# Re-examine the reliability of statistical learning tasks across domains and modalities

An Nguyen[1,2], Yoel Sanchez Araujo[3], Wendy Georgan[4] Joanne Arciuli[5], Zhenghan Qi[1,3]

[1] University of Delaware; [2] Johns Hopkins University; [3] Massachusetts Institute of Technology, [4] Harvard University, [5] University of Sydney

## Introduction

- Statistical learning (SL), the ability to detect and extract regularities from inputs, plays a key role in spoken language development (Saffran et al., 1996; Saffran, 2003) and is traditionally thought of as a unified capacity governed by the procedural memory system.
- However, recent studies have shown that adults' capabilities to learn temporal sequences vary across different stimulus types.
- The emergence of these individual difference studies raised questions regarding task reliabilities (Siegelman and Frost, 2015; Erickson et al., 2015; Siegelman et al., 2018).

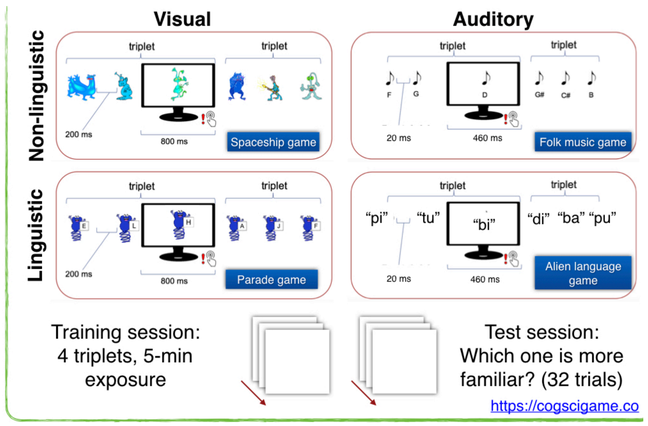
## Objectives

This study assessed:

- the relationship between online and offline SL measurements
- the reliability of SL tasks
- the relationship between SL performance across modalities and domains.

## Materials and Methods

145 native English speakers between 18-35 (mean age: 29.6; F = 69; M = 76 ) participated in the experiment. 79 came back for the re-test session two months later.



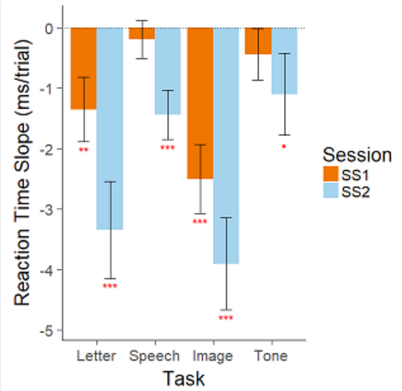

Familiarization phase:

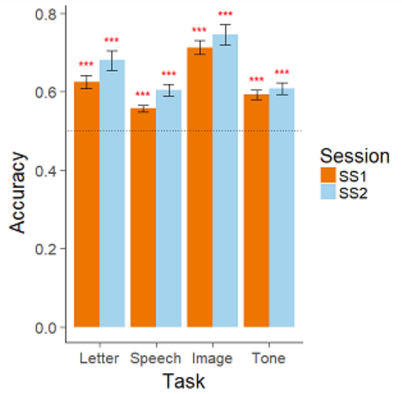
- Target detection cover task
- Each triplet is repeated 24 times in the visual tasks (SOA = 1000 ms) and 48 times in the auditory tasks (SOA = 480 ms.

Testing phase:

- Two-alternative forced choice task
- 4 foils for each triple
- 32 two-alternative forced choice test trials

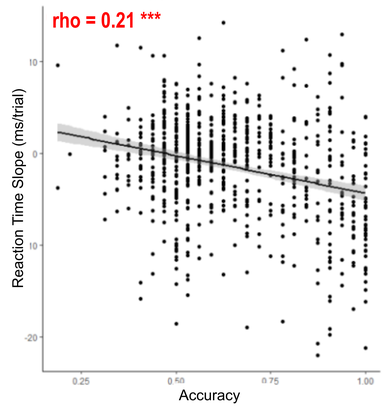
## Online-Offline Measurements



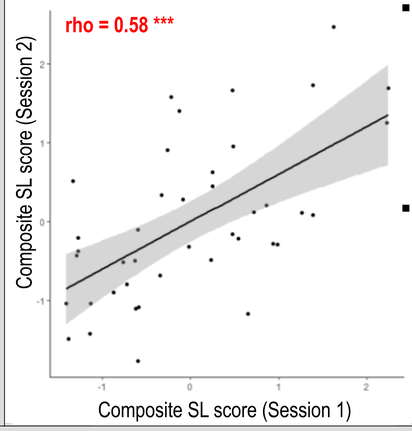

*Note: * p < .05; ** p < .01;*** p < .001.*

- Significant offline learning in all tasks (accuracy > 0.5, p's < .001)
- Offline performance in SS2 (M = 0.65) is slightly better than SS1 (M=0.62) (p = .02)
- The more rapid acceleration of RT was significantly associated with higher test accuracy (rho = 0.21, p < .001)




## Task Reliability

| | Letter | | Speech | | Tone | | Image | |
|---|---|---|---|---|---|---|---|---|
| **Internal Consistency (session 1 \| session 2)** | 0.84 | 0.84 | 0.63 | 0.68 | 0.68 | 0.81 | 0.89 | 0.9 |
| **Test-retest Reliability (Spearman Correlation)** | rho = 0.01 | | rho = 0.37 ** | | rho = 0.14 | | rho= 0.56 *** | |




- The test-retest reliabilities were low to moderate for individual SL tasks, but although reliability was high for the Image task (rho=.56). The highest test-retest reliability was obtained using the composite SL score (rho=.58, p < .001)
- Composite SL scores were the loading scores of the first principle component, explaining 45% of variance in Session 1 and 55% of variance in Session 2, representing the domain- and modality-general SL ability.

## Modalities and Domains

| Session 1 | | | |
|---|---|---|---|
| | Tone (N =131) | Letter (N=144) | Image (N=145) |
| Speech (N = 138) | **0.32** *** | -0.09 | **0.27** ** |
| Tone (N =131) | | 0.07 | **0.41** *** |
| Letter (N=144) | | | 0.07 |

| Session 2 | | | |
|---|---|---|---|
| | Tone (N=79) | Letter (N=75) | Image (N=78) |
| Speech (N=76) | 0.01 | **0.35** ** | **0.35** ** |
| Tone (N=79) | | 0.06 | 0.31 * |
| Letter(N=75) | | | **0.45** *** |

*Bold values indicate p < .008 (Bonferroni corrected α value)*

- Correlation between some tasks indicates that they share an underlying domain-general computational mechanism.
- However, the modest correlation also suggests individuals' statistical learning behavior can be partially constrained by domain- and modality-specific mechanisms.

## Conclusion

1.The online measure (RT slope) of statistical learning is mildly associated with the offline learning outcomes, suggesting the two measures represent related but separate aspects of SL processes.

2. A composite measurement of SL behavior is more reliable than any single measure of SL behavior, suggesting the domain- and modality-general computation mechanism is stable within an individual.

3. The modest correlation between SL tasks indicates learning outcomes are governed by both domain-general and modality-specific mechanisms.

## References

1.Conway, C.M., Christiansen, M.H. (2009). Seeing and hearing in space and time: effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology.*

2.Erickson, L.C., Thiessen, E.D. (2015). Statistical learning of language: theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review.*

3.Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science.*

4.Saffran, J. (2003). Statistical Language Learning: Mechanisms and Constraints. *Current Directions in Psychological Science.*

5.Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language.*

*6.* Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition.*

Contact: anguye60@jhu.edu ; zqi@udel.edu