The case for psychometrics

Alvin Tan, Stanford University

CogSci workshop, 2024-07-24

Outline

- Welcome + introduction
- A pitch for psychometrics
- A case study: Item and individual variability

Welcome!

differential item functioning

adaptive testing

cognitive variability

accumulator models

drift diffusion models

resource competition

development

Introduction

This workshop

- Why is psychometrics important to cognitive science?
- How can psychological theories be built and tested as psychometric models?
- What are some good examples of psychometrics being used in cognitive science?

Resources

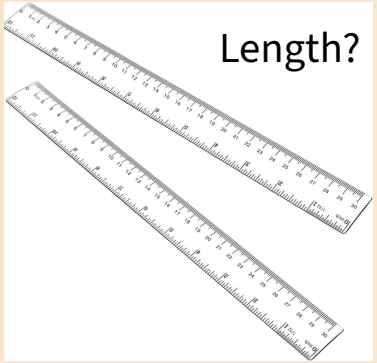
- Website: https://psychometrics-workshop.github.io/
- Pre-workshop tutorial on IRT

About you!

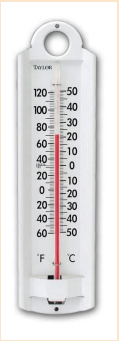
- Psychometricians
- Cognitive scientists who dabble in psychometrics
- Don’t currently work with psychometrics but would like to learn more
- Anyone else…

Psychometrics: Who cares?

How do we measure stuff?

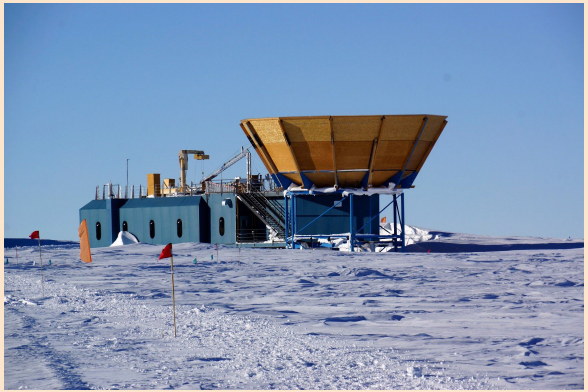


Length?



Temperature?

Cosmic microwave background radiation?



Personality?

Affect?

Lack of direct observational access

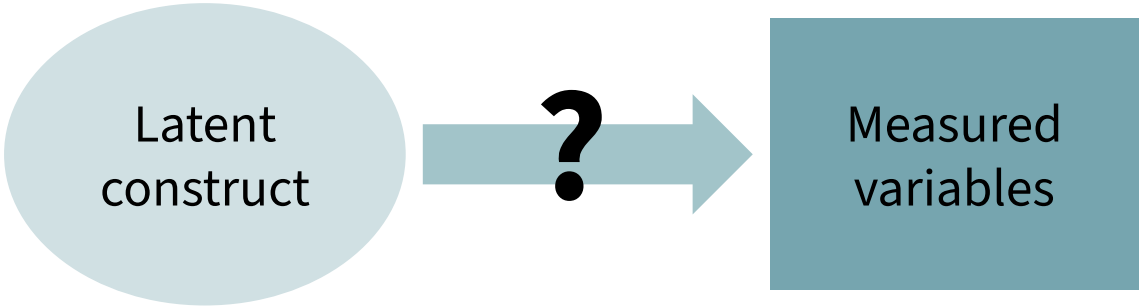
Memory?

Cognitive ability?

The place for psychometrics

What are the sources of variance?

Is my measurement consistent?



How are latent constructs organised?

How much error is there?

What is psychometrics?

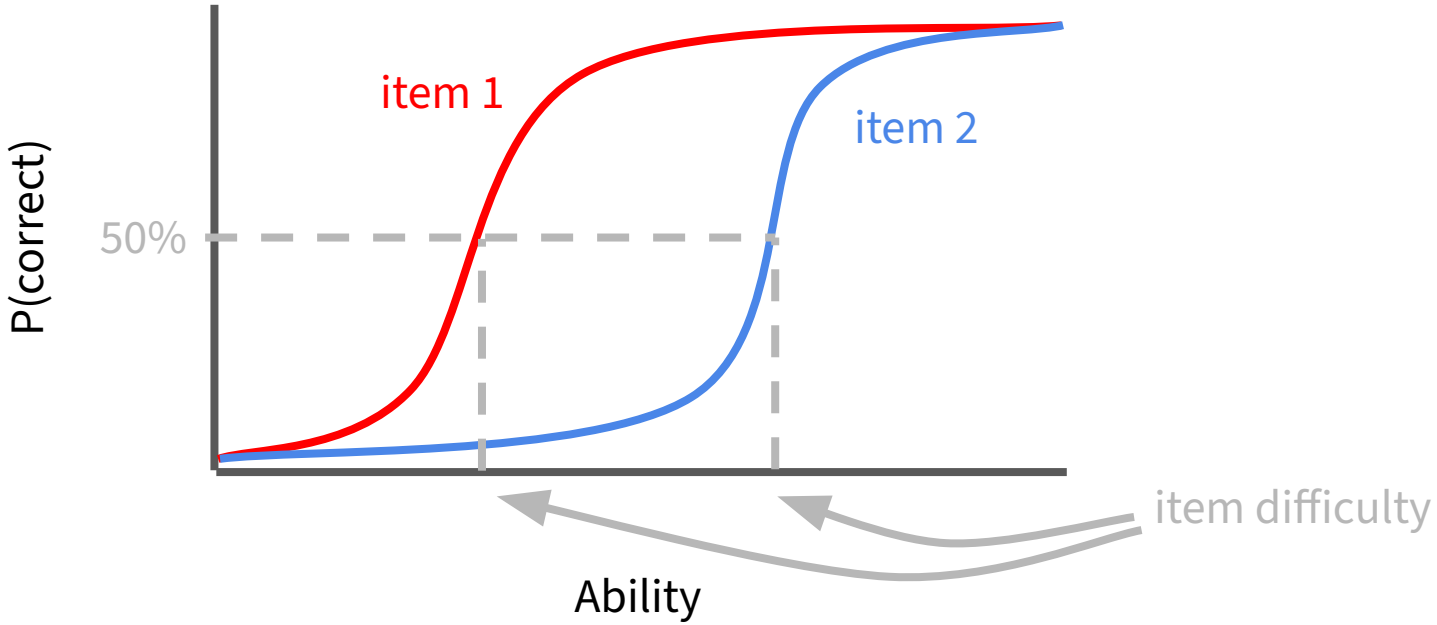
Validity

How do measurements relate to the underlying latent constructs?

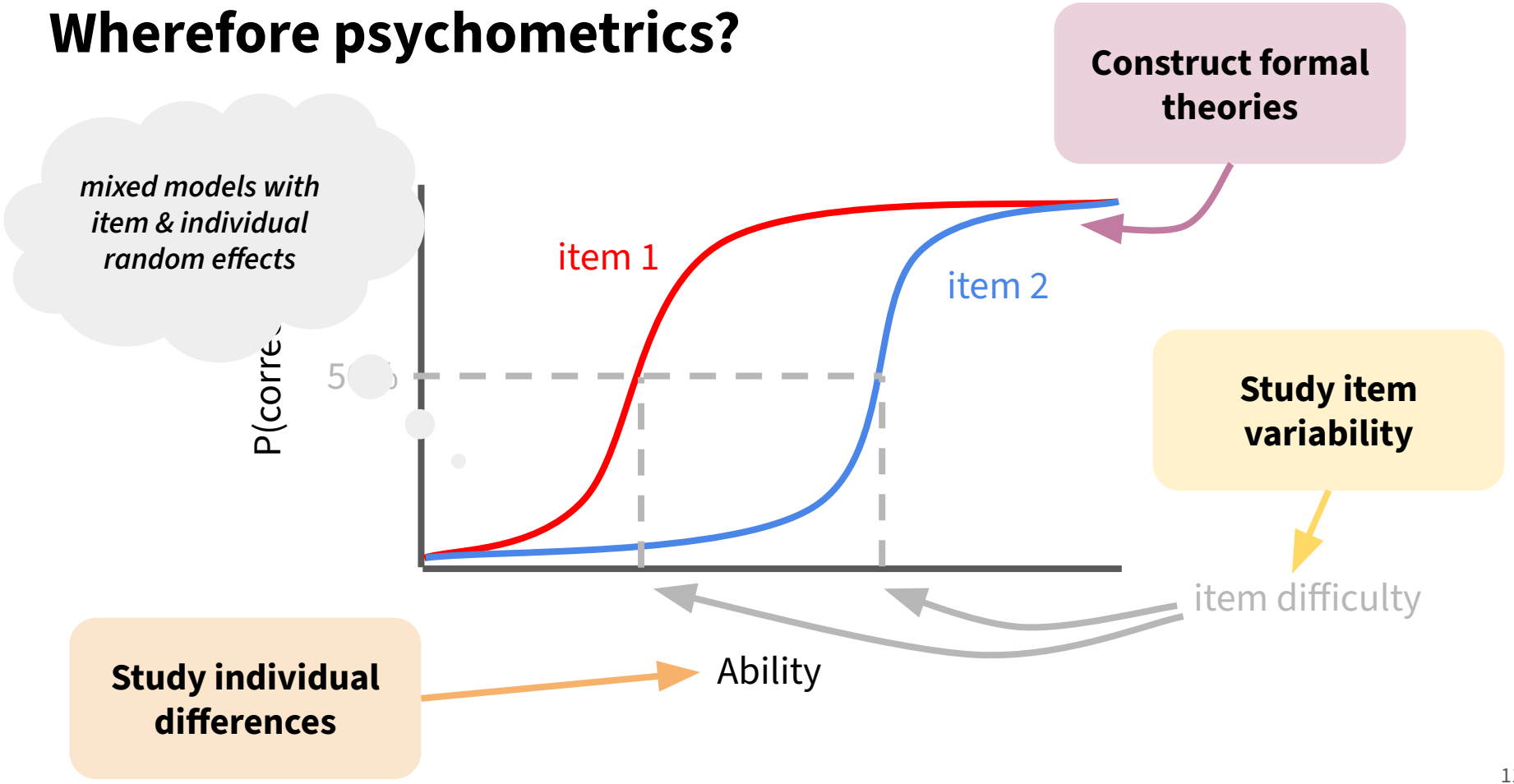
Instrumentation	*Measurement model*	*Latent structure*
• Scaling • Reliability • Bias • Measurement invariance • Differential item functioning	• Classical test theory • **Item response theory** • Generalisability theory	• Factor analysis • Latent variable models • Network theory • Structural equation modelling

Borsboom et al. (2004)

A psychometric approach



Wherefore psychometrics?



Studying item variability

Why?

- Understand dimensions of variability in the world
  (e.g., Erhardt et al., 2023; García et al., 2023; Judd et al., 2024)
- Develop effective scales and tests
  (e.g., Kachergis et al., 2022; McCowan & McCowan, 1999)
- Ensure item applicability
  (e.g., Kubinger, 2008; Peterson et al., 2003; Shahsavar et al., 2023)

How?

- Explicit item-wise and inter-item analyses
  (e.g., Piedmont, 2014; Rasch, 1993)
- Metrics for assessing scale/test properties
  (e.g., Cooper, 2024; Kalkbrenner, 2021; Magis et al., 2010)

Studying individual variability

Why?

- Suggest mechanisms and developmental pathways
  (e.g., Boogert et al., 2018; Fisher-Baum et al., 2018; Hofman et al., 2024; Oakes & Rakison, 2019; van der Maas et al., 2006)
- Understand natural variability → construct norms and diagnostics
  (e.g., Frank et al., 2021; de Ron et al., 2019; Habibzadeh et al., 2016; Lenhard et al., 2019; Schaaf et al., 2023)

How?

- Better measurement of latent traits
  (e.g., Cooper et al., 2017; Rouder & Haaf, 2019; Vermeent et al., 2024)

Constructing formal theories

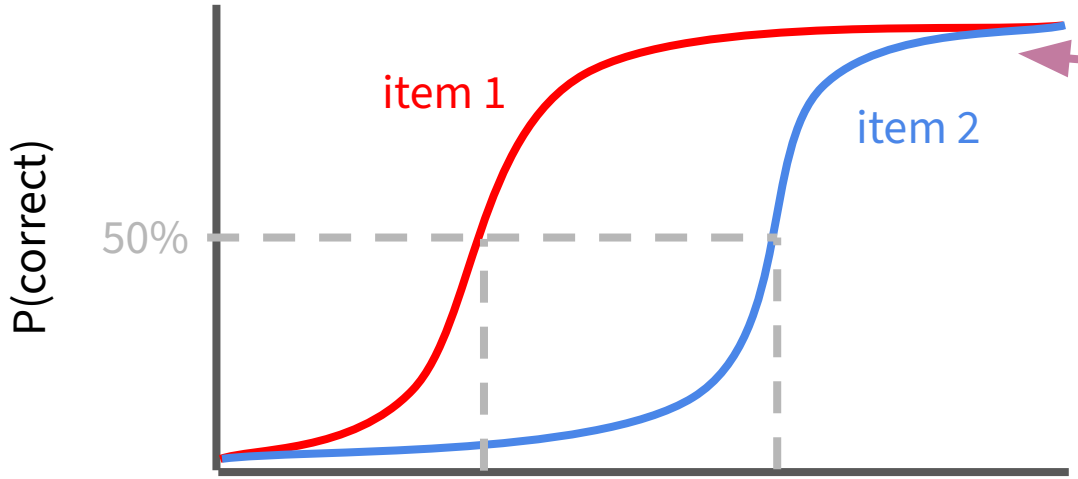
Why?

- Psychology lacks these!
  (e.g., Muthukrishna & Henrich, 2019; Oberauer & Lewandowsky, 2019)
- Greater clarity, testability, and comparability
  (e.g., Guest & Martin, 2021; Smaldino, 2020)
- Theory building
  (e.g., Borsboom et al., 2021; Fried, 2021; Robinaugh et al., 2021)

How?

- Specification of measurement model
  (e.g., Sijtsma & van der Ark, 2020; Wilson, 2013)
- Engagement with response processes and linking hypotheses
  (e.g., Padilla & Benítez, 2018; van Grinsven, 2023; Yurovsky et al., 2012)

Wherefore psychometrics!



Construct formal theories

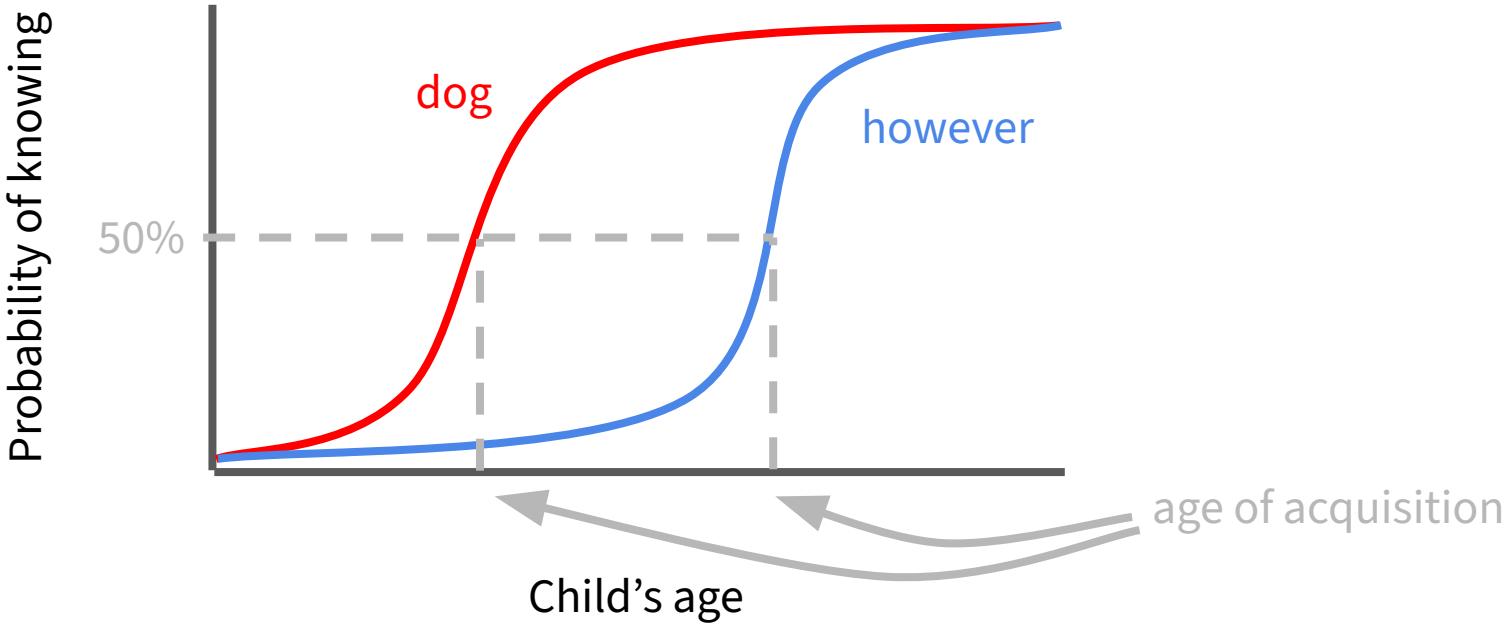
Study item variability

Study individual differences

Case study: Item response theory in early vocabulary learning

Tan, Marchman, & Frank (2024)

Early word learning



Goodman, Dale, & Li (2008)

Communicative Development Inventories (CDIs)

1. SOUND EFFECTS AND ANIMAL SOUNDS (12)		
baa baa	meow	uh oh
choo choo	moo	vroom
cockadoodledoo	ouch	woof woof
grrr	quack quack	yum yum

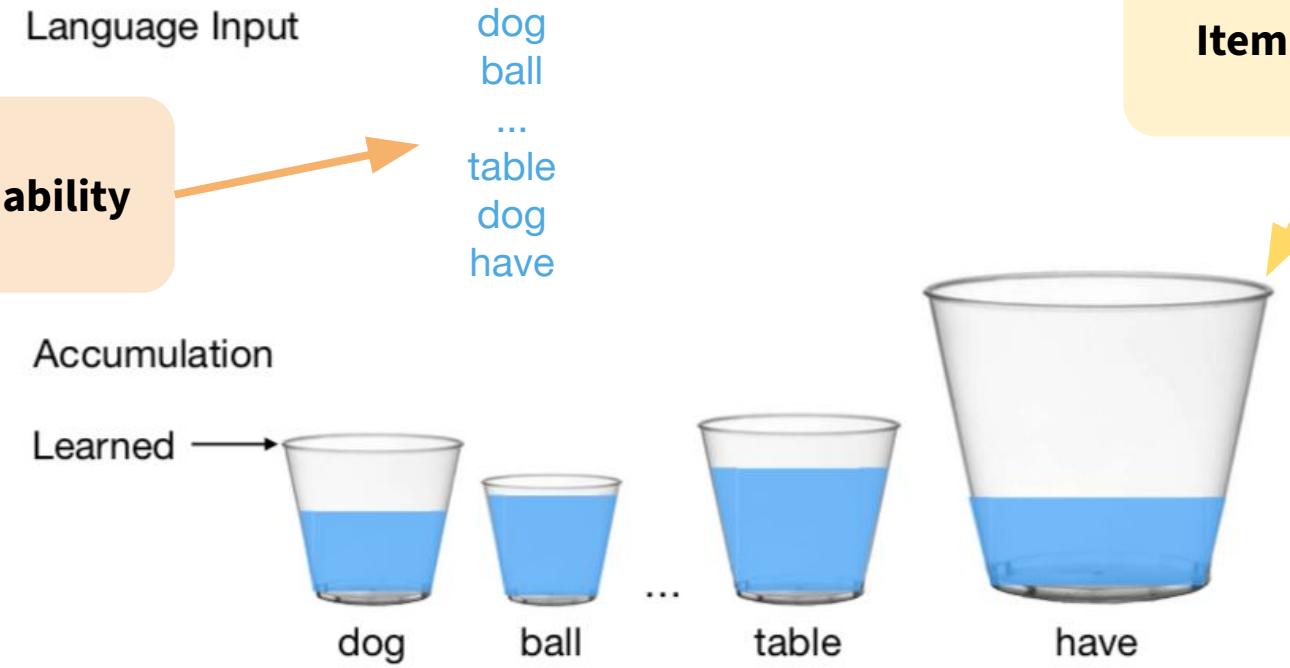
2. ANIMALS (Real or Toy) (43)		
alligator	duck	penguin
animal	elephant	pig
ant	fish	pony
bear	frog	puppy
bee	giraffe	rooster
bird	goose	sheep
bug	hen	squirrel
bunny	horse	teddybear
butterfly	kitty	tiger
cat	lamb	turkey
chicken	lion	turtle
cow	monkey	wolf
deer	moose	zebra
dog	mouse	
donkey	owl	

Accumulator models

Formal theory

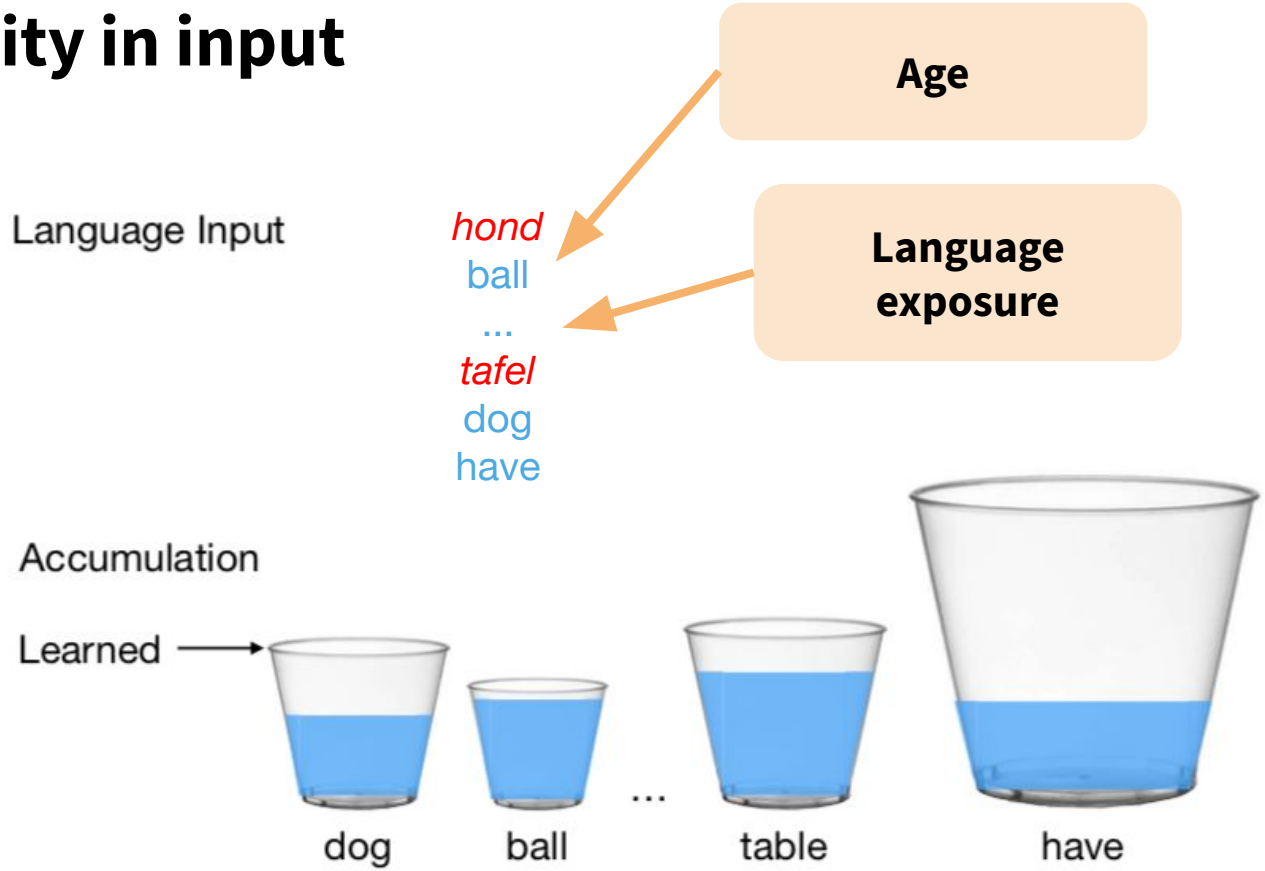
Item difficulty

Individual ability

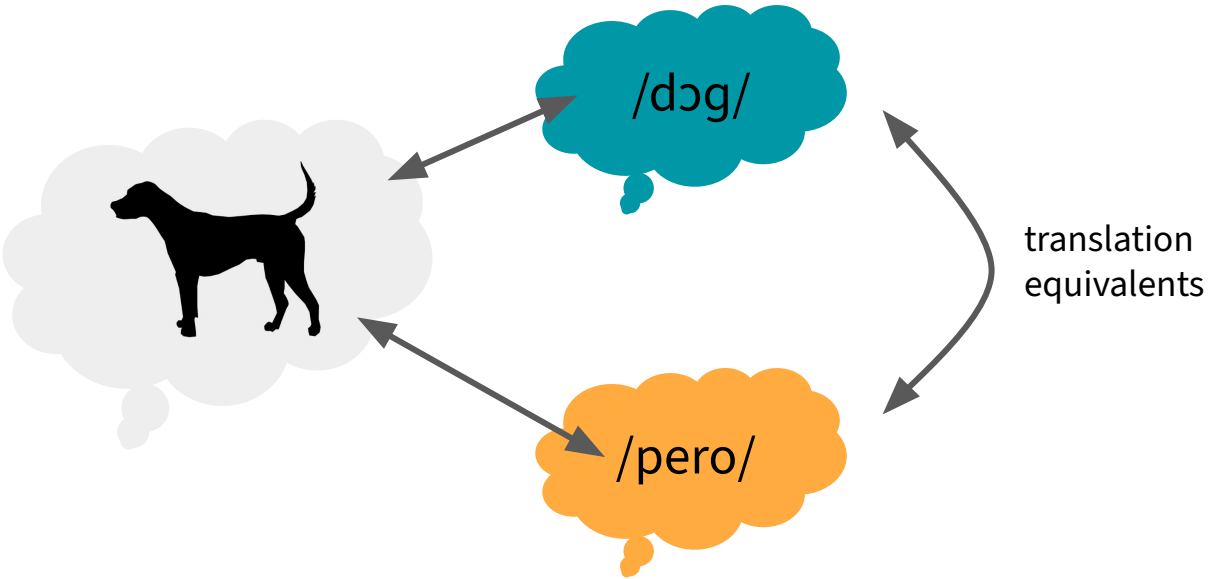


Kachergis, Marchman, & Frank (2022)

Variability in input



Translation equivalents



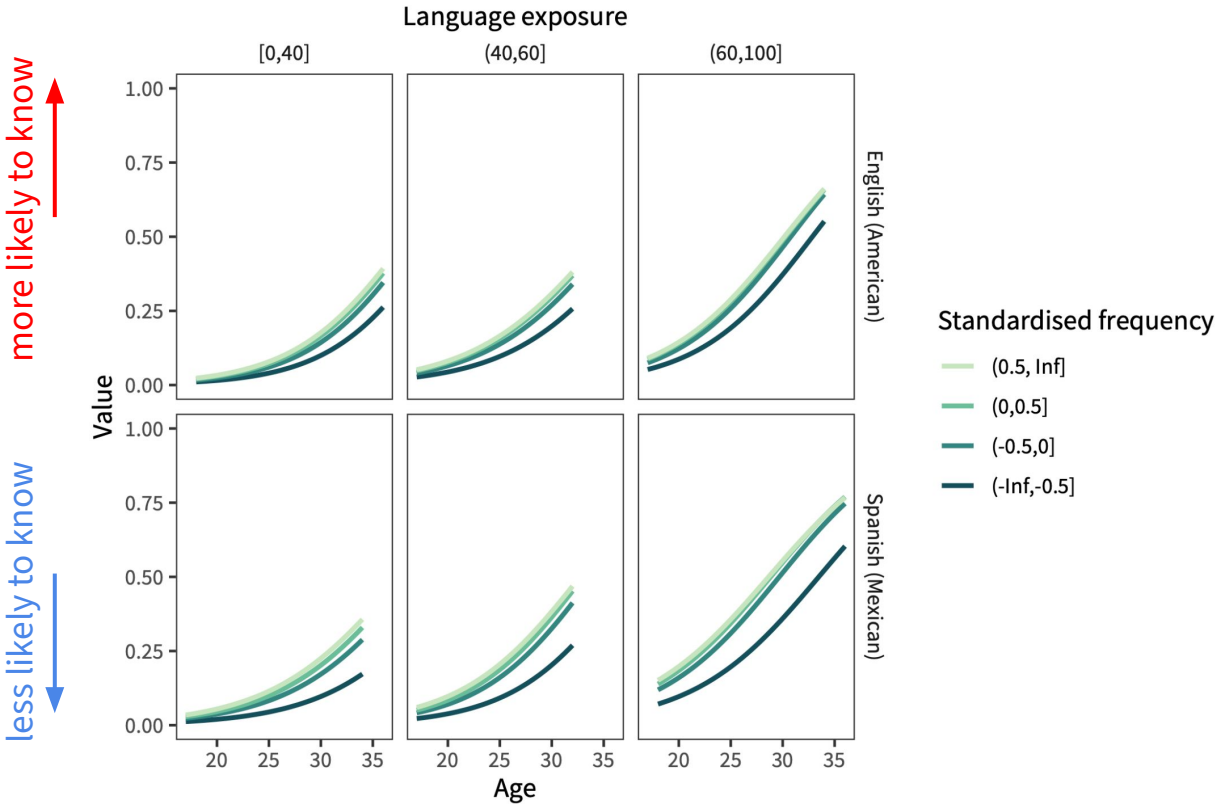
Method: Predictors

Individual-level	*Individual x Item*	*Item-level*
Age	TE knowledge	Frequency
Proportion exposure	TE x Phonological similarity	Concreteness
		MLU-w

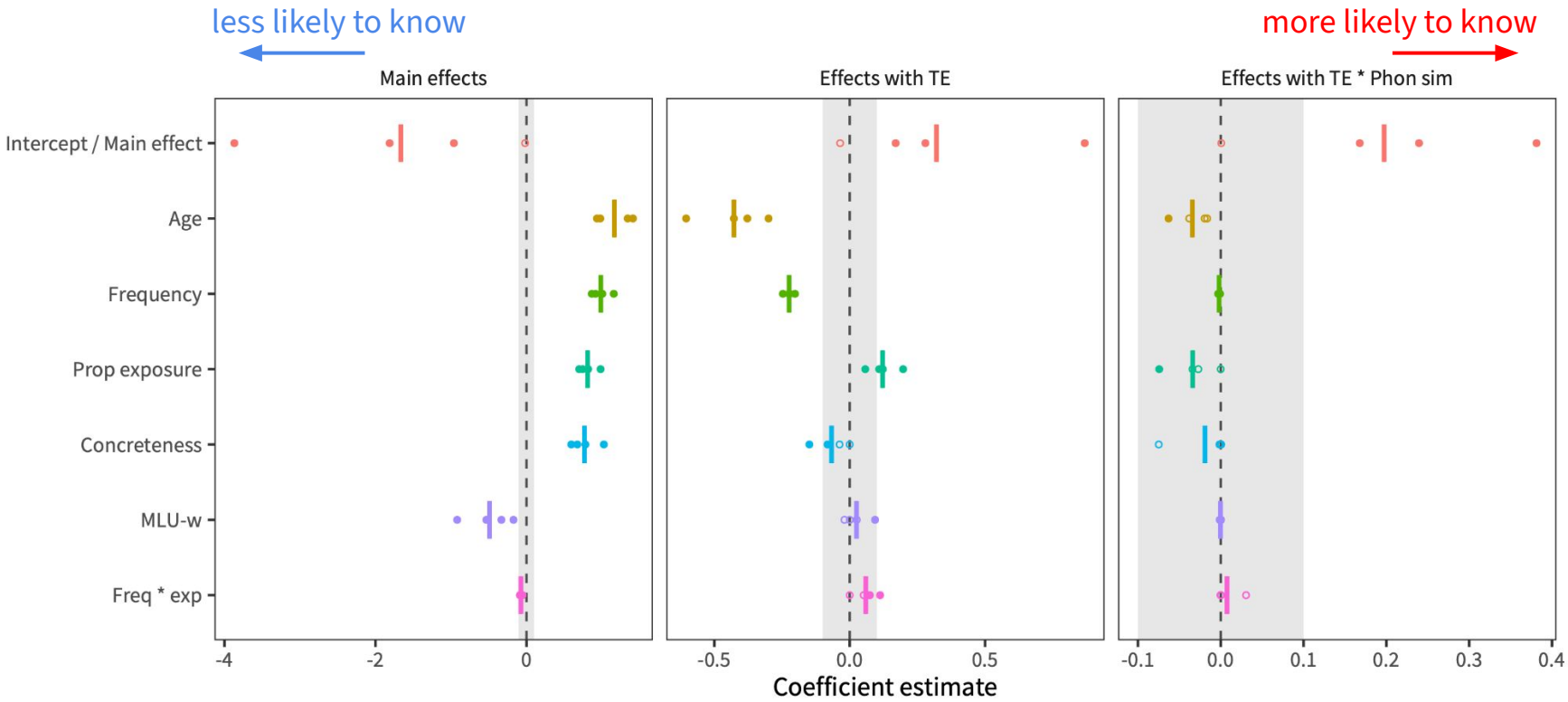
Method: Datasets

Languages	Contributor	*N*
English–Spanish	Virginia Marchman	147
English–Spanish	Erika Hoff	165
English–French	Diane Poulin-Dubois	59
English–French	Mitchell et al. (2022)	48

Analysis: The role of input



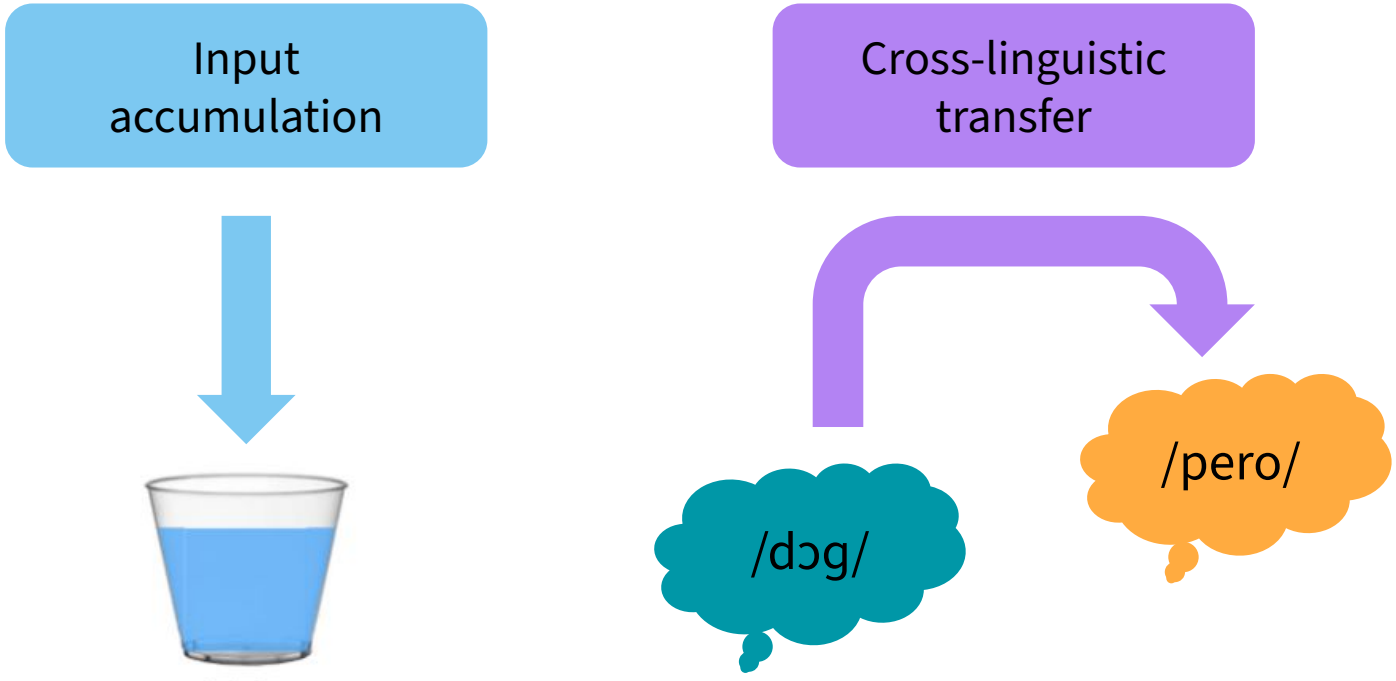
Analysis: Predictor effects



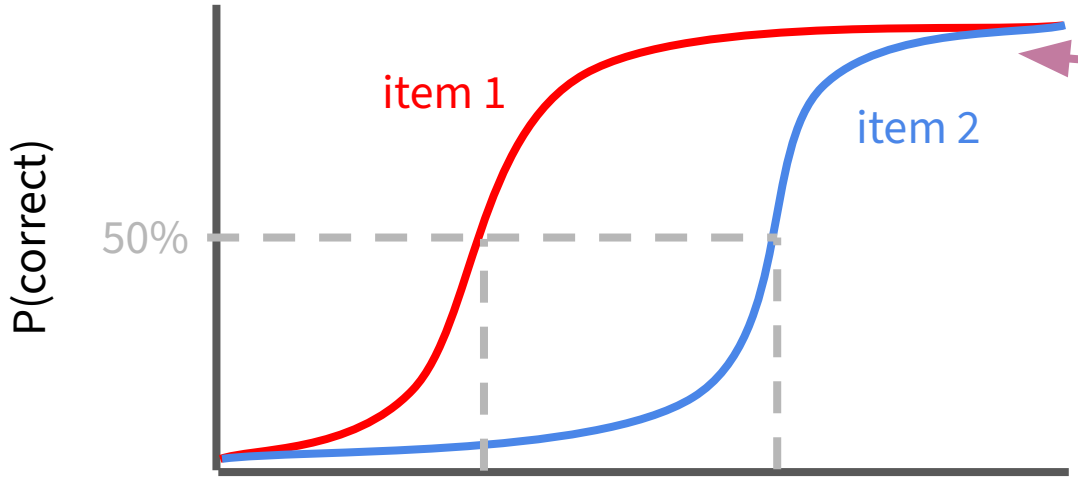
Discussion

- Bilingual children’s word learning largely aligns with monolingual word learning
- Predictors are consistent across datasets and languages
- TE knowledge gives an advantage for younger children, for less frequent words, and for TE pairs that are more phonologically similar

Discussion: Two-route model



Wherefore psychometrics!



Construct formal theories

Study item variability

Study individual differences

George Kachergis

Virginia Marchman

Michael C. Frank

ty+q?