

Online Experiments for Language Scientists

Lecture 6: Word learning

Kenny Smith

kenny.smith@ed.ac.uk

Matters arising: office hour

- I am available for extra help (for if you are behind on labs only!) –
Thursday at 10am

Matters arising from the self-paced reading lab

“Is there a way to write a long experiment that doesn’t involve copying, pasting and editing the same code over and over?”

- Yes, several! My preferred method: **functions that build trials**, coming in week 7, this week is the last of copy-and-paste code

“Is there a way to automatically randomise the order of things (e.g. buttons on a button-response trial)”

- Yes! Also coming in week 7

Standing business: Any questions about annotated bibliography assessment?

Chua & Pan (2026)

Chua, T. J. E., & Pan, S. C. (2026). Duolingo-inspired pretesting with words and pictures improves vocabulary learning. *Cognitive Research*, 11, 20.

Four duolingo-like word learning experiments on Prolific

- Does **pretesting** (guess with feedback) facilitate word learning?



Tabitha Chua
(National University of Singapore)



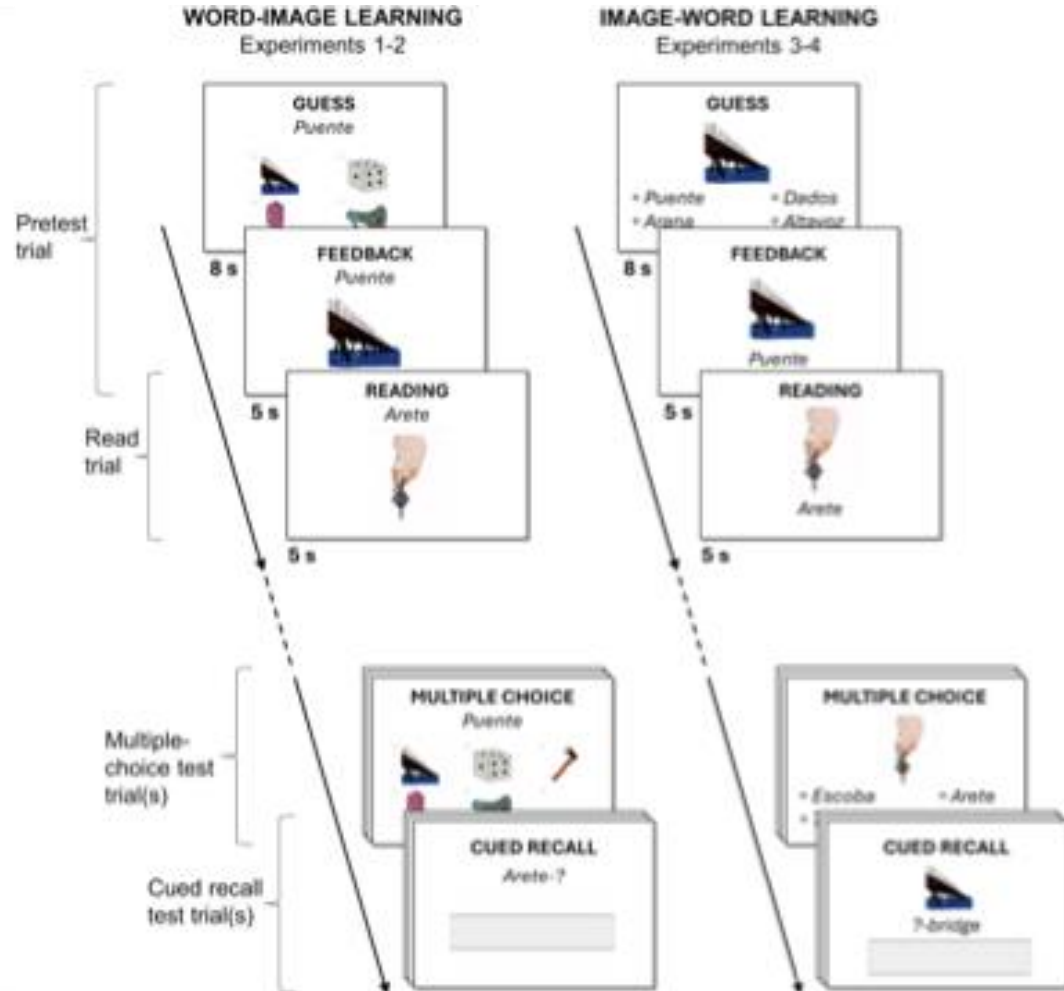
Steven Pan
(National University of Singapore)

Procedure and design

Exp 1: word-to-image, training trial types randomly intermixed, test trial types randomly intermixed

Exp 2: word-to-image, **blocked vs randomized testing**

Exps 3-4: Same as 1-2, but image-to-word



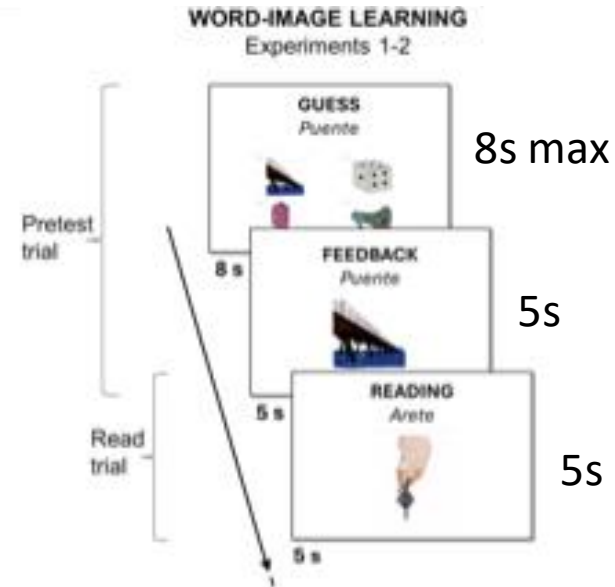
Word learning

Demo with this week's lab code

Image-to-word, blocked testing

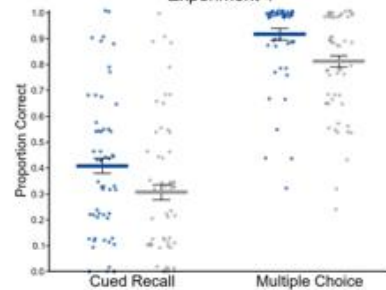
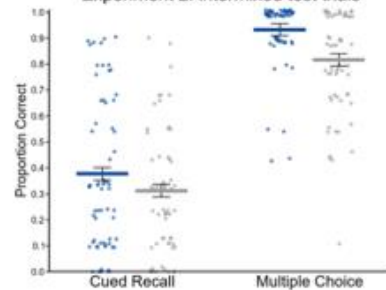
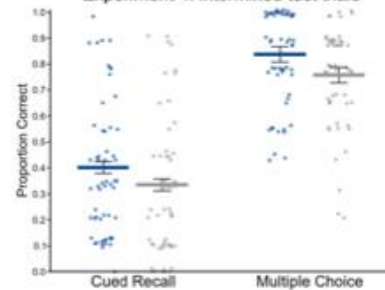
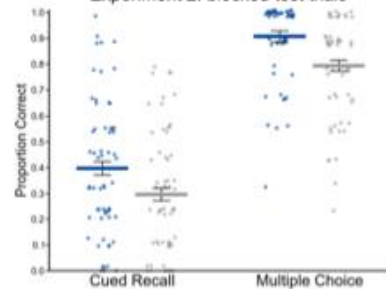
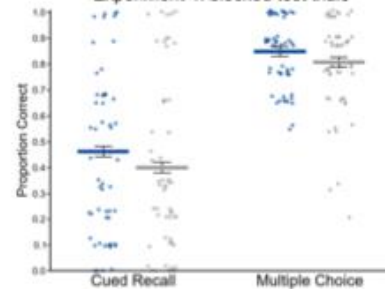
A question from the pre-lecture form

“In Experiment 1, the authors acknowledged that “participants in the pretesting condition had longer exposure to the Spanish word alone, presumably without knowing its correct meaning.” If we were to replicate the experiment, how could we modify the experimental design to control for this potential confounding factor of the longer exposure time?”



Sample size, study duration etc

- Self-reported fluent speakers of English, resident in an English-speaking country, >95% Prolific approval, age 21-45
- N=78>58 (Exp 1), 153>123 (Exp 2), 65>47 (Exp 3), 162>113 (Exp 3)
- ?? minutes, \$2.67

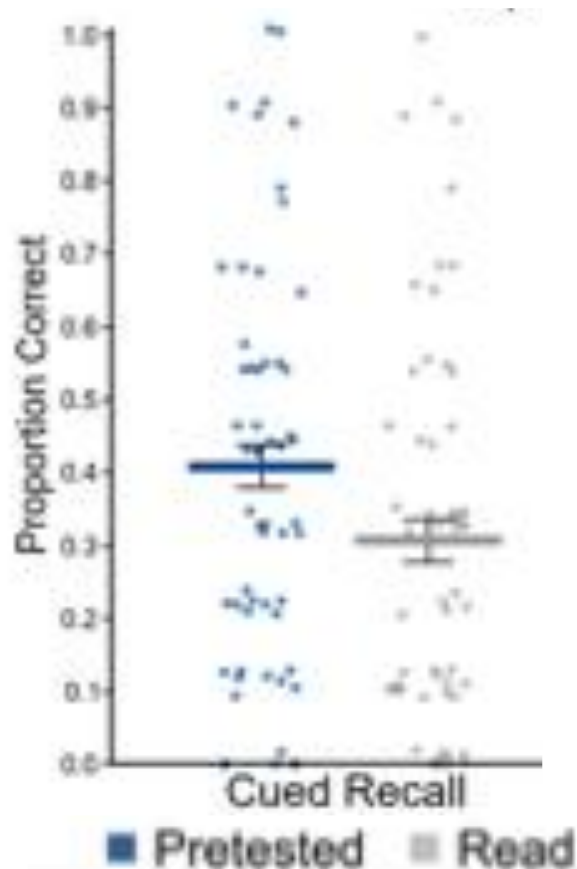
WORD-IMAGE LEARNING**Experiment 1****IMAGE-WORD LEARNING****Experiment 3****Experiment 2: intermixed test trials****Experiment 4: intermixed test trials****Experiment 2: blocked test trials****Experiment 4: blocked test trials**

A question from the pre-lecture form

“I understand the general findings, but I’d like to understand how the statistical values are calculated and interpreted. For example, in Experiment 1, how do the researchers calculate the t-value from the data, and how is the p-value obtained from it? I’m also not quite sure how Cohen’s d, the 95% confidence interval, and the Bayes factor (BF_{10}) are calculated, or what additional information each provides.”

WORD-IMAGE LEARNING

Experiment 1



$$t = \frac{\overline{dif}}{s_{dif}/\sqrt{n}}$$

$$d = \frac{m_1 - m_2}{s}$$

A question from the pre-lecture form

“I'm curious why the researchers chose $d=0.49$ as the expected effect size for the power analysis in Experiment 1 (p.3). Although they cited Potts and Shanks (2014), I'm not sure how this specific value was determined or why it was considered appropriate for this study.”

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General*, 143, 644–667.

Chua & Pan's conclusions

Pretesting does facilitate word learning

- For both word-to-image and image-to-word learning
- The effects seem not to depend on the details of test trial ordering

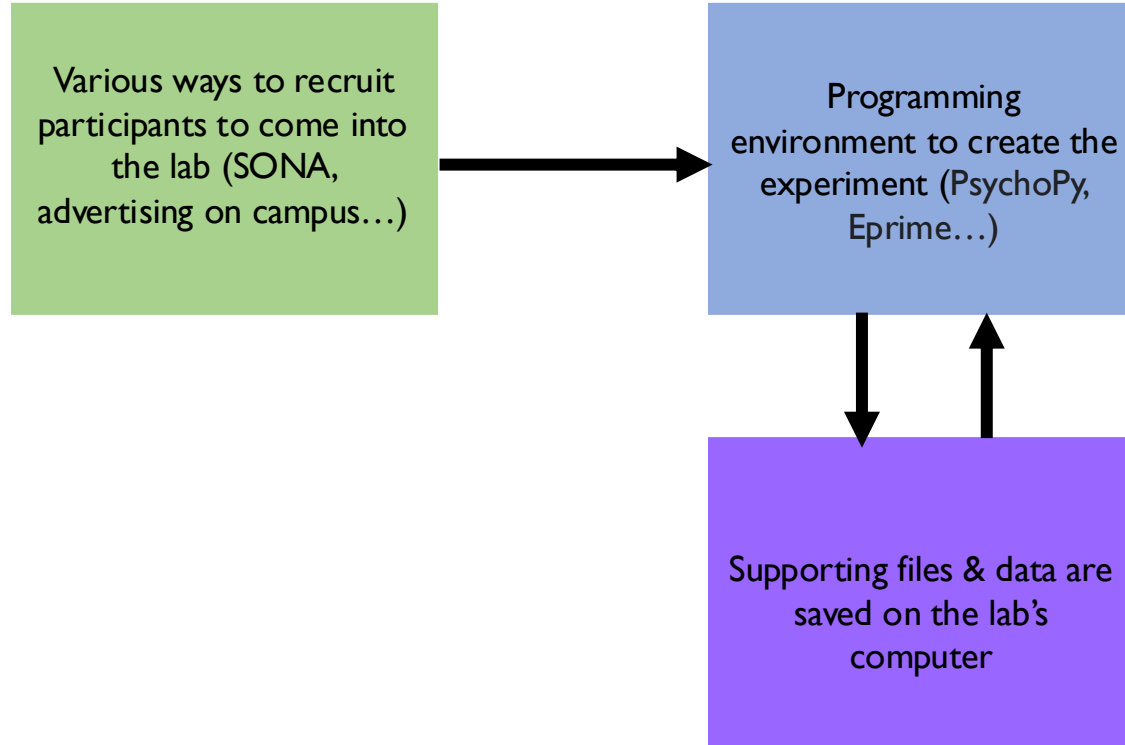
Additional note

- Successful use of online data collection for a task involving sustained attention and effort
- But there's always something to keep participants engaged

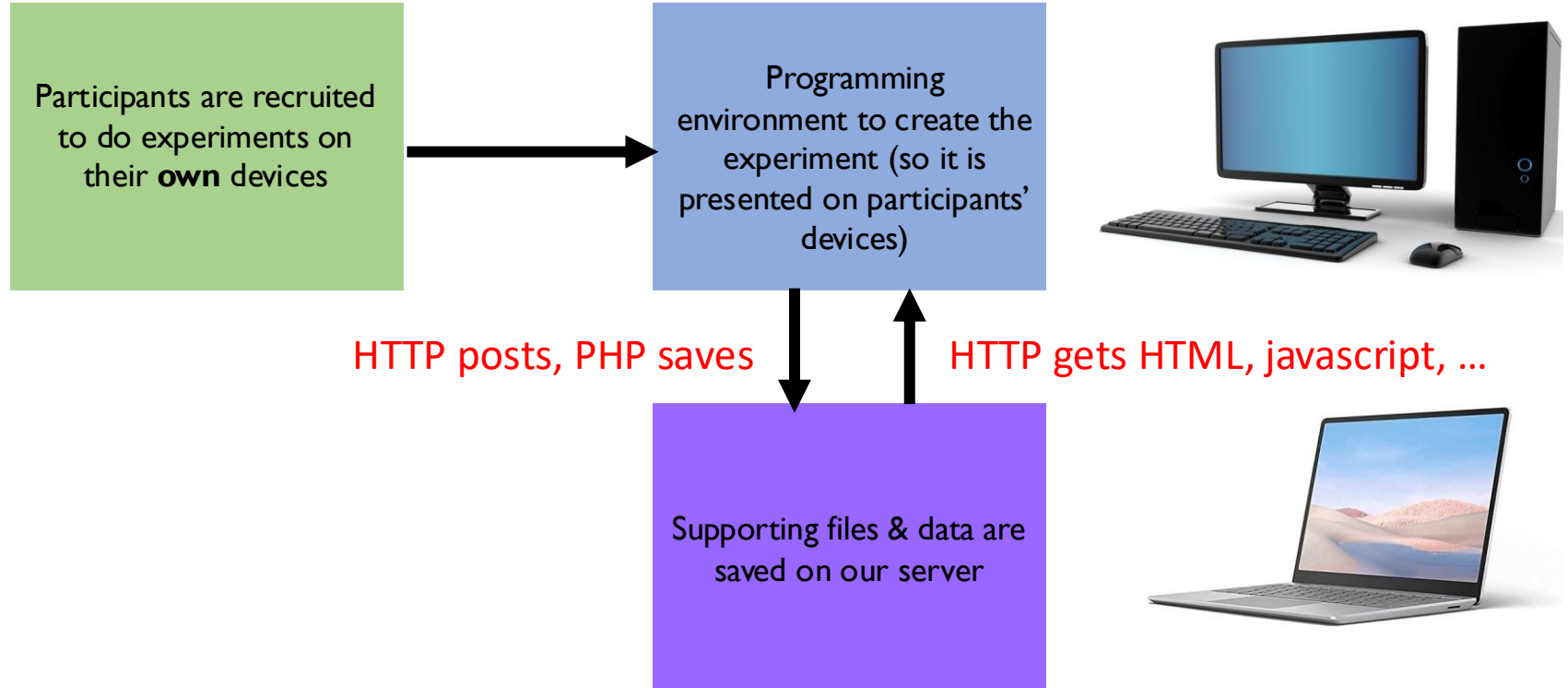
Additional time for Q&A/discussion on this week's reading

Coming up in labs: saving data to the server

Running experiments in the lab



Running experiments online



News

Opinion

Sport

Culture

Lifestyle



The Guardian

UK ~

UK US politics World Climate crisis Middle East Ukraine Football Newsletters Business Environment UK politics Science Tech Global development Obituaries

Data and computer security

Josh Taylor and Robert Booth

Thu 22 Oct 2026 17:42 BST

Share

Prefer the Guardian on Google

Explainer

What is PHP and how serious is the computer language flaw causing website shutdowns?

An estimated 70% of websites use the open-source language. What is the identified vulnerability in it?



📷 Vulnerabilities discovered in the programming language PHP could allow hackers to gain control over servers. Photograph: ANP/Shutterstock

Advertisement



aws

Powering the world's most innovative companies.

AWS AI is how →

How it works

1. Our experiment code collates the data on the participant's computer
 - jsPsych stores data for every trial
2. We send it to the server using fetch and POST

```
function simple_save_data(name, data_in){  
  var url = 'save_data.php';  
  var data_to_send = {filename: name, filedata: data_in};  
  fetch(url, {  
    method: 'POST',  
    body: JSON.stringify(data_to_send),  
    headers: new Headers({  
      'Content-Type': 'application/json'  
    })  
  });  
}
```

How it works

3. Our PHP script on the server **processes and saves** the incoming data

- This is the dangerous bit!
- save_data.php needs to include protections to prevent people running arbitrary commands on our server!
- The save_data.php we give you includes these (and there are substantially more than in previous versions!)

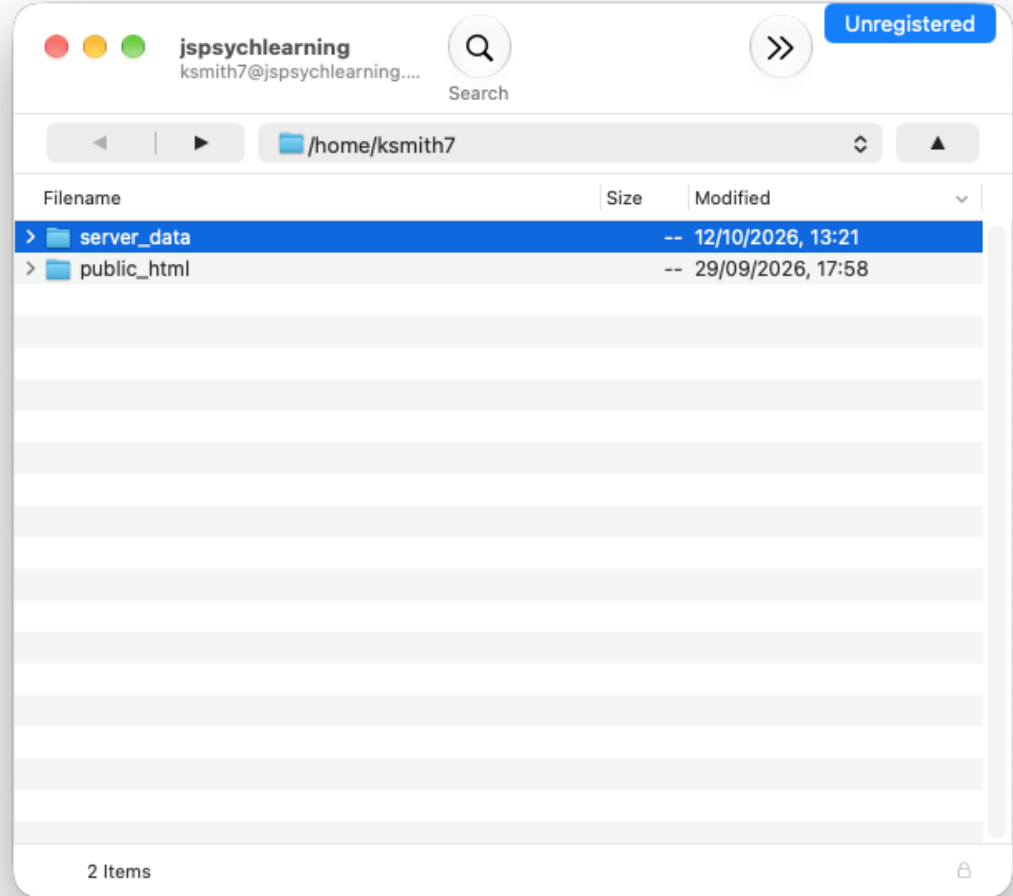
Where to save data

Not in public_html!

- Data usually shouldn't be public

In server_data

- Our PHP script can write there
- But only you can access it



When to save data?

For now: at the end of the experiment, in a single file

```
var jsPsych = initJsPsych({  
  on_finish: function () {  
    var all_data = jsPsych.data.get(); //get all data  
    var all_data_as_csv = all_data.csv(); //convert to csv format  
    save_data("wordlearning_data.csv", all_data_as_csv); //save it using save_data function defined in utilities.js  
  },  
});
```

Better (week 8 on):

- At the end of each trial
- Seperate file per participant

Next up

Wednesday lab

- **Same (different) venue this week ! G.06, 50 George Square**

Next week:

- Frequency learning (reading is **technical**, see my notes re. not worrying too much about the later sections!)