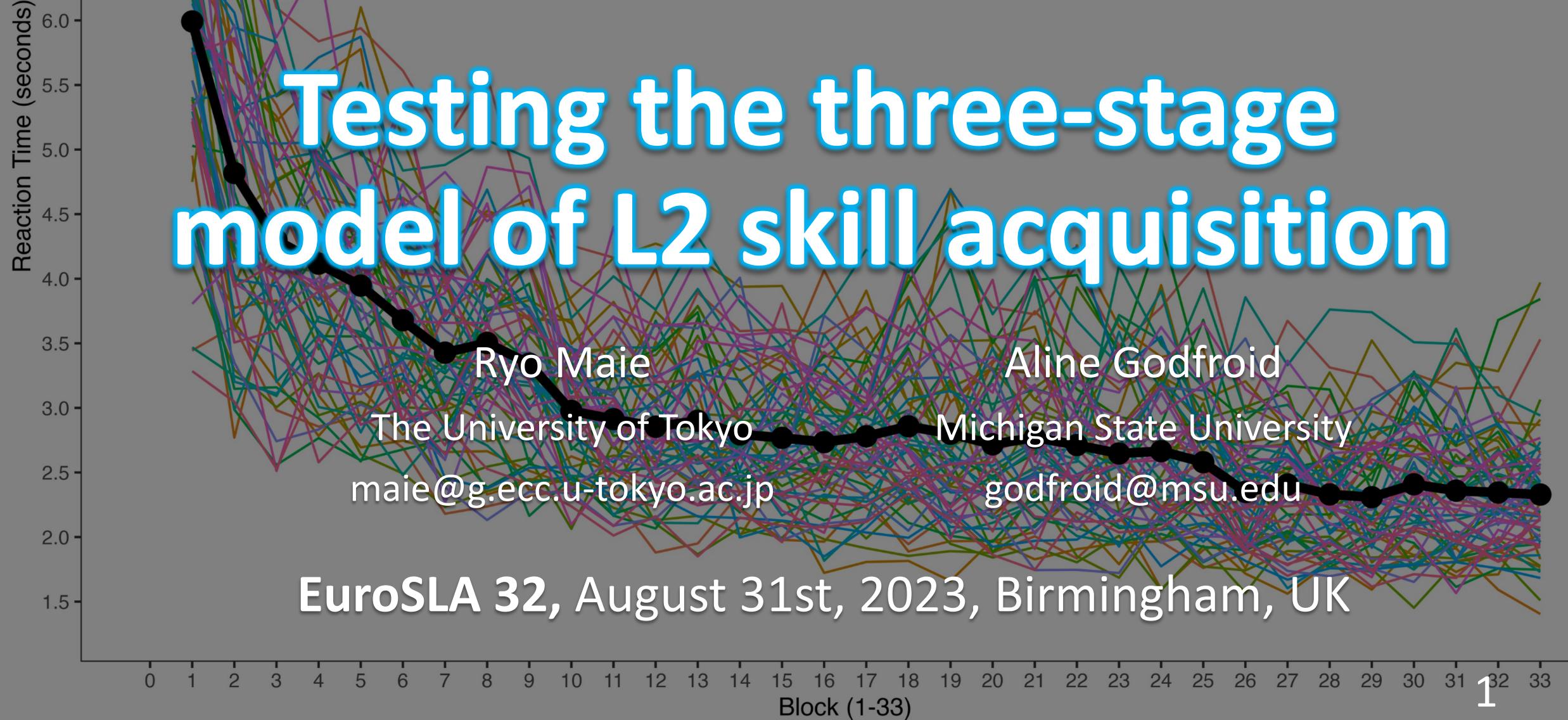




Testing the three-stage model of L2 skill acquisition



L2 learning as skill acquisition



Skill acquisition theory

- Learning as acquisition of **skills**
- Learning done by domain-general mechanisms

L2 skill acquisition

- L2 learning = skill acquisition
- John Anderson (1983):
 - *“language is cut from the same cloth as the other cognitive processes, but it differs in its cognitive content”* (p. 261)

Evidence on L2 skill acquisition

Skill acquisition ...

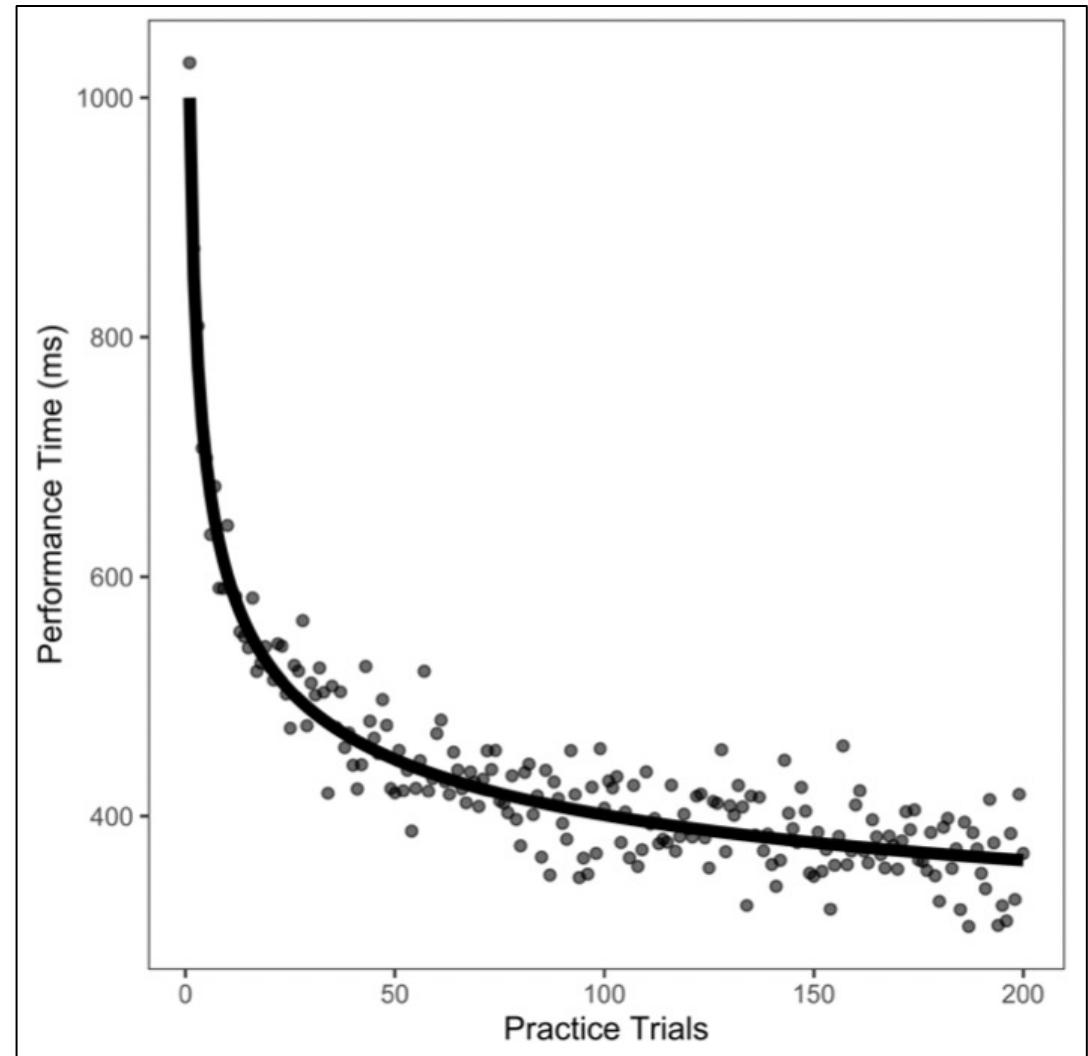
1. follows **the power-law of practice**

DeKeyser (1997), Ferman et al. (2009), Hui (2020), Maie (2020)

2. leads to **skill-specific** competence

- comprehension vs. production

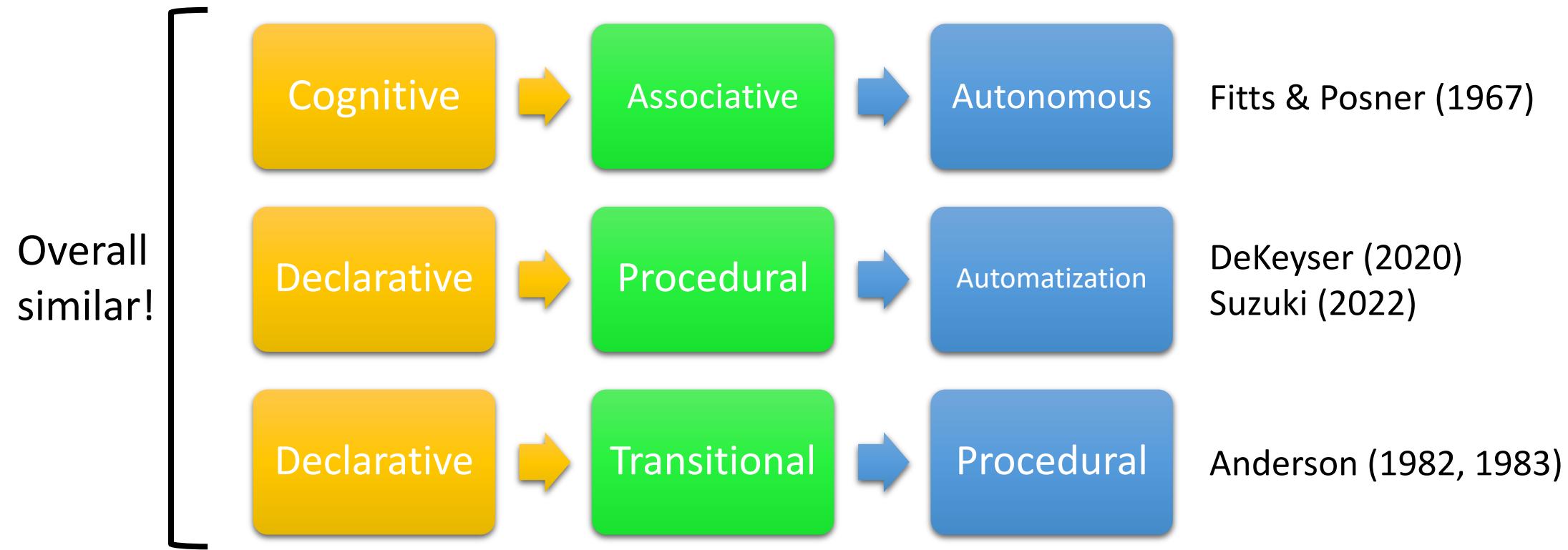
DeKeyser (1997), DeKeyser & Sokalski (1996), Li & DeKeyser (2017), Suzuki & Sunada (2019)



The three-stage model

L2 skill acquisition is **a three-stage process**

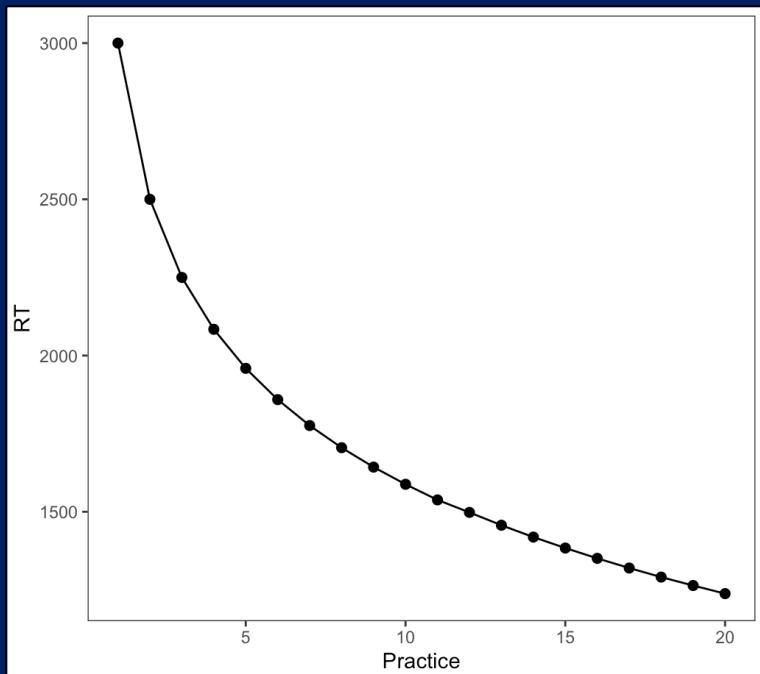
- skill “development from [1] initial representation of knowledge [2] through initial changes in behavior [3] to eventual fluent, spontaneous, largely effortless, and highly skilled behavior” (DeKeyser, 2020, p. 83)



Race Model (Instance Theory)

Logan (1988)
Logan (2002)

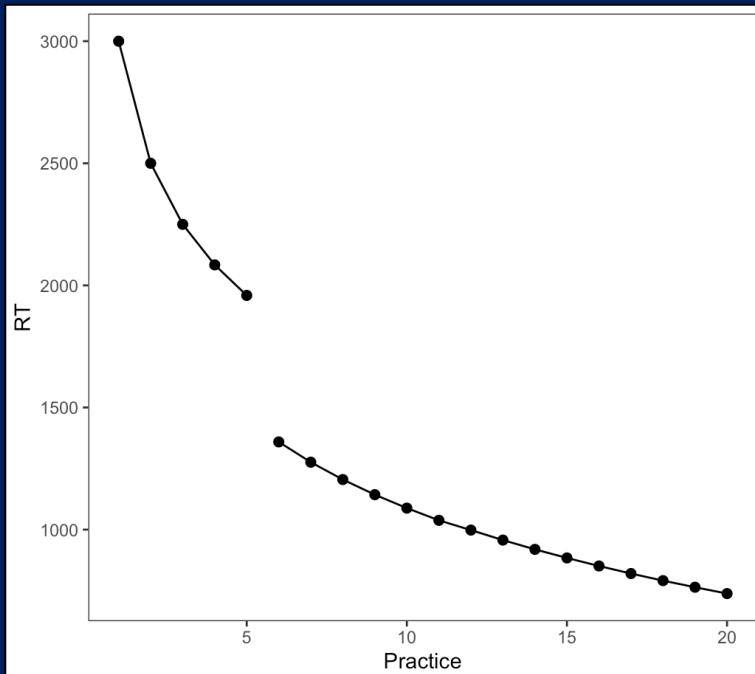
Instance accumulation



CMPL

Rickard (1997)
Rickard (2004)

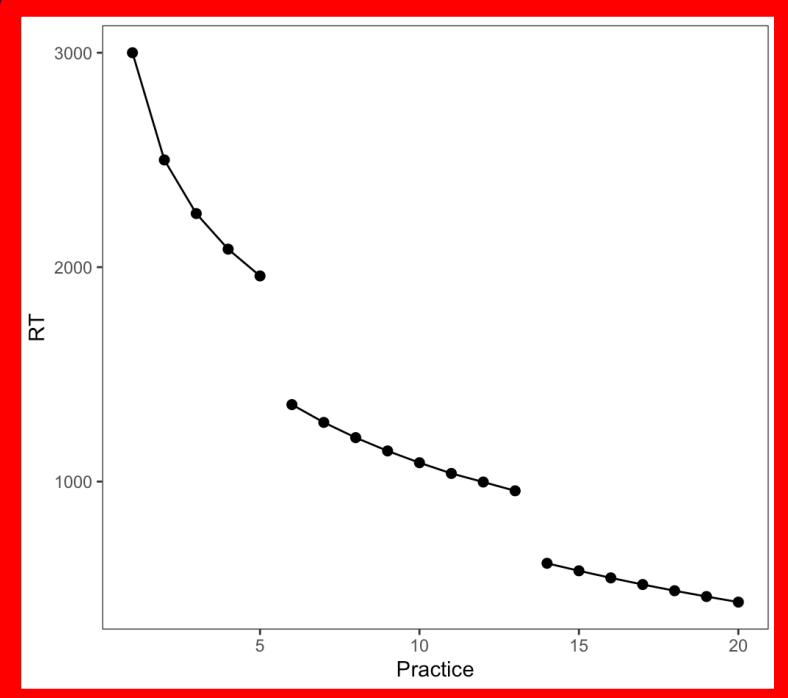
Algorithm -> Item retrieval



ACT-R

Anderson (1982)
Anderson (2007)

Declarative -> Transitional
-> Procedural



One-stage
model

Two-stage
model

Three-stage
model

Individual differences in L2 Skill acquisition

The Declarative-Procedural Model (Ullman, 2004, 2014, 2020)

Grammar: Declarative → Procedural with proficiency

→ Meta-analysis by Hamrick et al. (2018) (but see Oliveira et al., 2023)

L2 automatization research

- Procedural learning ability predicting the degree of automatization from practice (Pili-Moss et al., 2020; Suzuki, 2017)

→ Number of stages? Mechanisms underlying each stage?

The Study: Research Questions

When practicing a novel foreign language ...

1. How many stages of skill acquisition do L2 learners go through?
2. Which memory systems, declarative and/or procedural memory, are implicated in each learning stage?



6 days (6 hours in total)

Procedure ($n = 65$)

Day 1 Cognitive Tests	Day 2 Explicit Instruction	Day 3 Practice Session 1	Day 4 Practice Session 2	Day 5 Practice Session 3	Day 6 Practice Session 4
-----------------------------	----------------------------------	--------------------------------	--------------------------------	--------------------------------	--------------------------------

Day 1 (39 mins)	Min
1. Background questionnaire	1
2. Two-choice response task	3
3. Alternating serial reaction time task	15
4. Statistical learning task	20

Day 2 (60 mins)	Min
1. Continuous visual memory task	10
2. LLAMA-B	10
3. Explicit instruction of Mini-Nihongo	20
4. Vocabulary and grammar tests	5
5. Warmup practice of Mini-Nihongo	15

Day 3 (70 mins)	Min
1. Vocabulary and grammar tests	5
2. Comprehension practice	20
3. Production practice	45

Day 4 (65 mins)	Min
1. Vocabulary and grammar tests	5
2. Production practice	40
3. Comprehension practice	20

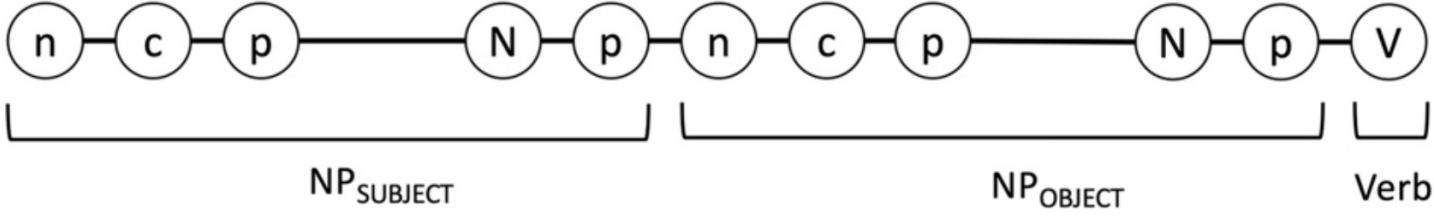
Day 5 (60 mins)	Min
1. Vocabulary and grammar tests	5
2. Comprehension practice	20
3. Production practice	35

Day 6 (55 mins)	Min
1. Vocabulary and grammar tests	5
2. Production practice	30
3. Comprehension practice	20

Language

Mini-Nihongo (Mueller, 2006)

- A miniature language based on Japanese
- Only S-O-V order (canonical in Japanese)

Grammar structure of Mini-Nihongo			
			
Vocabulary items and case-markers of Mini-Nihongo			
N [noun]	=	hato (pigeon), amo (duck), nezumi (mouse), neko (cat)	
V [verb]	=	tobikoeru (jump over), tsukamaeru (capture), oikakeru (chase away), otozureru (visit)	
n [number]	=	ichi (one), ni (two)	
c [classifier]	=	wa (bird class), hiki (small animal class)	
p [postposition]	=	ga (nominative), o (accusative), no (genitive)	

Mueller, J. L. (2006). L2 in a nutshell: The investigation of second language processing in the miniature language model. *Language Learning*, 56(s1), 235–270.

Language Instruction



Explicit instruction of Mini-Nihongo (Day 2)

Explicit-deductive instruction in the form of a 19-minute video

Grammar explanation with examples and exercises

<https://osf.io/vh6ap>



Vocabulary and grammar knowledge test (Day 2 - 6)

Vocabulary: A picture-word matching task for nouns and verbs ($k = 24$)

Grammar: A metalinguistic knowledge test with a fill-in-the-blank format ($k = 9$)

Day 1
Cognitive
Tests

Day 2
Explicit
Instruction

Day 3
Practice
Session 1

Day 4
Practice
Session 2

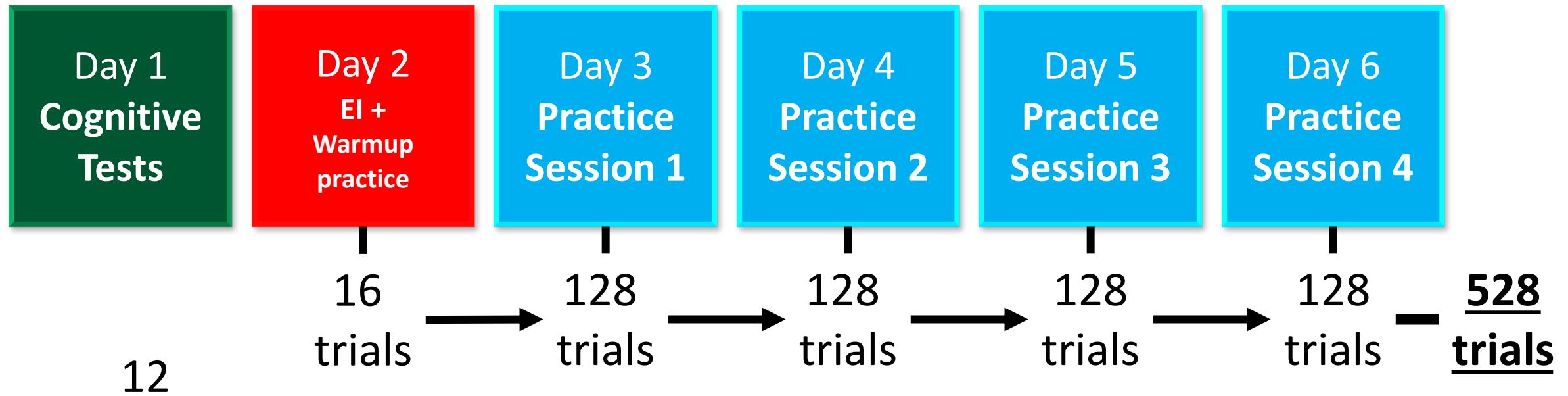
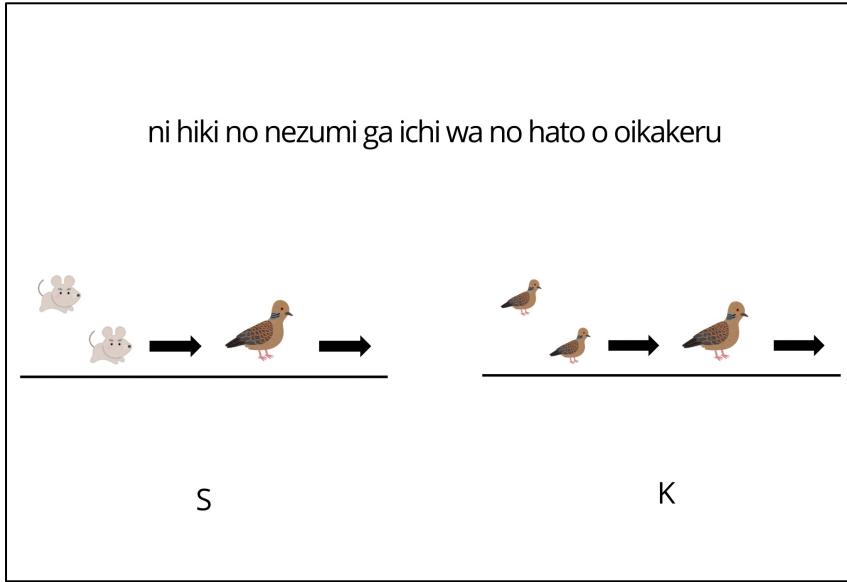
Day 5
Practice
Session 3

Day 6
Practice
Session 4

Language Instruction

Mini-Nihongo
(Mini-Japanese)

Comprehension Practice



Comprehension Practice



Cognitive Tests

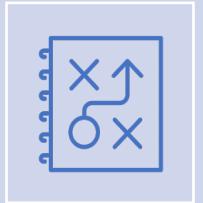


Declarative Memory

Continuous Visual Memory

Task (domain-general)

LLAMA-B (domain-specific)



Procedural Memory

Alternating reaction time task

(domain-general)

Statistical learning task

(domain-specific)

Measured Variables



Accuracy (0 or 1)



Reaction Time
(seconds)

We looked at
coefficient of
variation (CV)
too but ...



CVMT (d-prime)



LLAMA-B (0-100)

Factor
Scores!

Combined!



ASRT (milliseconds)



SL (0-24)

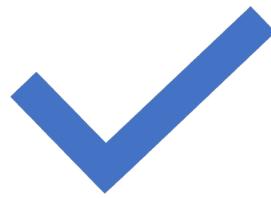
▷ Analysis

1. Hidden Markov Modeling (Python)

- takes RT as the dependent variable
- estimates the probability of each participant staying in each learning stage on each practice trial
- identifies **the number** of skill acquisition stages by comparing one, two, and three-state models

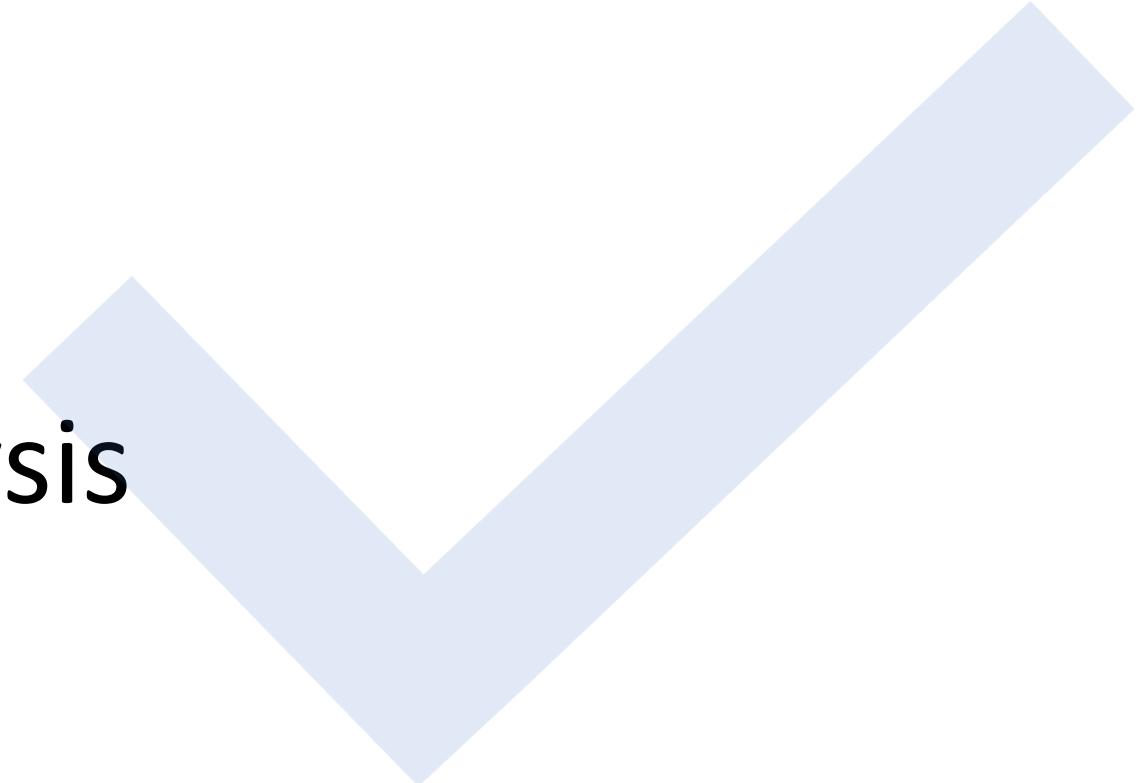
2. Regression Modeling (R)

- identifies **the nature** of skill acquisition stages by investigating which cognitive abilities predict learning in each learning stage

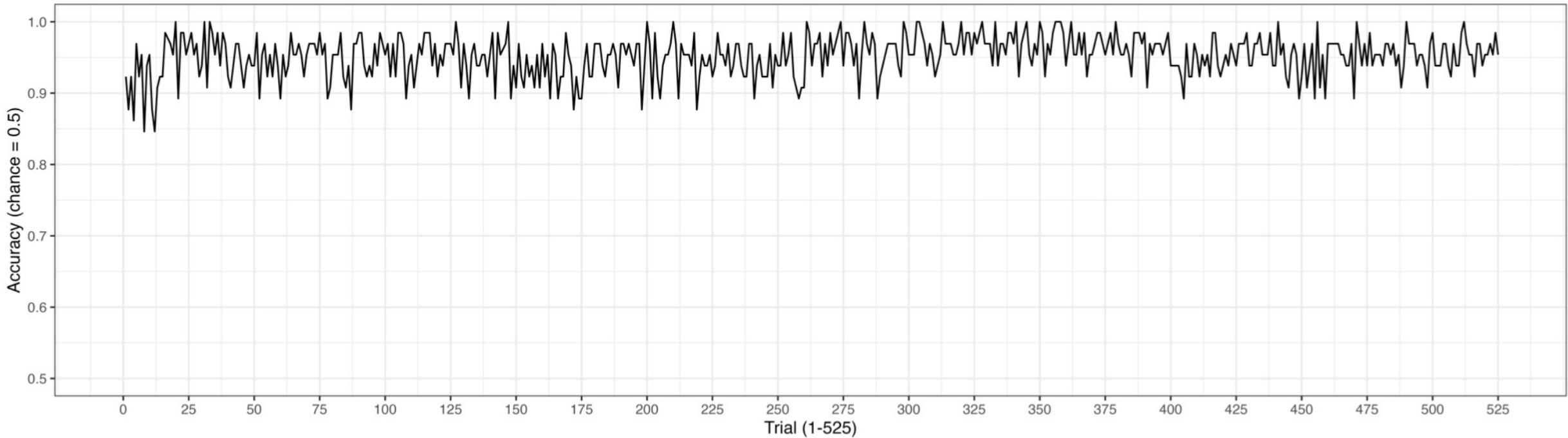


Results

Descriptive Analysis

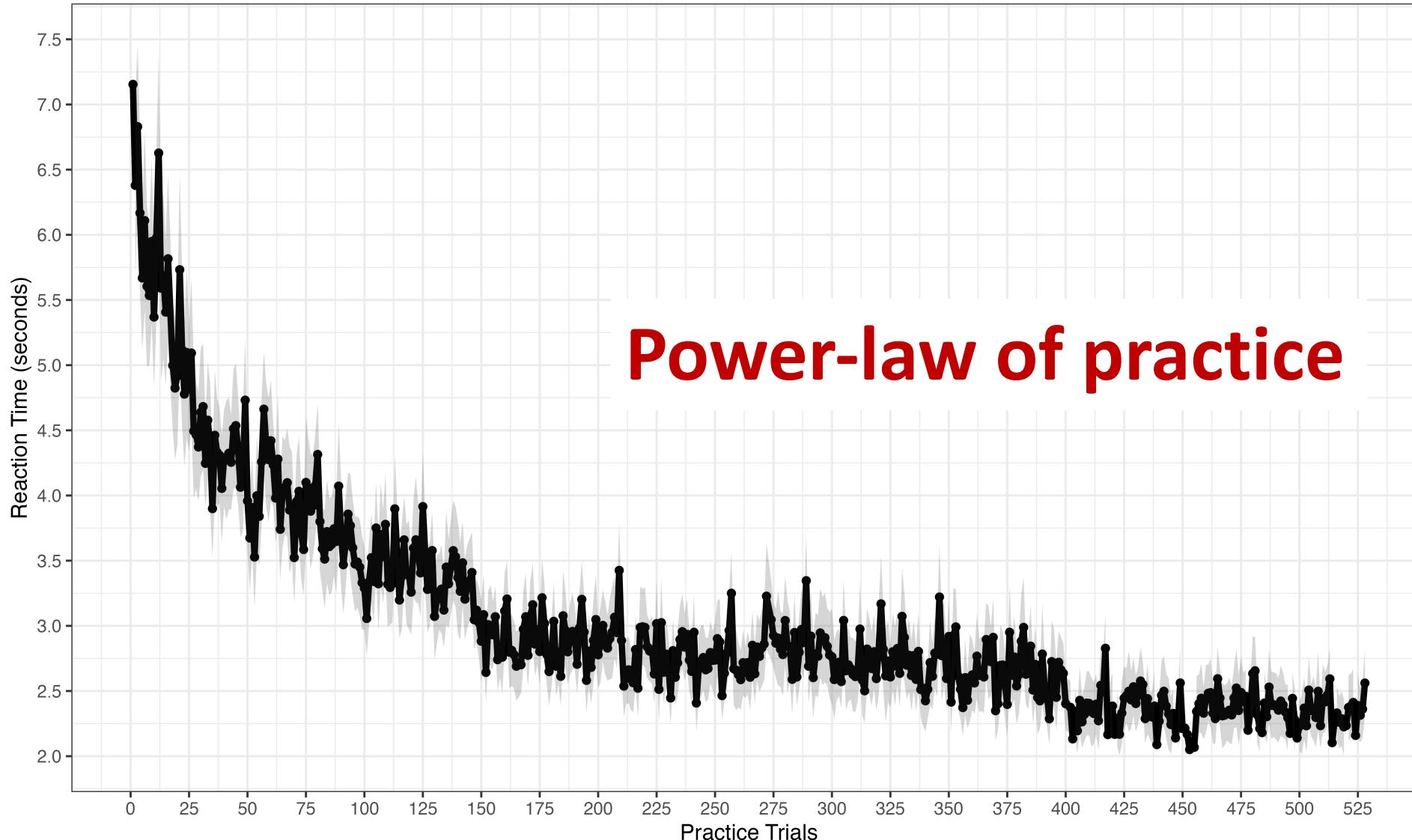


Accuracy – comprehension practice



Even from the first trial, participants showed very accurate performances (90%)

Reaction Time



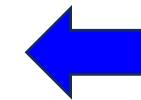


Hidden Markov Modeling

The *number* of stages

HMM: Model Comparison

	BIC	Diff	Pr(M)
One-stage	39902	3239	.000
Two-stage	36700	37	.000
Three-stage	36663	0	≈ 1.000



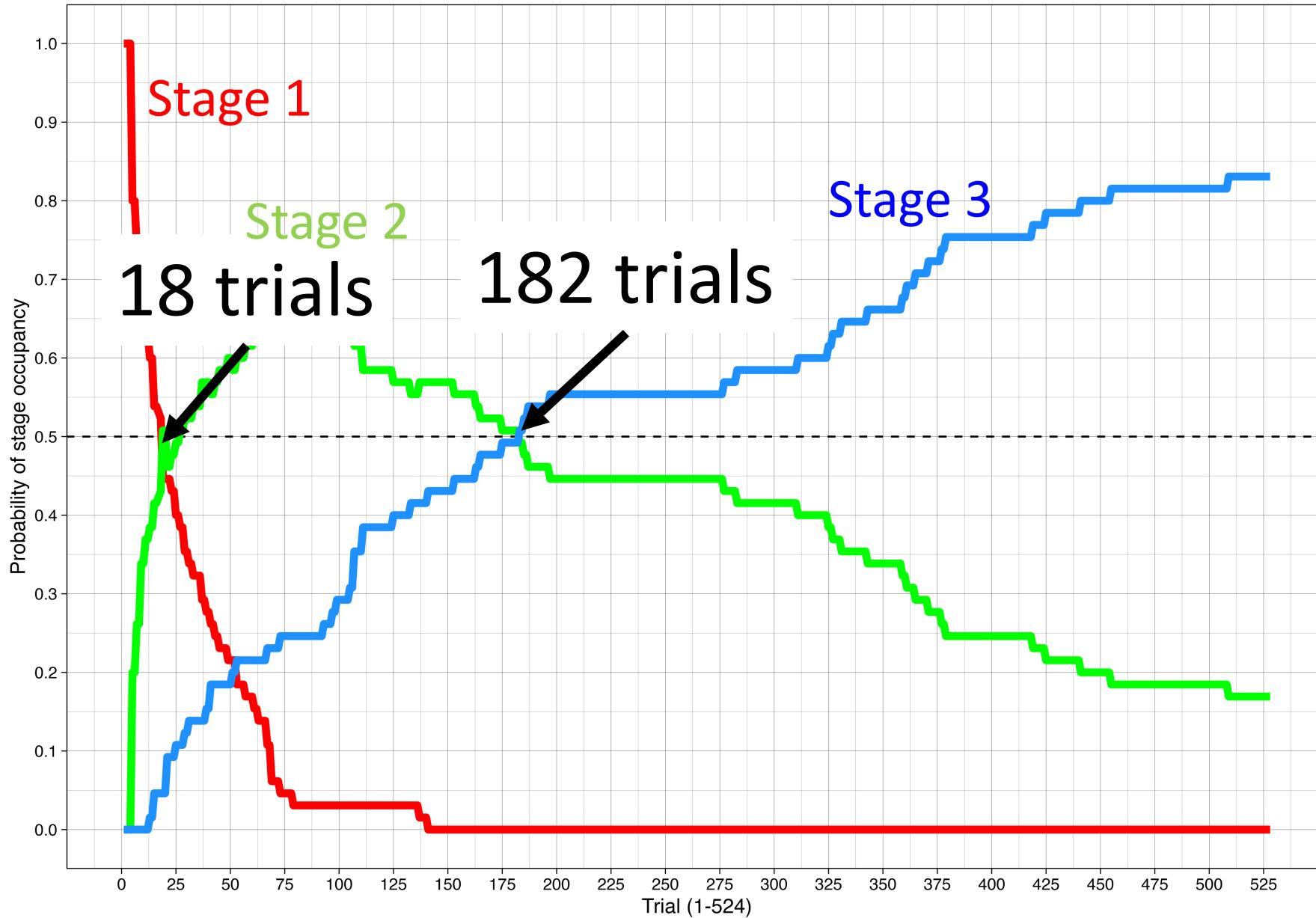
The best fitting model
&
Way(!) more probable
than the other models

BIC = Bayesian information criterion (lower, better)

Diff = Difference in BIC

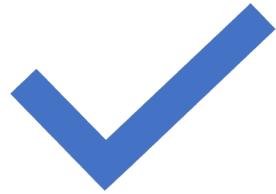
Pr(M) = relative likelihoods of each model against the others

HMM: Stage Occupancy (Group)



Stage Occupancy (Individual)





Regression Modeling

The *nature* of stages

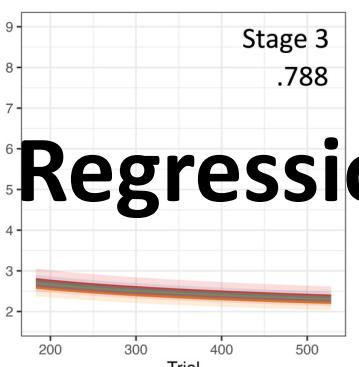
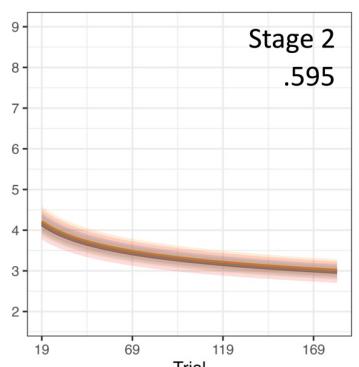
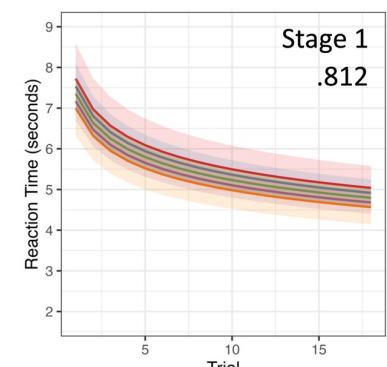
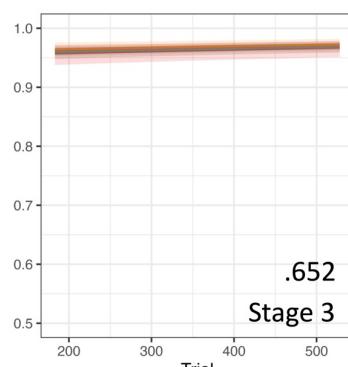
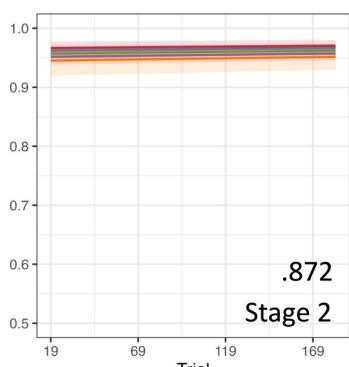
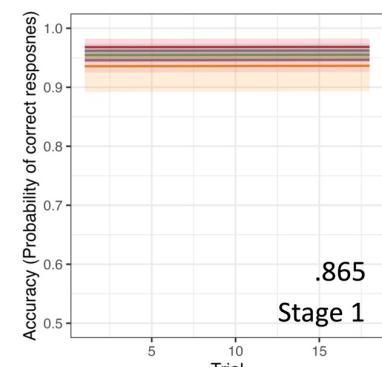
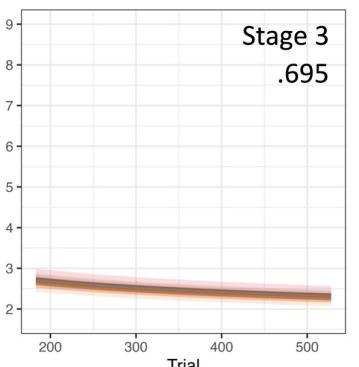
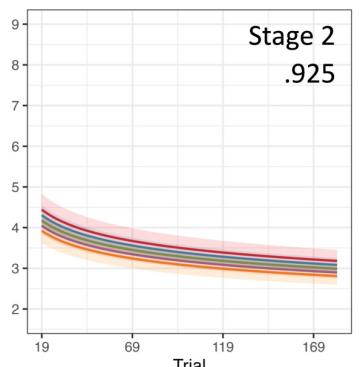
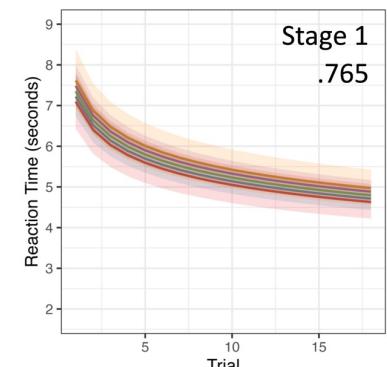
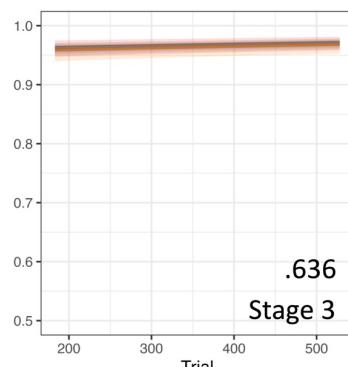
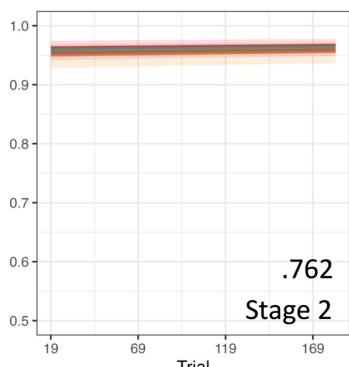
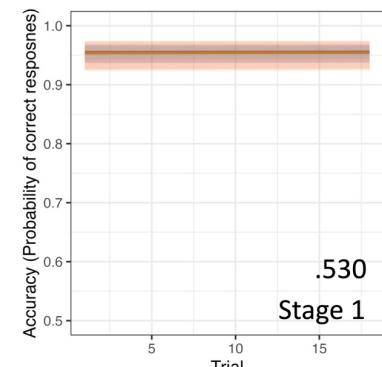
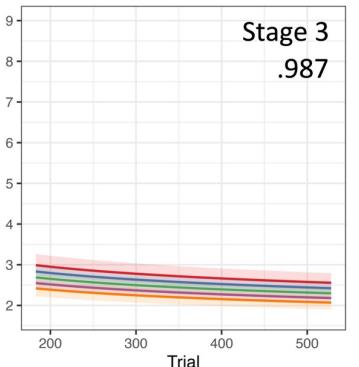
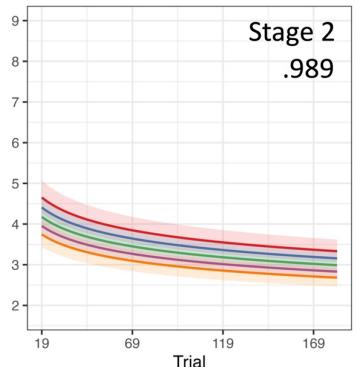
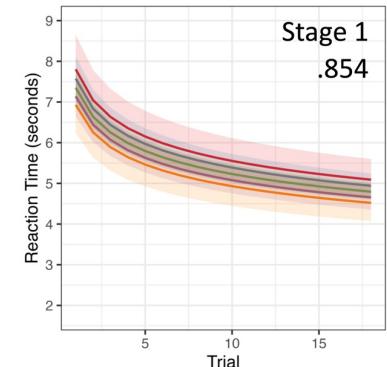
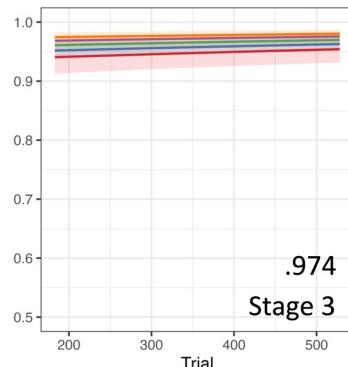
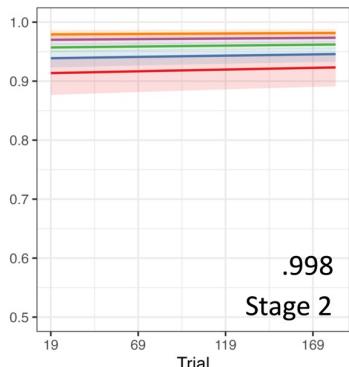
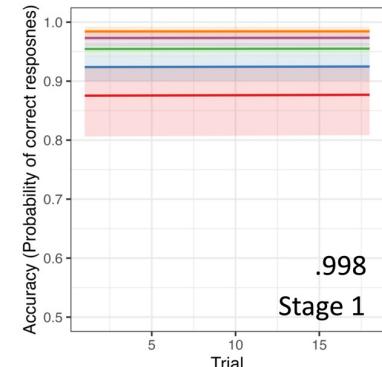
Regression Modeling: Hypothesis



		Accuracy		
		Stage 1	Stage 2	Stage 3
Declarative	Declarative	✓	✓	X
	Procedural	X	✓	✓

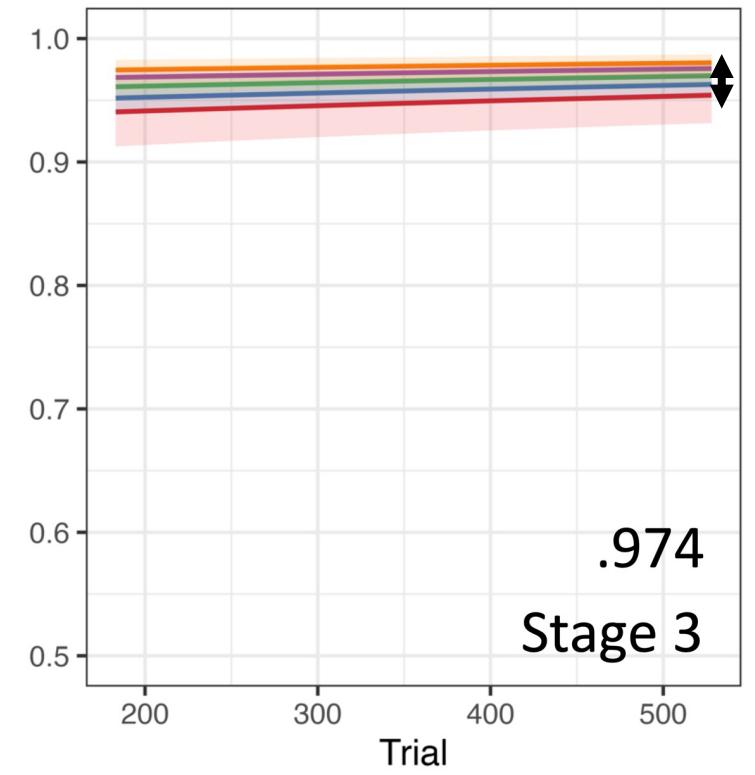
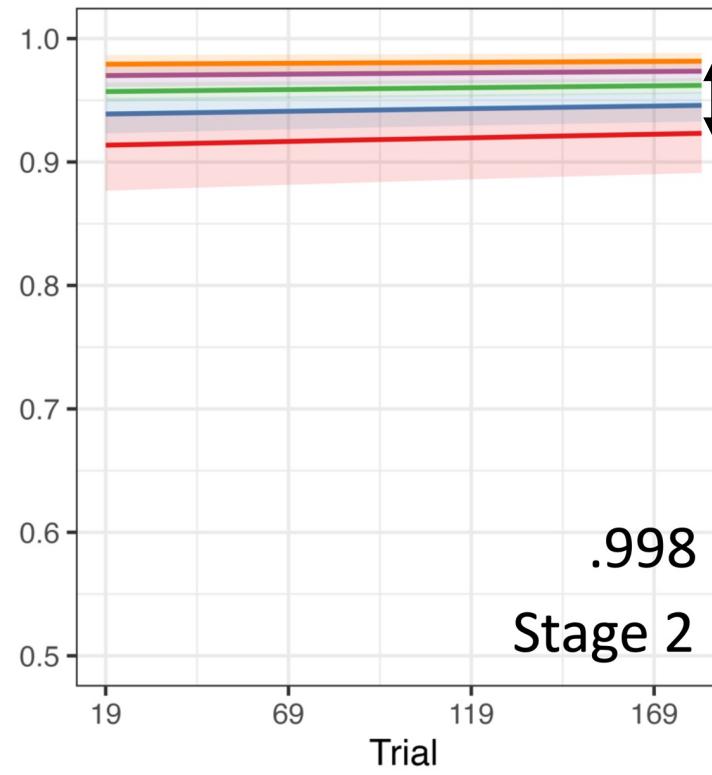
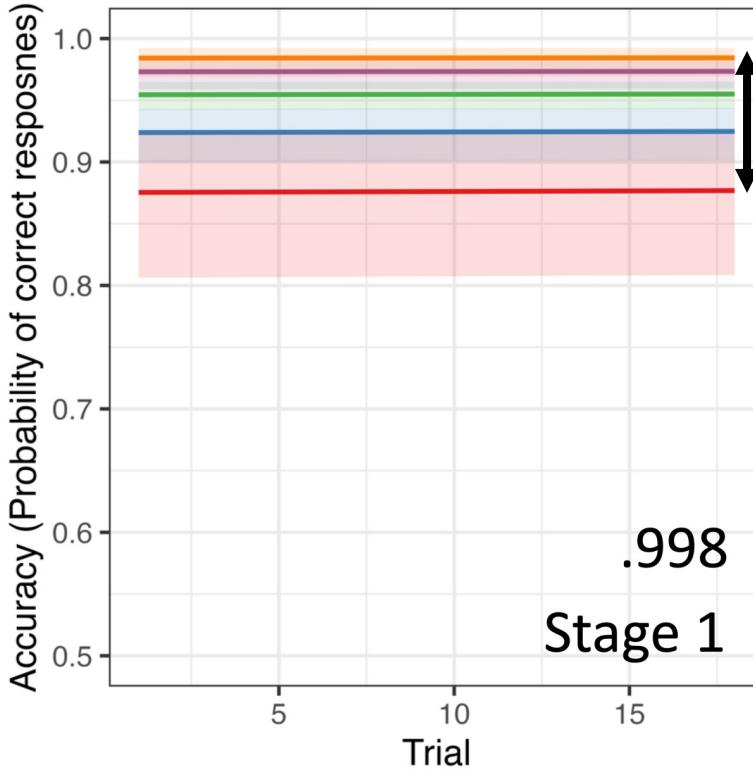
		Reaction Time		
		Stage 1	Stage 2	Stage 3
Declarative	Declarative	✓	✓	X
	Procedural	X	✓	✓

		Reaction Time		
		Stage 1	Stage 2	Stage 3
Declarative	Declarative	✓	✓	X
	Procedural	X	✓	✓



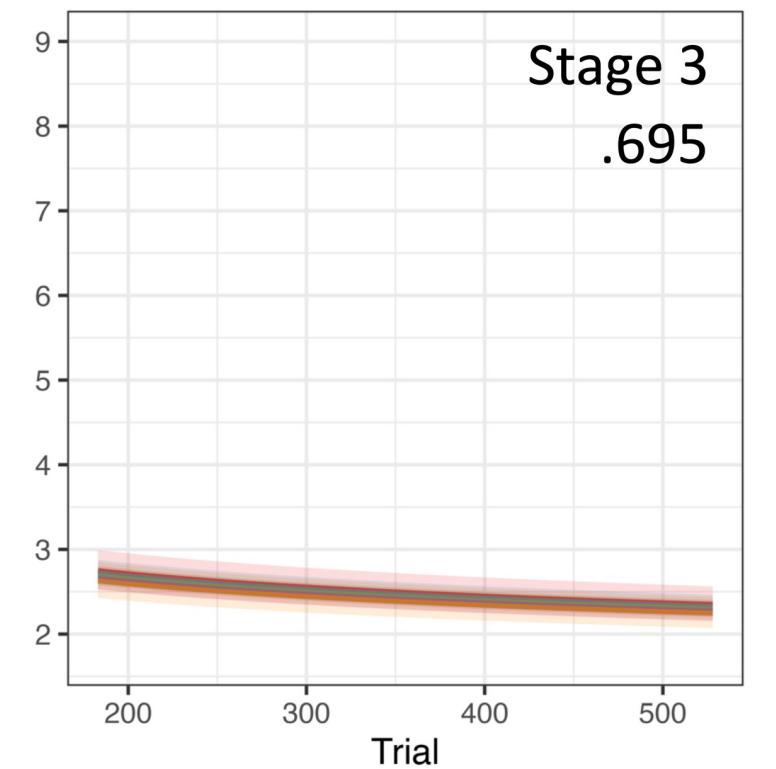
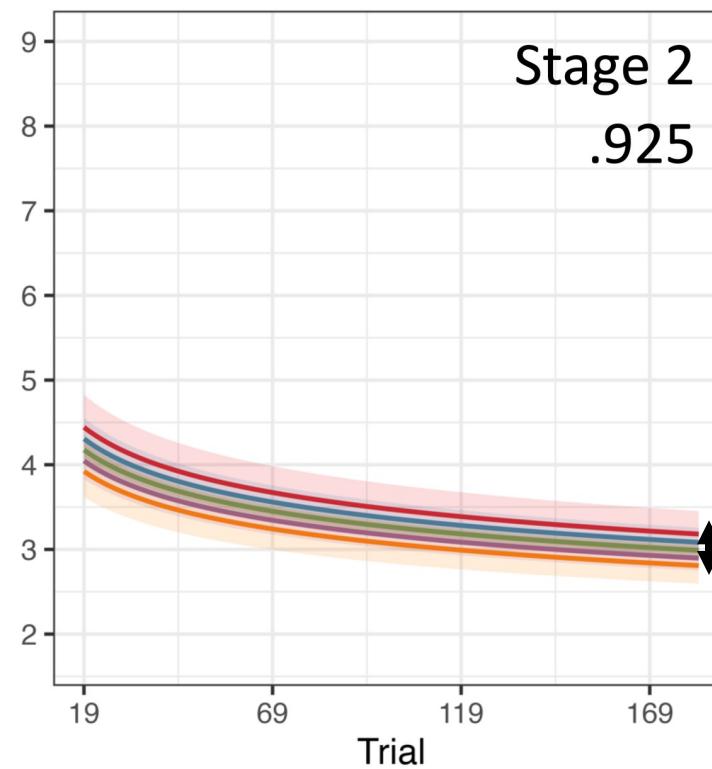
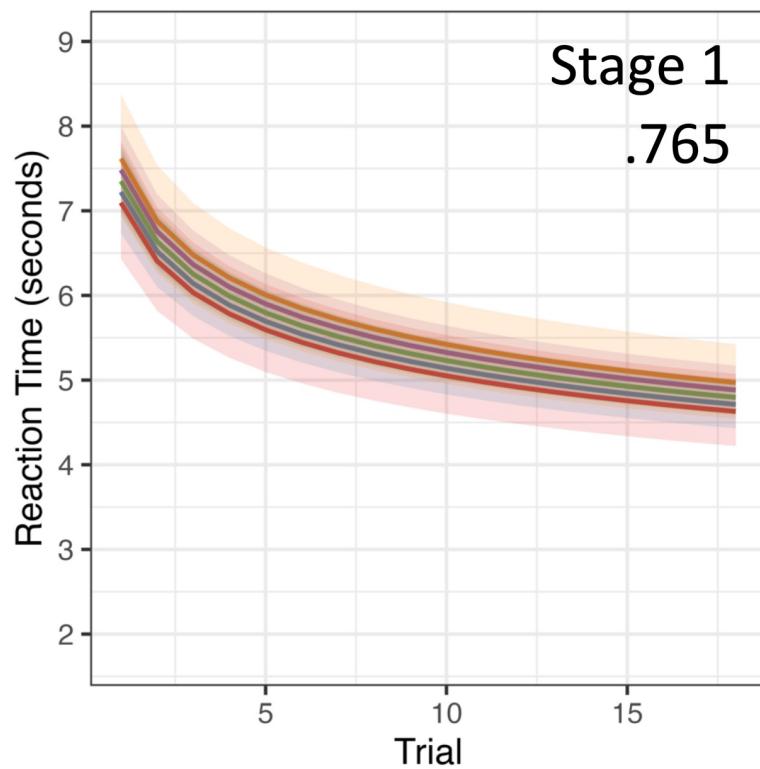
Regression

Accuracy: Declarative



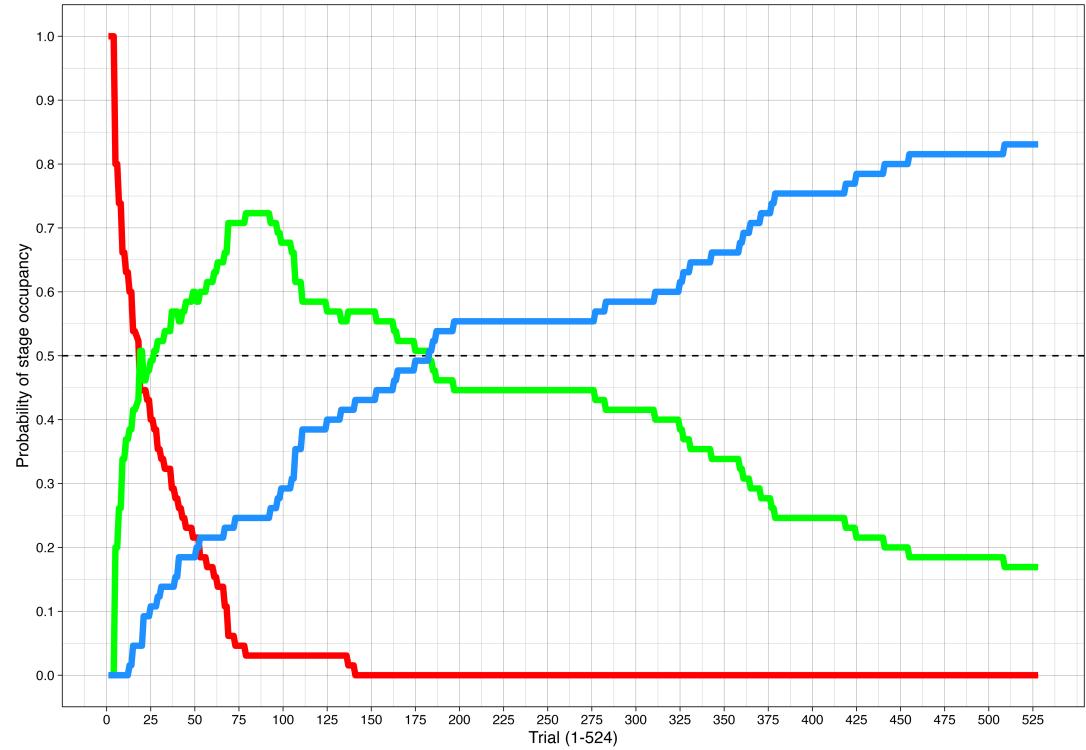
Consistent but decreasing role of declarative memory

Reaction Time: Procedural



Procedural memory only (appears to be) relevant in Stage 2

Discussion: Evidence



L2 skill acquisition in three stages

- 18 trials for Stage 2
- 182 trials for Stage 3

For intentional learning of grammar

- ACT-R seems to be the best account

Discussion: Implication



➤ The Three-stage model

If explicitly learned, L2 skill acquisition follows a **three-stage process** with the assumed **declarative-procedural** transition

➤ Trials to proceduralization / automatization

Our participants required:

- 18 trials for **proceduralization** (DeKeyser, 1997 claimed 16 – good call!)
- 182 trials for performance to become **automatic**

➤ Individual differences

- Stage 1 – **declarative** memory is critical
- Stage 2 – **procedural** memory may be critical

} Stage 3 = very few individual differences
→ Extensive practice removes IDs!?

Discussion: Limitations and Future Directions

➤ Deductive vs. Inductive learning conditions

- Comparing L2 skill acquisition under deductive vs. inductive conditions
 - Currently research in progress (supported by National Science Foundation)

➤ Neurocognitive approach

- Using fMRI for direct evidence of learning mechanisms
 - Currently research in progress as Maie, Jeong, & Uchihara (in progress)

➤ Updating and refining skill acquisition theory

- ACT-theory (often cited in SLA) was developed in 1980s (40 years ago!!)
- SAT needs more sophistication esp. from more L2 perspectives

Thank you!

Contact

Ryo Maie (The University of Tokyo)

maie@g.ecc.u-tokyo.ac.jp

If you want to ask anonymous
questions ...

[https://docs.google.com/presentation/d/12MggvtnkhlI7lI8LcudC8qRzzmB7ig-Trz8Rm37P0Fc/edit?usp=sharing](https://docs.google.com/presentation/d/12MggvtnkhlI7lI8LcudC8qRzzmB7ig-Trz8Rm37P0Fc/edit?usp=ssharing)



Acknowledgements

Phillip Hamrick (Kent State University)

Caitlin Tenison (Educational Testing Service)

Surya Narayan (The University of Tokyo)

Funding



National
Science
Foundation



MICHIGAN STATE
UNIVERSITY

References

- Anderson, J. R. (1983). *The architecture of cognition*. Mahwah, NJ: Lawrence Erlbaum.
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press.
- DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19(2), 195–221.
- DeKeyser, R. M. (2017). Knowledge and skill in ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 15–32). New York, NY: Routledge.
- DeKeyser, R. M. (2020). Skill acquisition theory. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition. An introduction* (3rd ed., pp. 83–104). New York, NY: Routledge.
- DeKeyser, R. M., & Sokalski, K. J. (1996). The differential role of comprehension and production practice. *Language Learning*, 46(4), 613–642.
- Ferman, S., Olshtain, E., Schechtman, E., & Karni, A. (2009). The acquisition of a linguistic skill by adults: Procedural and declarative memory interact in the learning of an artificial morphological rule. *Journal of Neurolinguistics*, 22(4), 382–412.
- Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 243–285). Cambridge, MA: Academic Press.
- Hamrick, P., Lum, J. A. G., & Ullman, M. T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences*, 115(7), 1487–1492.
- Hui, B. (2020). Processing variability in intentional and incidental word learning: An extension of Solovyeva and DeKeyser (2018). *Studies in Second Language Acquisition*, 42(2), 327–357.

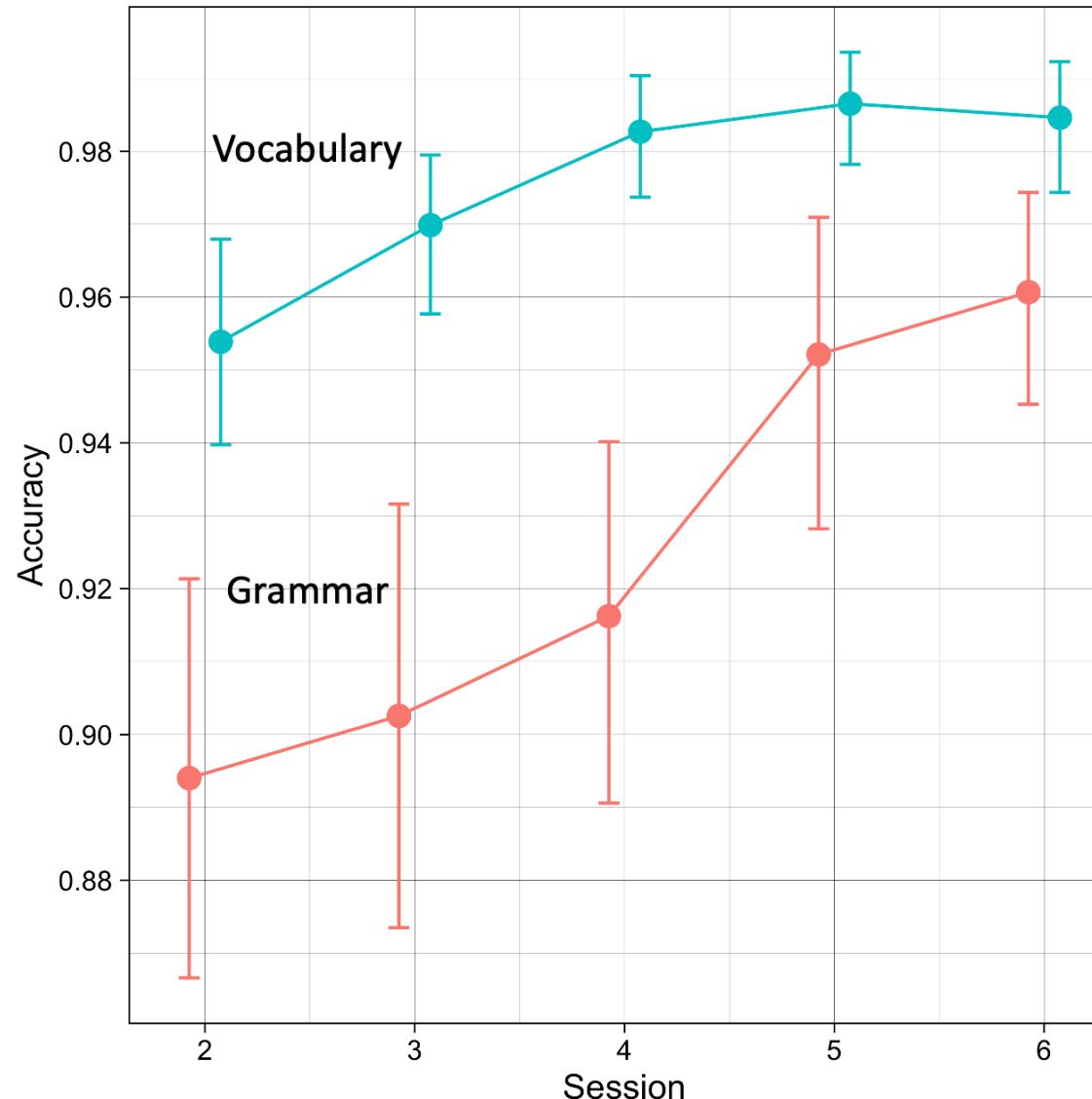
References

- Logan, G. D. (1988). Towards an instance theory of automatization. *Psychological Review*, 95(4), 492–527.
- Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review*, 109(2), 376–400.
- Maie, R. (2020). *Testing skill acquisition stages in language learning: A case of L2 vocabulary learning and practice*. Unpublished qualifying paper. Michigan State University, East Lansing, MI.
- Mueller, J. L. (2006). L2 in a nutshell: The investigation of second language processing in the miniature language model. *Language Learning*, 56(s1), 235–270.
- Oliveira, C. M., Henderson, L. M., Hayiou-Thomas, M. E. (2023). Limited evidence of an association between language, literacy, and procedural learning in typical and atypical development: A meta-analysis. *Cognitive Science*, 47(7), 1–34.
- Pili-Moss, D., Brill-Schuetz, K., Faretta-Stutenberg, M., & Morgan-Short, K. (2020). Contributions of declarative and procedural memory to accuracy and automatization during second language practice. *Bilingualism: Language and Cognition*, 23, 639–651.
- Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, 126(3), 288–311.
- Rickard, T. C. (2004). Strategy execution in cognitive skill learning: An item-level test of candidate models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), 65–82.
- Suzuki, Y. (2018). The role of procedural learning ability in automatization of L2 morphology under different learning schedules: An exploratory study. *Studies in Second Language Acquisition*, 40(4), 923–937.
- Suzuki, Y. (2022). Automatization and practice. In A. Godfroid & H. Hopp (Eds.), *The Routledge handbook of second language acquisition and psycholinguistics* (pp. 308-321). New York, NY: Routledge.

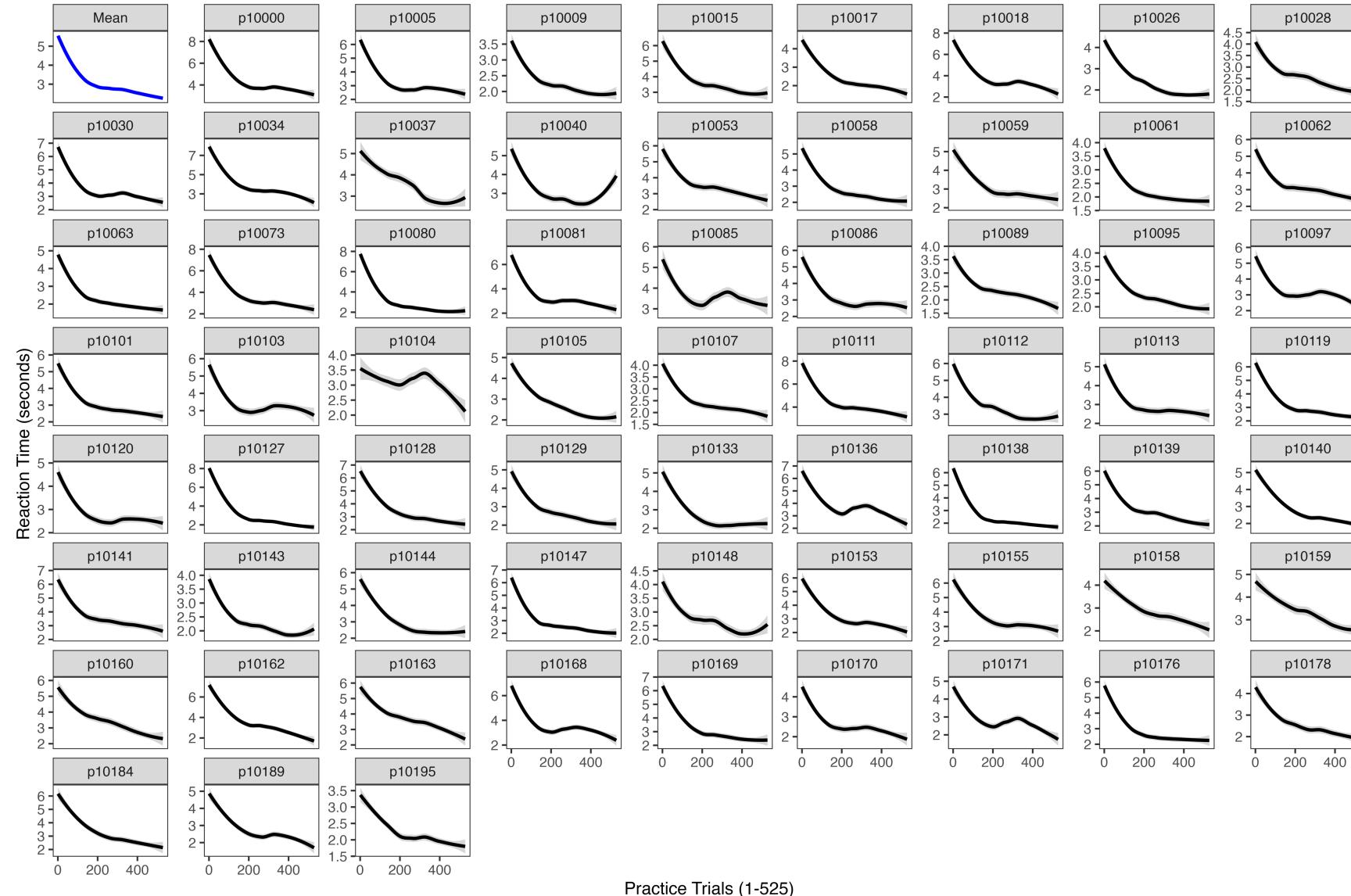
References

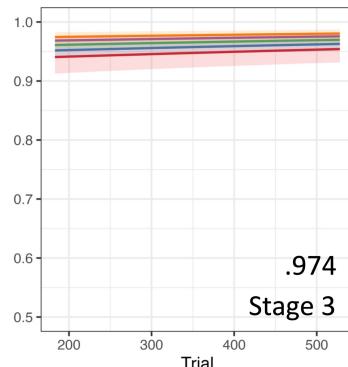
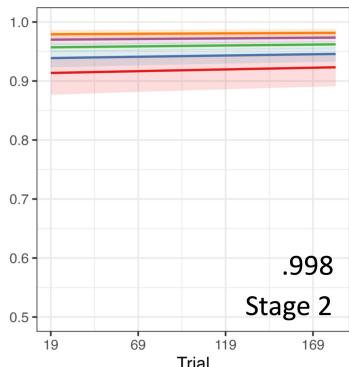
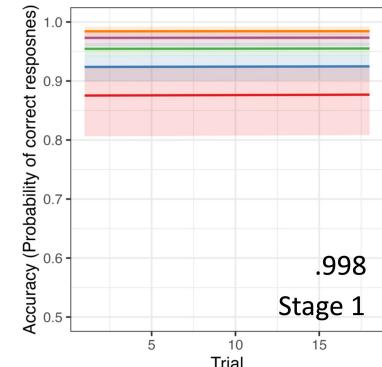
- Suzuki, Y., & Sunada, M. (2018). Automatization in second language sentence processing: Relationship between elicited imitation and maze tasks. *Bilingualism: Language and Cognition*, 21(1), 32–46.
- Tenison, C., & Anderson, J. R. (2016). Modeling the distinct phases of skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(5), 749–767.
- Tenison, C., Fincham, J. M., & Anderson, J. R. (2016). Phases of learning: How skill acquisition impacts cognitive processing. *Cognitive Psychology*, 87, 1–28.
- Ullman M. T. (2004) Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1–2), 231–270.
- Ullman, M. T. (2016). The declarative/procedural model: A neurobiological model of languagelearning, knowledge, and use. In G. Hickok & S. Small (Eds.), *Neurobiology of language* (pp. 953–968). Cambridge, MA: Academic Press.
- Ullman M. T. (2020) The declarative/procedural model: A neurobiologically-motivated theory of first and second language. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition. An introduction* (3rd ed., pp. 83–104). New York, NY: Routledge.

Vocabulary and Grammar Tests



Trial by trial RT - individual

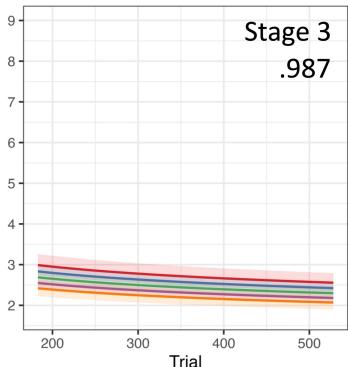
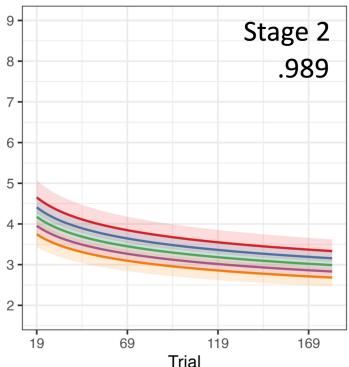
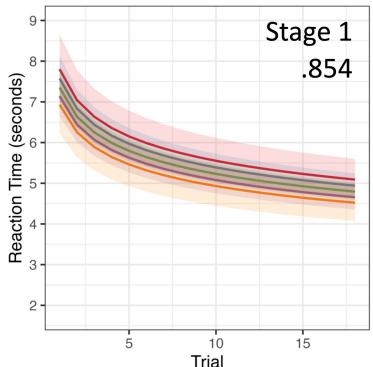




Declarative [-2] [-1] [0] [1] [2]

Declarative [-2] [-1] [0] [1] [2]

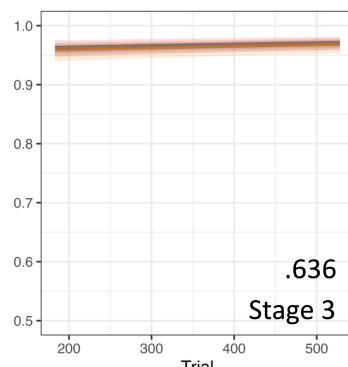
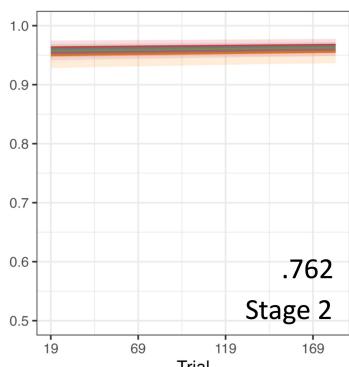
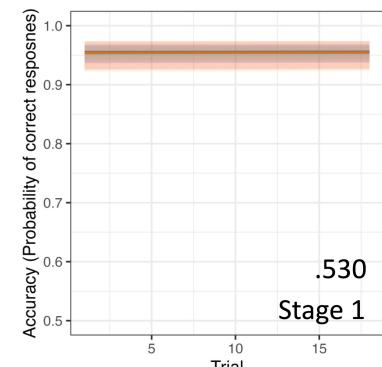
Declarative [-2] [-1] [0] [1] [2]



Declarative [-2] [-1] [0] [1] [2]

Declarative [-2] [-1] [0] [1] [2]

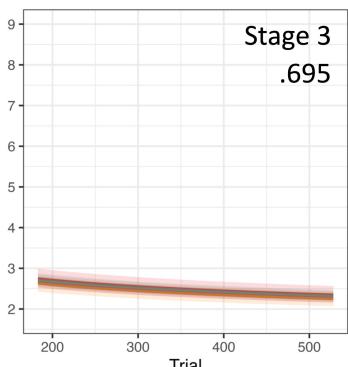
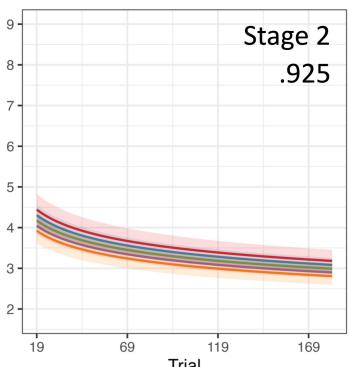
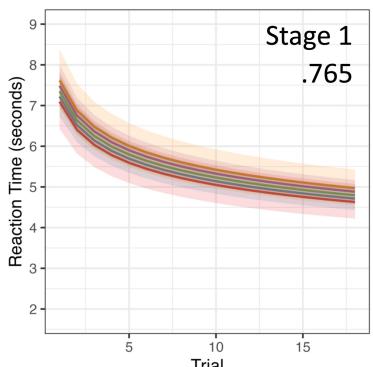
Declarative [-2] [-1] [0] [1] [2]



zASRT15 [-2] [-1] [0] [1] [2]

zASRT15 [-2] [-1] [0] [1] [2]

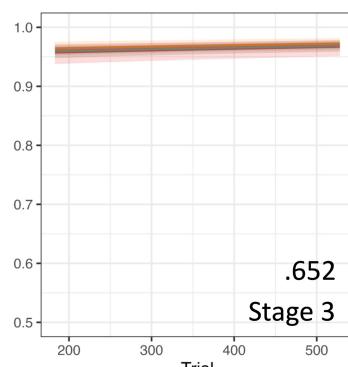
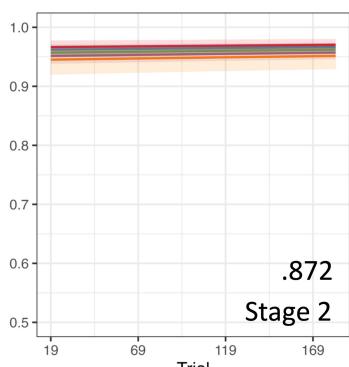
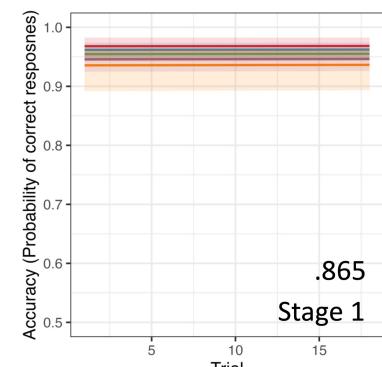
zASRT15 [-2] [-1] [0] [1] [2]



zASRT15 [-2] [-1] [0] [1] [2]

zASRT15 [-2] [-1] [0] [1] [2]

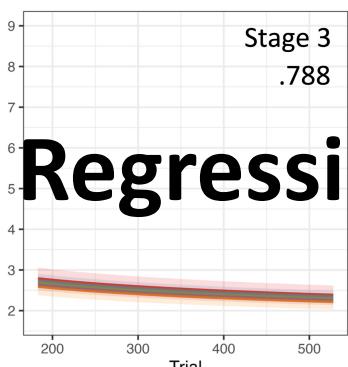
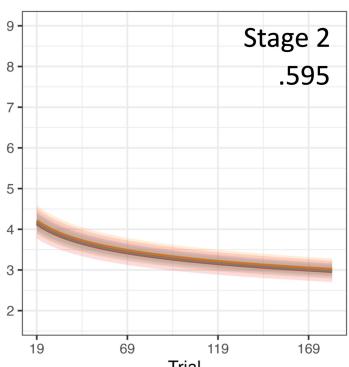
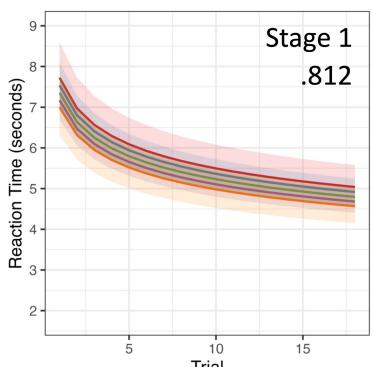
zASRT15 [-2] [-1] [0] [1] [2]



zSL [-2] [-1] [0] [1] [2]

zSL [-2] [-1] [0] [1] [2]

zSL [-2] [-1] [0] [1] [2]



zSL [-2] [-1] [0] [1] [2]

zSL [-2] [-1] [0] [1] [2]

Regression

Correlation Matrix

