the game works before you can play!

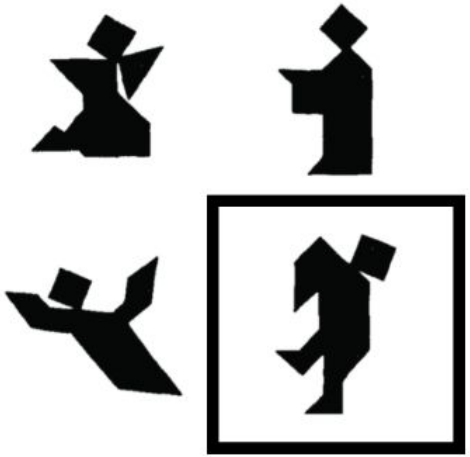
In this task, you will be assigned to a team with 3 other people (**4 people including yourself!**). You will play a series of communication games with people on your team. Everyone on your team will see the same set of pictures, which will look something like this:



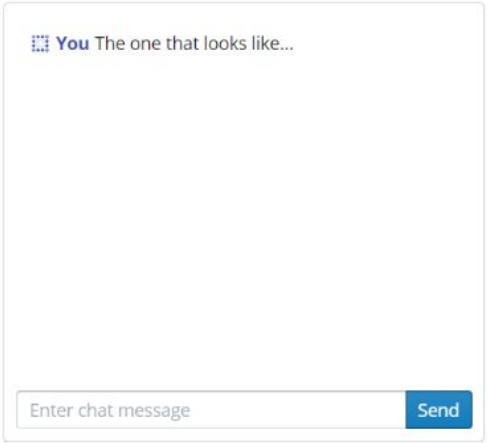
One of the people on the team will be randomly assigned the **Speaker** role and the others will be assigned the **Listener** role. The Speaker role will rotate to a new person each round.

One of the people on the team will be randomly assigned the **Speaker** role and the others will be assigned the **Listener** role. The Speaker role will rotate to a new person each round.

If you are the Speaker, you will see a black box secretly marking one of the pictures as the **target**.



The Speaker's job is to send a description of the target through the chatbox so that the Listeners are able to pick it out of the set. You can write whatever description you think will best allow your partners to identify the target. Please note that **the order of the pictures on your screen is scrambled**, so descriptions like "the one on the left" or "the third one" will not work. Also, please limit your description to the current target picture: **do not discuss previous trials or chat about any other topics!**



After the Speaker sends a message, the Listeners will read it and **each click the picture they believe is the target**. They are also allowed to respond by sending messages back through the chatbox until they are ready to make a selection.

« Previous Next »

Team Details

To help you identify yourself and differentiate each other in the team, we will assign an icon and a name to you when the game starts (as shown in the following example). This also shows who has what role.

You and your teammates have 3 minutes to select an image on each repetition. If you do not select an image in this time frame, you will automatically **progress to the next task when the time is up** and will not get a bonus, so please stay focused.

There will be **12 pictures shown at a time**. As a group, you will go through the all the pictures 6 times, so each picture can appear as the target multiple times.

You may communicate with your teammates through the in-game chat. Whatever you write will appear to your teammates. You can use the chat function however you like, but please note that **the Speaker must send a message before Listeners can make their selections**.

Note that the game allows for simultaneous and real-time actions. Each trial will only end after all the listeners have made a selection (or the timer runs out).

At that time, you will see feedback on what everyone selected and whether they were correct or not.

« Previous Next »

Scores and Bonuses

In each task, we use "score" to evaluate the quality of the selections that you and your partners make. Your total score will be calculated as the sum of all your scores on each round.

Each time a **Listener** makes a correct selection, they get **4 points**.

Incorrect selection and no selection (timing out) earn no points.

The **Speaker** gets a bonus equal to **the average** of the Listeners's scores.

For instance, if there are two Listeners and both make a correct selection, they each get 4 points, and the Speaker gets 4 points as well. If one Listener chooses correctly and one doesn't, the correct Listener gets 4 points, the other Listener gets 0 points, and the Speaker gets 2 points.

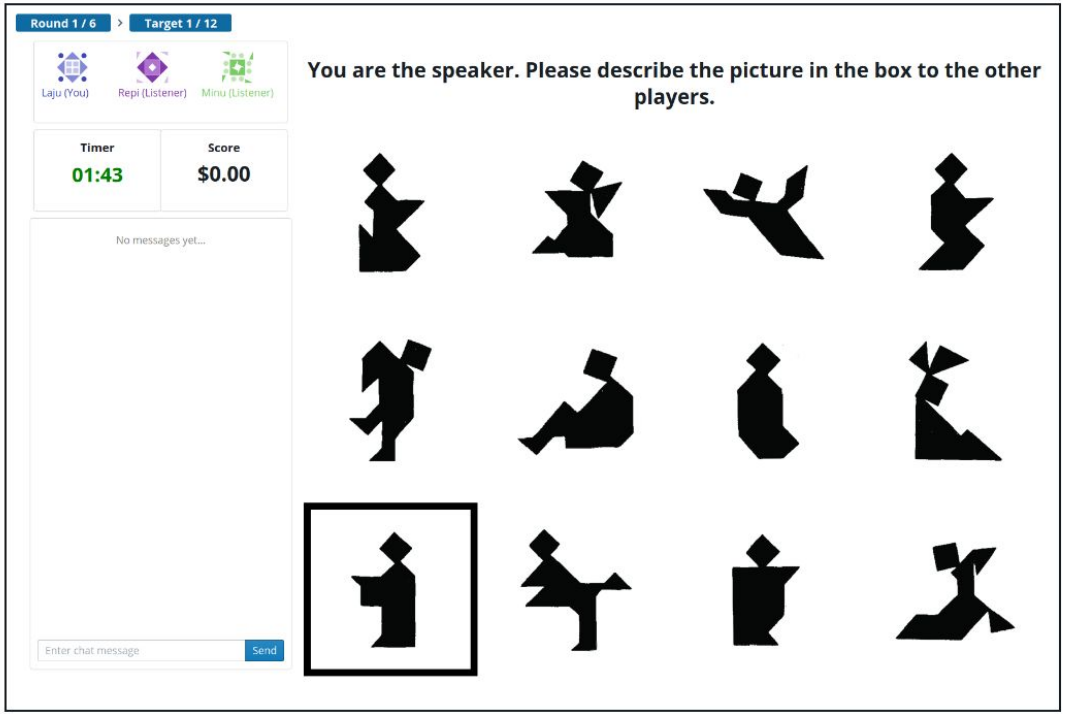
Your performance bonus will be based on your score at the end of the experiment.

Remember, free riding is not permitted. If we detect that you are inactive during a task, you will not receive a bonus for that task.

« Previous Next »

Game Interface

We are almost there! Please take a second to familiarize yourself with the game Interface shown here:



Now you know where everything goes and are ready to take the quiz! Good luck.

« Previous Next »

Quiz

How many participants will play at the same time on your team, including yourself?

How many pictures will be shown at a time?

Select the true statement about the score:

- The Speaker gets more points if more Listeners make the right choice.
- The Speaker gets the same number of points if 1 or 2 Listeners makes the right choice.

Select the true statement about the chat:

- Anyone can send messages through the chat.
- Only the Speaker can send messages through the chat.

Select the true statement about the pictures:

- Everyone will see the same pictures in the same places in the grid.
- Pictures will be mixed up and in different places for different people.

If a Listener does NOT choose a tangram before the time is up then their score in that task will be:

« Back to instructions Submit ↵