

How to crowdsource behavioural data in the social sciences

Jon W. Carr & Jasmeen Kanwal

Centre for Language Evolution

School of Philosophy, Psychology and Language Sciences

University of Edinburgh



CSCS@ED

Today's workshop

About Us

Why run experiments online?

CrowdFlower & Mechanical Turk

Ethics & Payment

Web Programming Basics

Live Demo & Hands-on

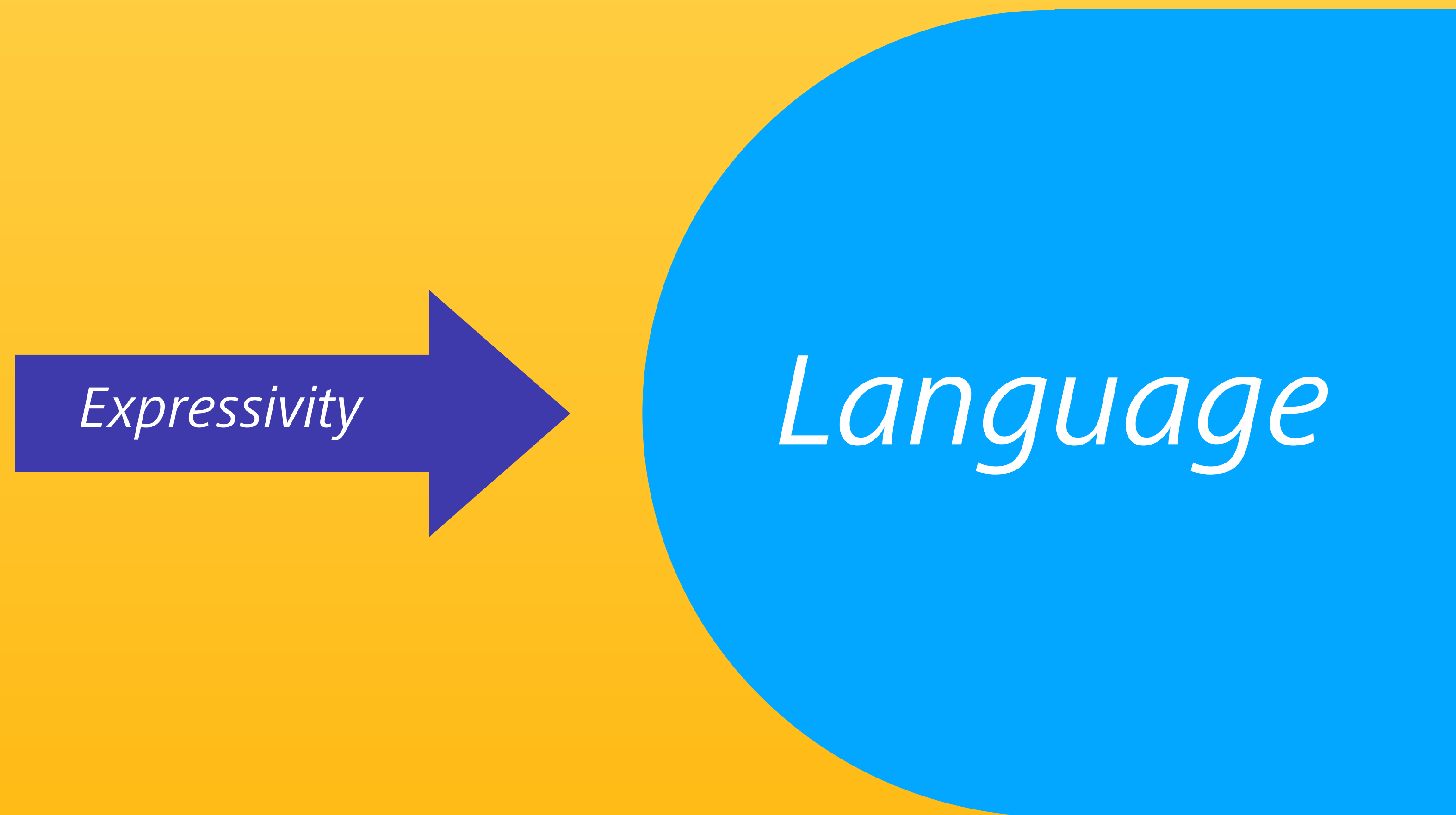
Question & Answer

What shapes language?



Language

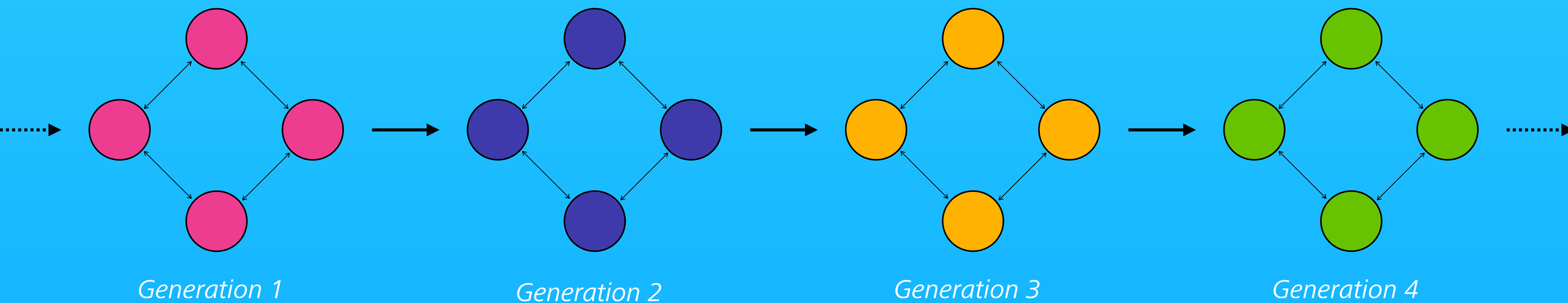
What shapes language?



What shapes language?



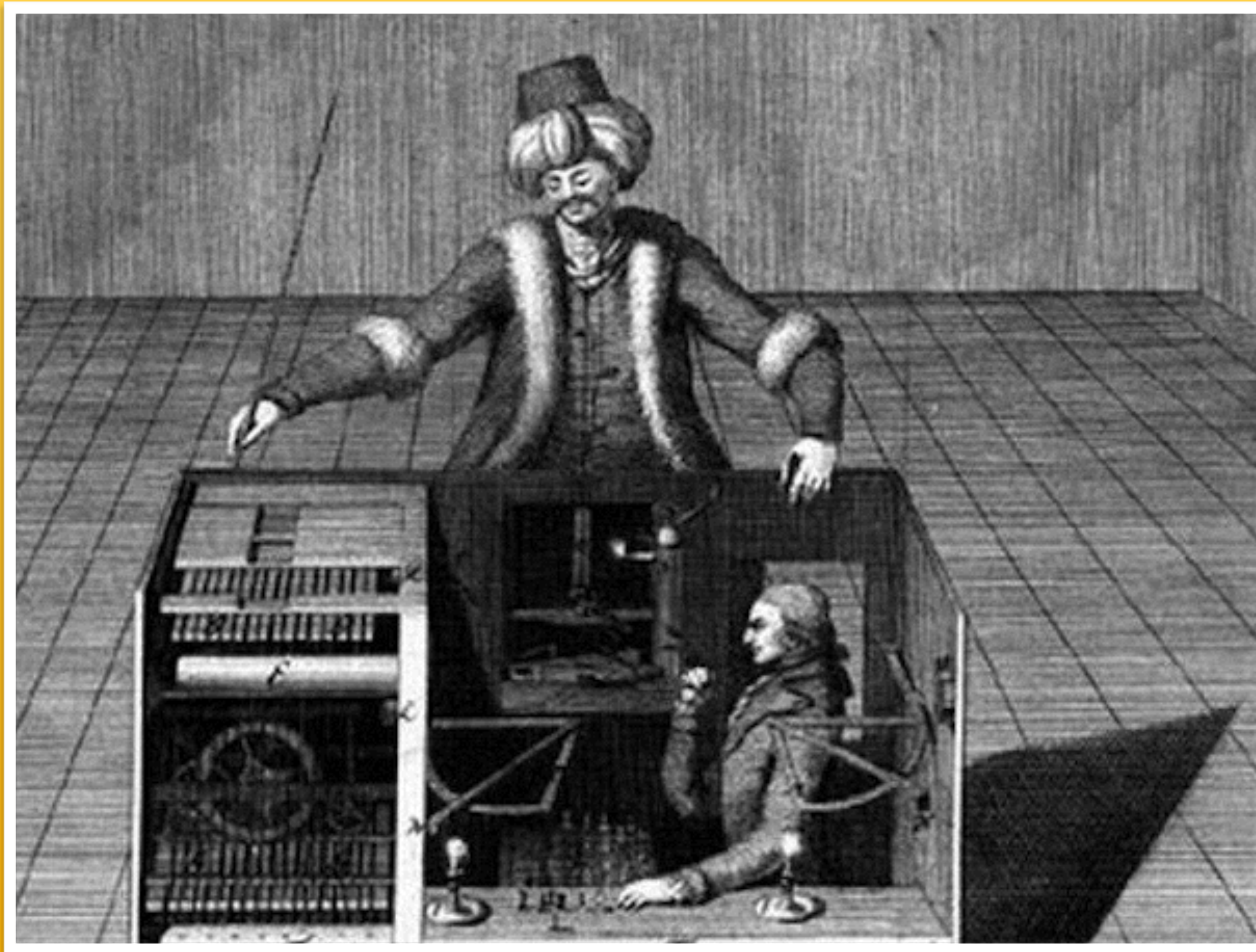
Language Evolution in the Lab



Why Run Experiments Online?

- Get results fast. You can have a few hundred people do your experiment within half an hour.
- Save time. You don't need to spend weeks in the lab explaining the task over and over to participants.
- Cheaper. Each participant spends less time because they don't need to come into the lab.
- More diverse population. Not as WEIRD?
- As good as lab results. Reproduce classic results.

Mechanical Turk



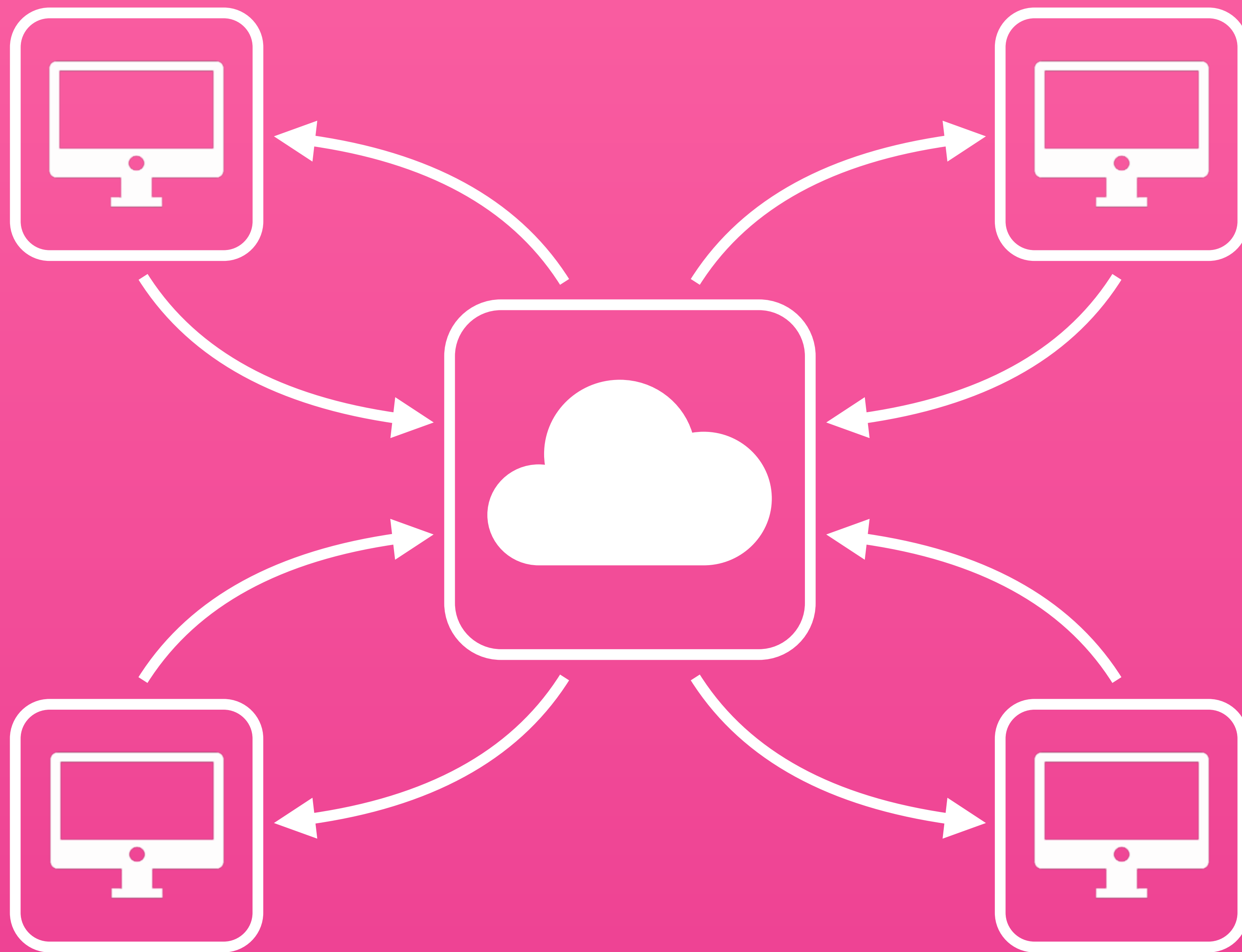
CrowdFlower

- Open to people outside US
- Participants from all over the world, so potentially a more diverse population than MTurk
- CF mostly outsources to other (possibly questionable) companies, and the fee taken by these companies is not transparent
- The platform is not ideal for running your own custom experiments, so you have to resort to slightly hacky methods to get it to work

Ethics & Payment

- Standard ethics apply. You still need to get ethical approval.
- Store online participants data securely and anonymously just like you would with lab participants.
- Set wages equivalent to minimum wage or minimum wage set by your ethics body.
- Generate a completion code and give it to your participant, so that you can verify that they completed the task.
- Be aware of potential payment issues.

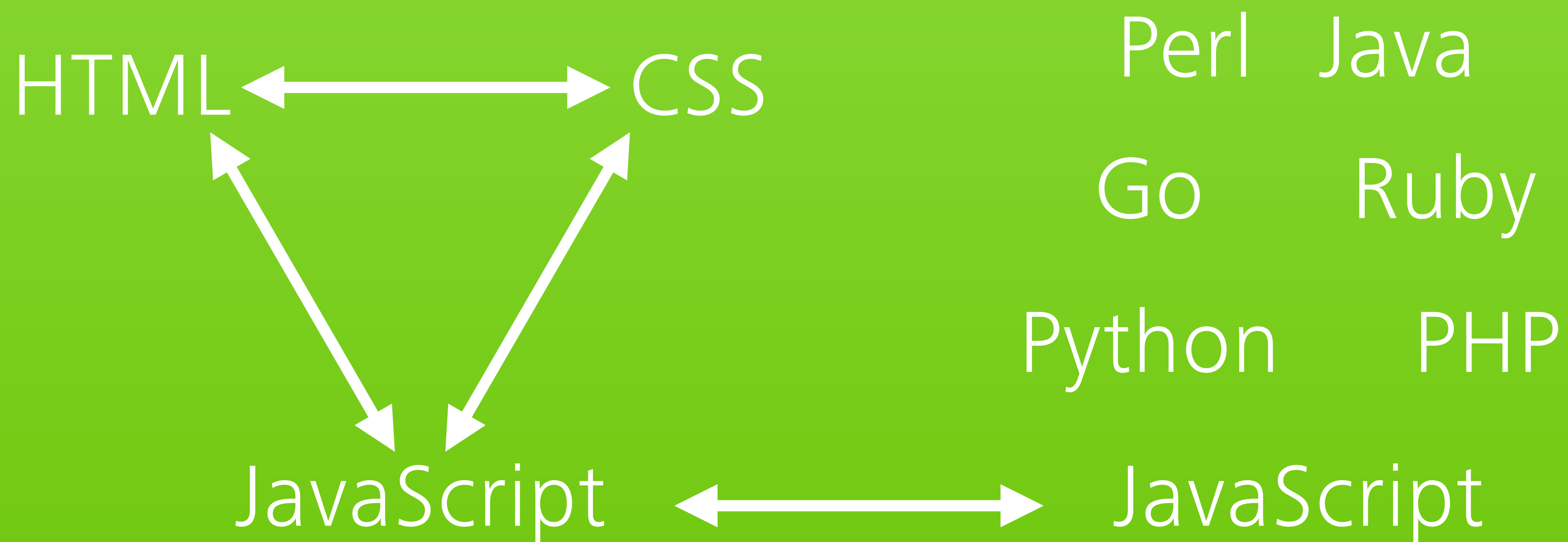
Server-Client Basics



Web Programming

Client-side

Server-side



Client-Side

HTML: Markup language describing the elements of the page (DOM)

CSS: Markup language describing how things should be styled (colours, sizes, positions, etc)

JavaScript: Programming language where you do the client-side logic

Subitizing

Rapid, accurate, and confident estimation of numerosity, especially for small numbers

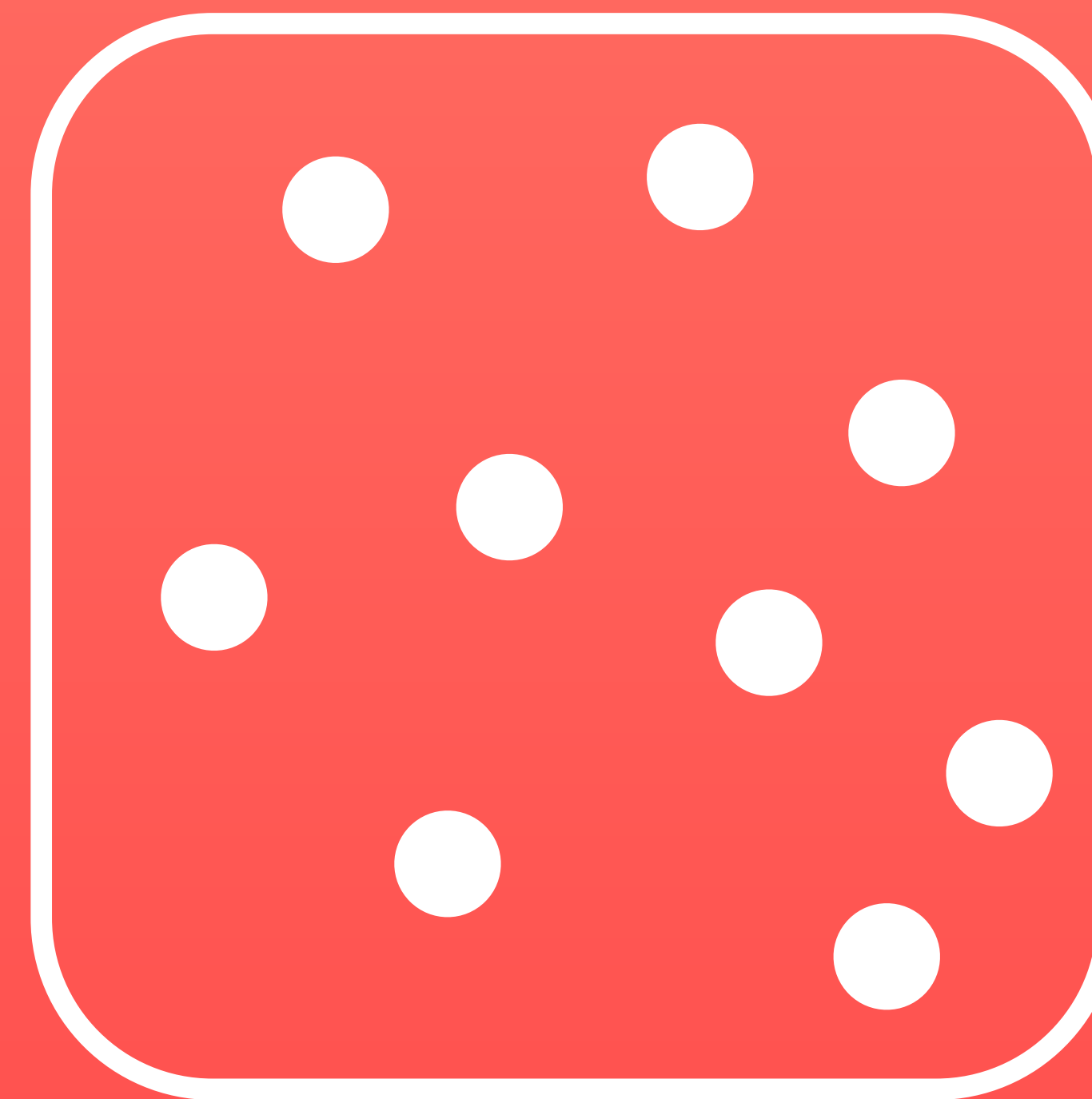
Subitizing

Rapid, accurate, and confident estimation of numerosity, especially for small numbers

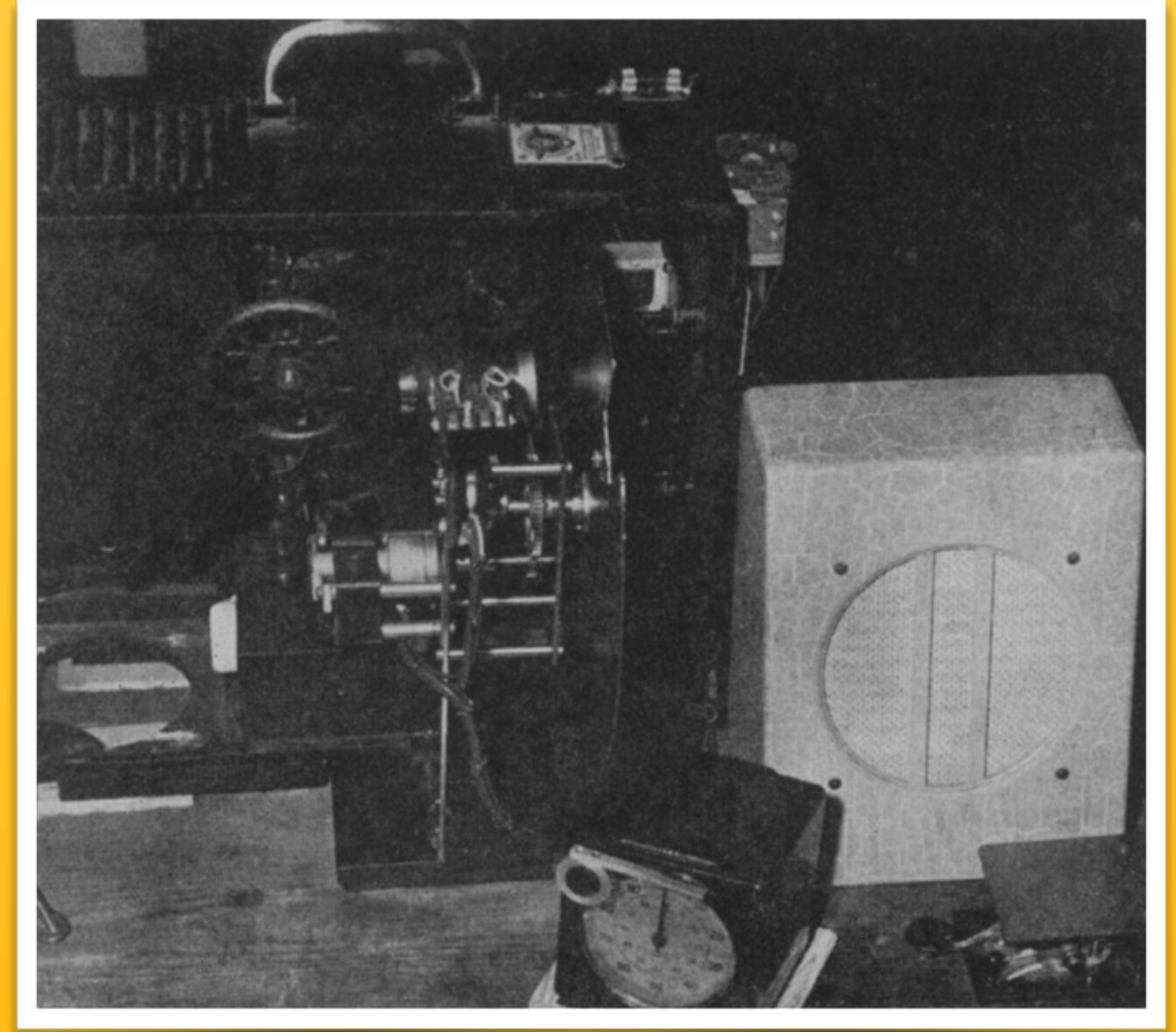


Subitizing

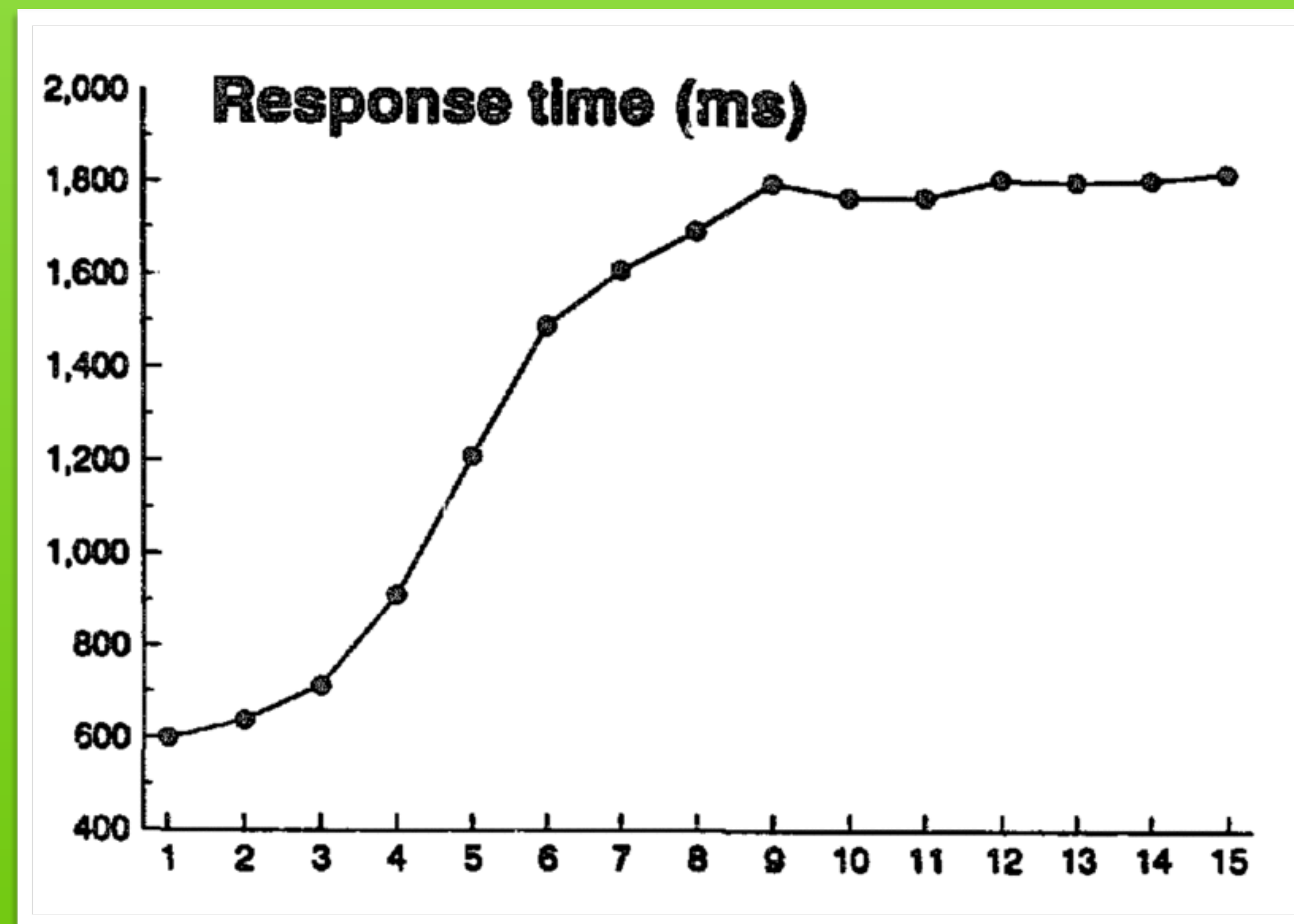
Rapid, accurate, and confident estimation of numerosity, especially for small numbers



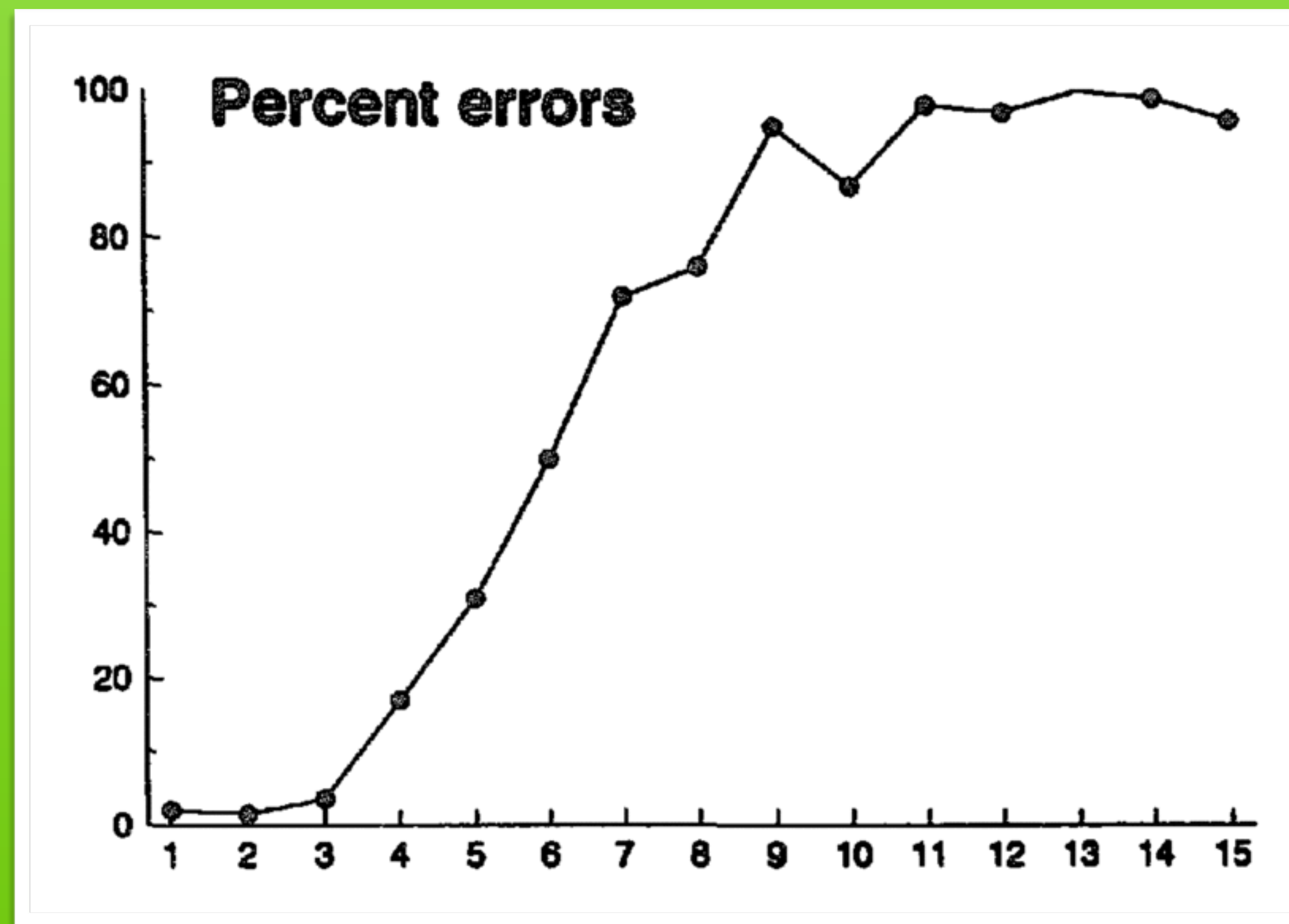
Subitizing



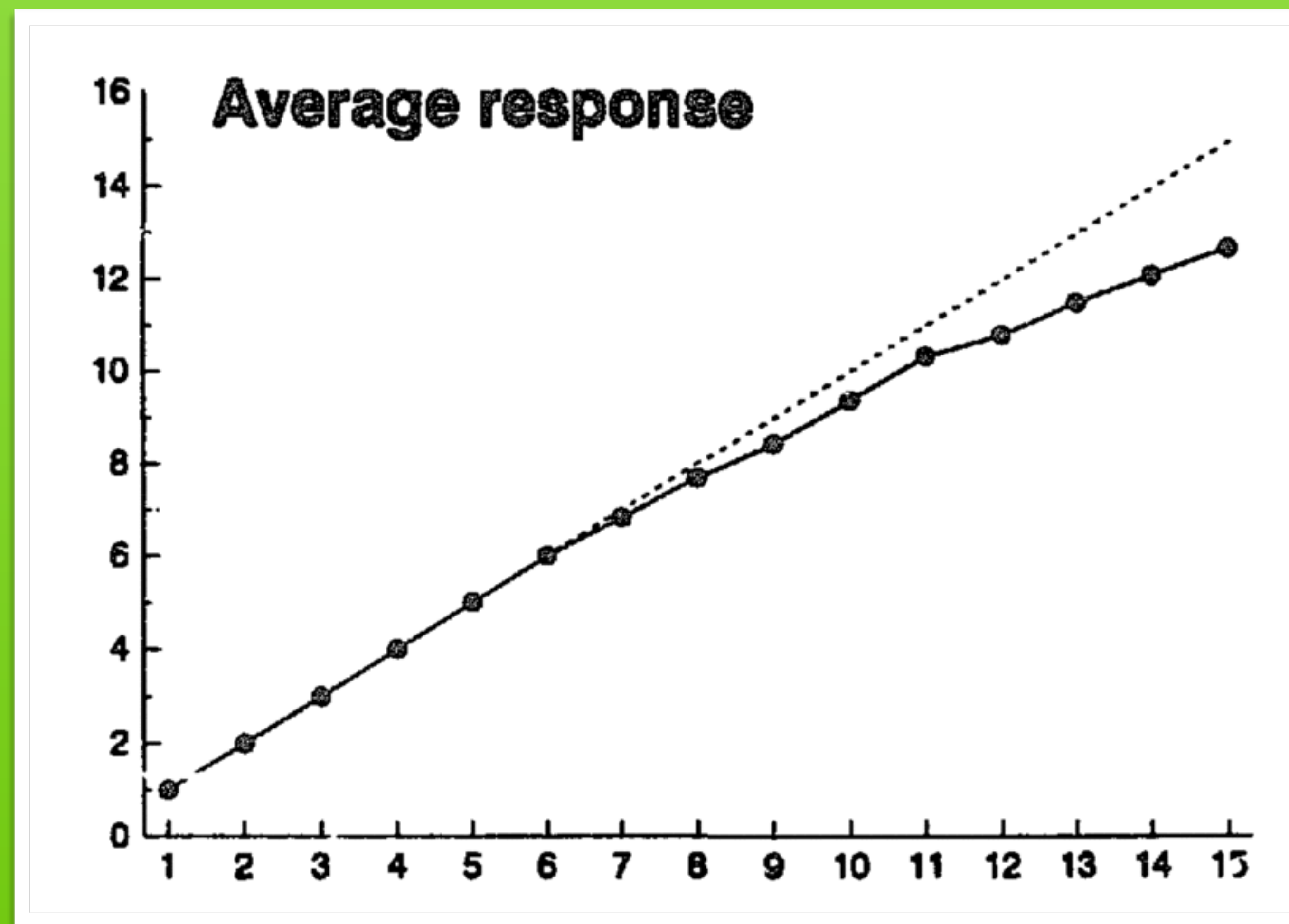
Subitizing



Subitizing



Subitizing



Hands-on

Try the experiment

blake.ppls.ed.ac.uk/~s1153197/sub/cf.php

Then set up the task

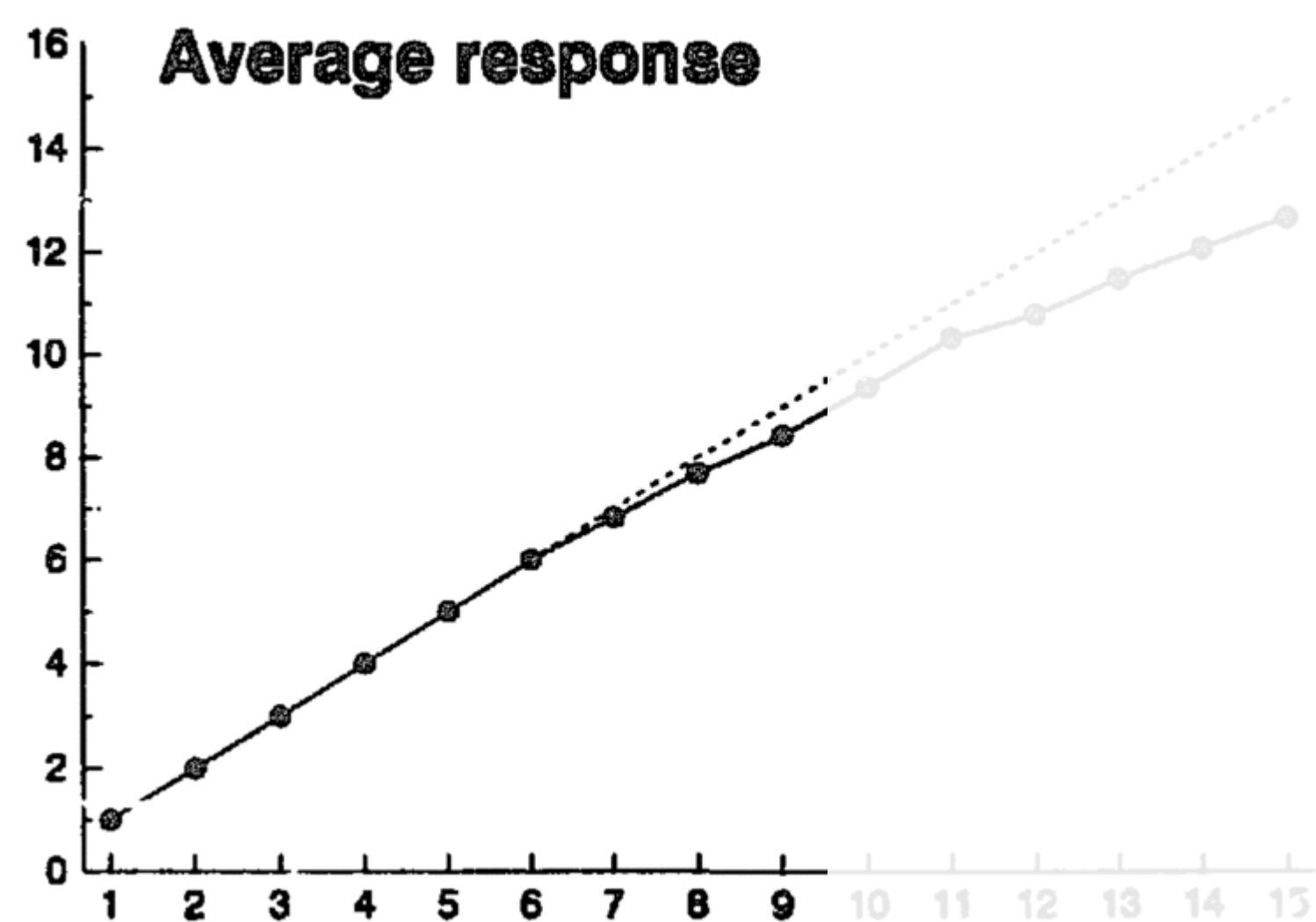
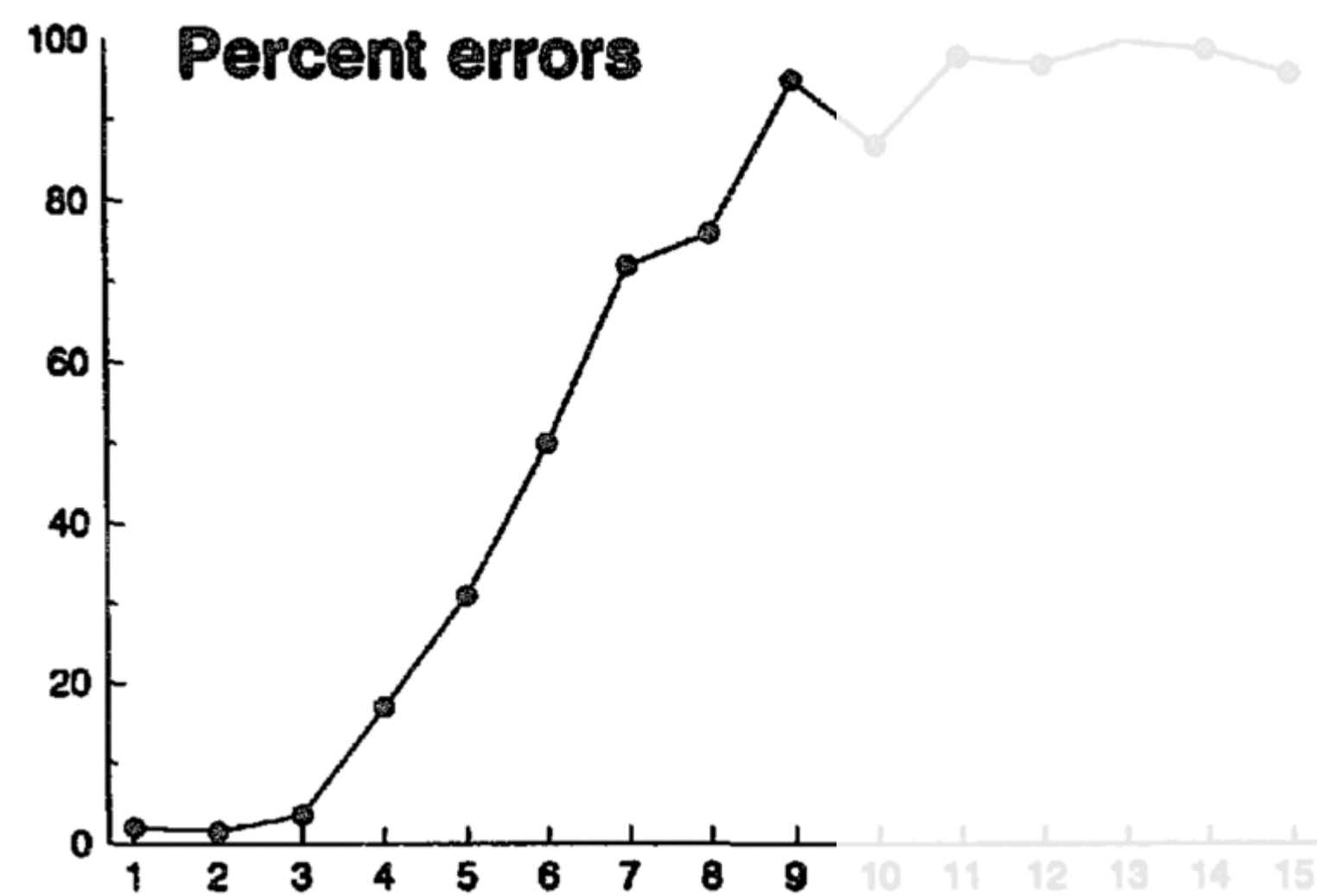
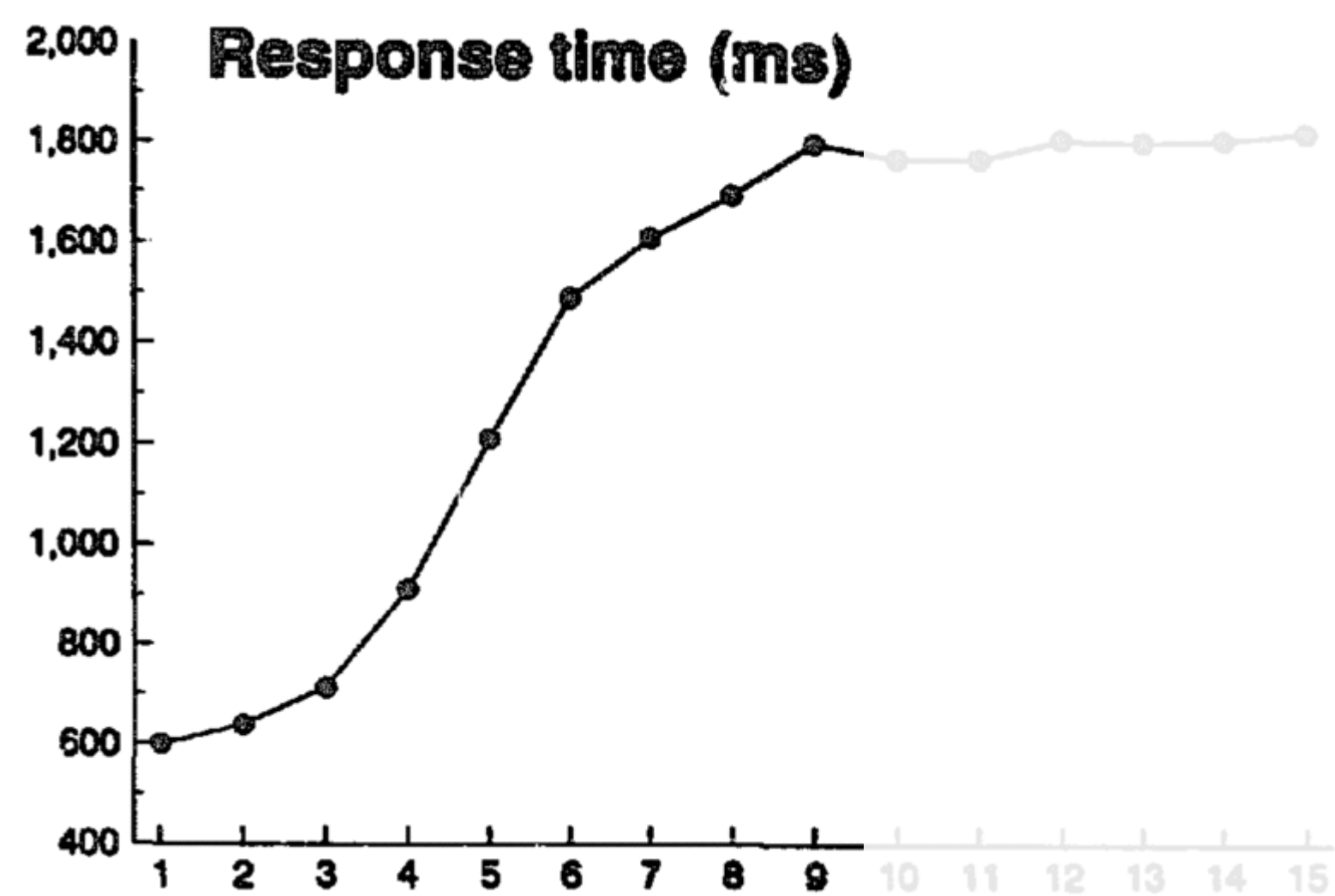
crowdflower.com OR mturk.com

If you're feeling adventurous... play with the code

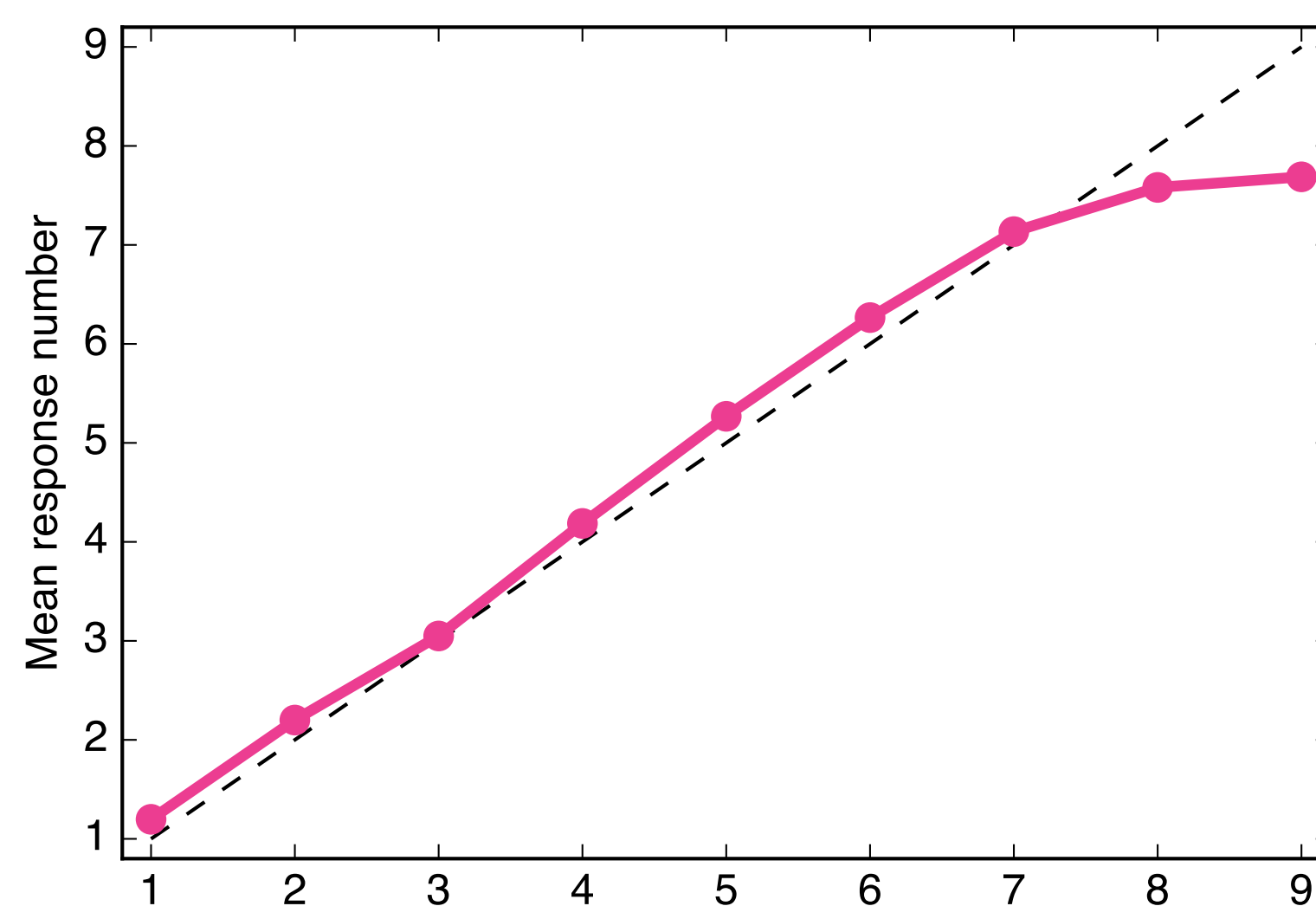
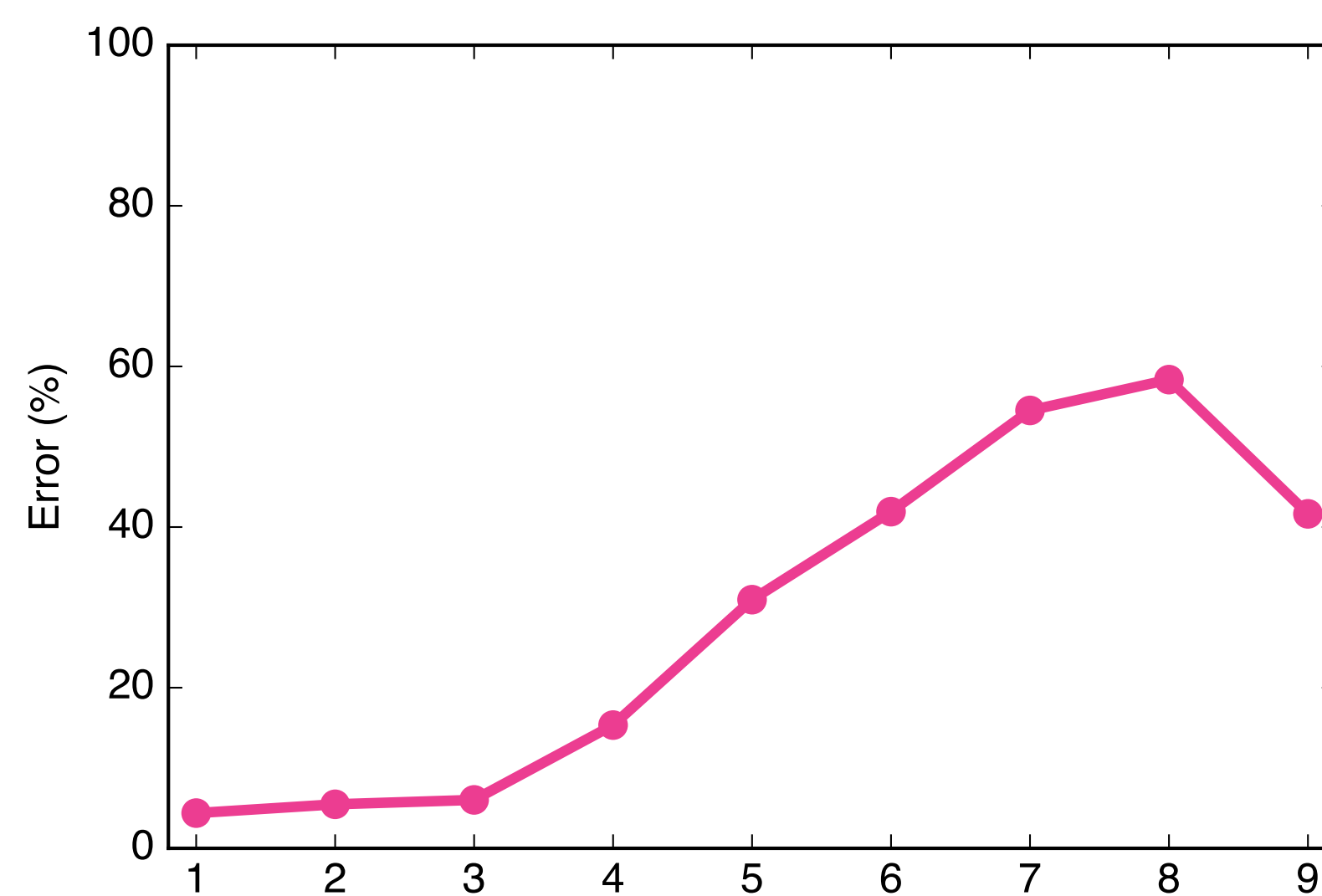
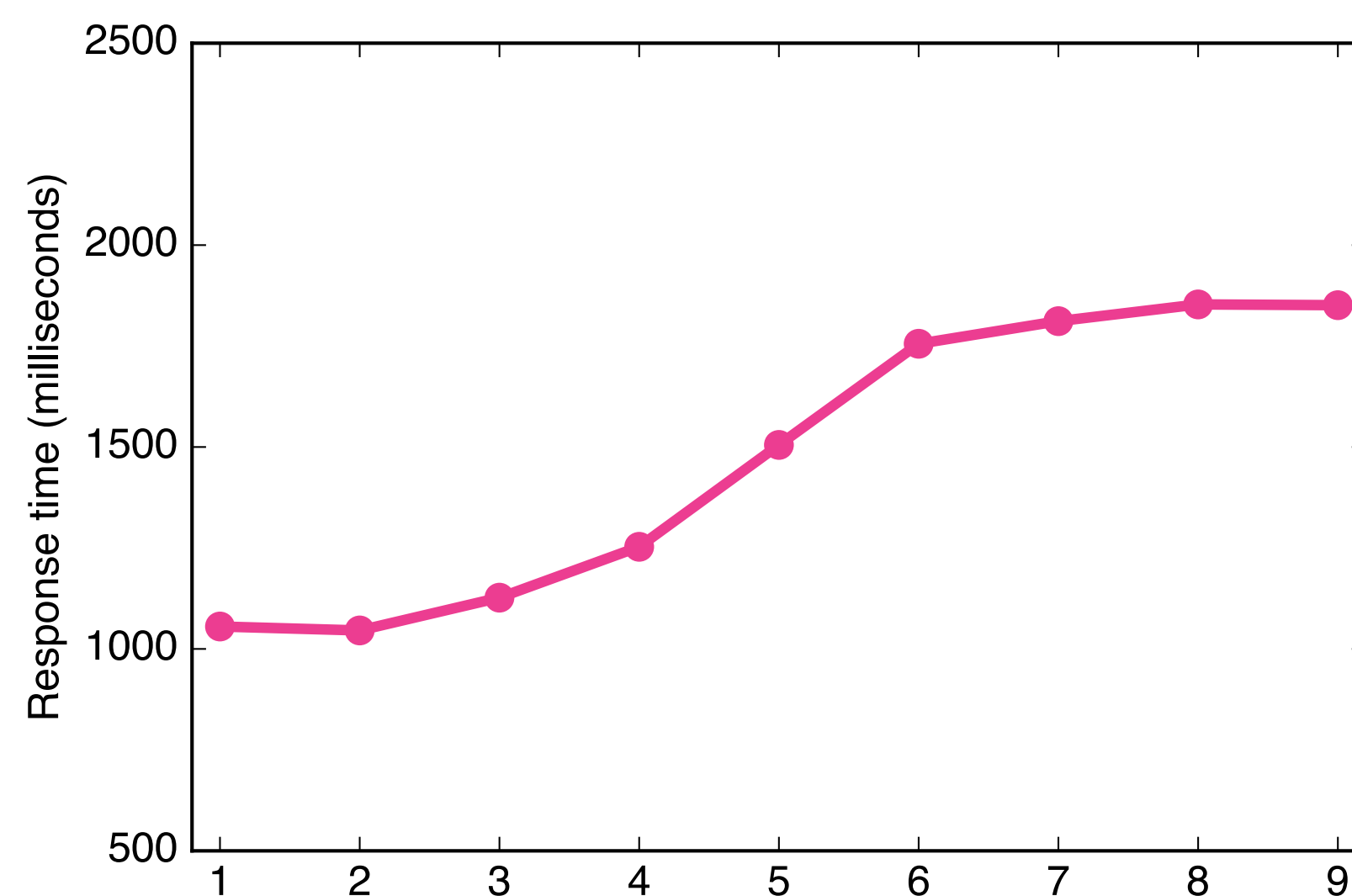
github.com/jwcarr/subitizingOnline

Results...

Results reported in Dehaene (1992)



Our CrowdFlower results



Our Mechanical Turk results

