# The VST and other measures of L2 vocabulary knowledge

**Irina Elgort**
**Victoria University of Wellington**
irina.elgort@vuw.ac.nz

## Test validity measures

- Validity is not the property of the test or assessment as such, but rather of the ***meaning of the test scores***. Hence, what is to be validated is … the ***inferences*** derived from test scores …

  *Messick (1996)*

- The original (14k) VST measures how many word-families English language learners know (i.e., an aspect of L2 lexical proficiency / development).

➢ VST scores should have various degrees of correlation with other measures of lexical development.

## What is being measured?

- The VST is designed to measure learners' **written receptive vocabulary size** for the first 14 1000-word families of English.
- RASCH validations of the VST with Japanese (Beglar, 2010) and Russian (Elgort, 2013) learners of English show:
  - a very high degree of **reliability, technical quality** and
  - **psychometric unidimensionality =>**
    "the VST primarily measures a **single latent variable**, which is ***presumably*** written receptive vocabulary knowledge"
    (Beglar, 2010: 112)

- **The VST <u>does not measure</u>:**
  - Productive vocabulary size
  - Spoken vocabulary
  - Quality (depth) of vocabulary knowledge
  - Fluency of access to vocabulary knowledge

## Predictions

Behavioural measures of lexical knowledge (LDT)

1. Accuracy of responses (d-prime, to account for bias) * * *
2. Response latencies (RT) to L2 words * * *
3. Fluency (automaticity) of lexical processing $CV_{RT}$ (SD/RT) * * *

Reading comprehension * * *

New word learning * * *

## Lexical decision task

You will briefly see letters in the middle of the screen.
If it is a word, say “yes”, if it is not a word, say “no”.
Try to respond as quickly & as accurately as you can.

**SUNCTION**

## Sources of data: L2 vocabulary studies

1. **Incidental vocabulary learning from reading a long connected text**
   - 48 adult L2 participants

| | Mean | SD | Range |
|---|---|---|---|
| Vocabulary size (VST) | 10135 | 1618 | 6900 – 13100 |

2. **Incidental vocabulary learning from sentence contexts**
   - 26 adult L2 participants

| | Mean | SD | Range |
|---|---|---|---|
| Vocabulary size (VST) | 7792 | 2451 | 3800 – 12100 |

3. **Deliberate vocabulary learning from word cards**
   - 41 adult L2 participants

| | Mean | SD | Range |
|---|---|---|---|
| Vocabulary size (VST) | 9444 | 1689 | 5100 – 13800 |

# Behavioural measures (L2 LDT)

**Pearson's product-moment correlations of VST and …**

- *Study 1 (pre-study LDT)*

| | r | *t*-value | df | p-value |
|---|---|---|---|---|
| D-prime | **0.62** | 5.36 | 46 | p<.001 *** |
| RT | **-0.50** | -3.94 | 46 | p<.001 *** |
| CV | **-0.39** | -2.83 | 46 | p<.01 ** |

- *Study 2 (pre-study LDT)*

| | r | *t*-value | df | p-value |
|---|---|---|---|---|
| D-prime | **0.77** | 5.66 | 22 | p<.001 *** |
| RT | **-0.58** | -3.31 | 22 | p<.01 ** |
| CV | **-0.42** | -2.18 | 22 | p<.05 * |

- *Study 3 (pre-study LDT)*

| | r | *t*-value | df | p-value |
|---|---|---|---|---|
| D-prime | **0.64** | 5.16 | 39 | p<.001 *** |
| RT | **-0.21** | -1.32 | 39 | p=.193 |
| CV | **-0.28** | -1.80 | 39 | p=.080 . |

# L2 Reading

**Pearson's product-moment correlations:**

- *Study 1: long connected text – deep understanding / interpretation*

| | r | *t*-value | df | p-value |
|---|---|---|---|---|
| Reading comp. scores (Day & Park, 2005): reorganizational understanding | **0.60** | 5.11 | 46 | p<.001 *** |

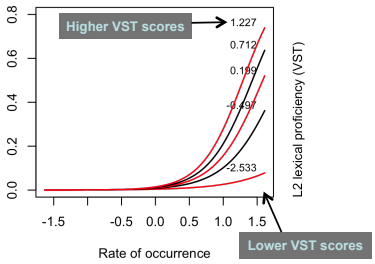
- *Study 2: sentence comprehension – true-false format*

| | r | *t*-value | df | p-value |
|---|---|---|---|---|
| Self-rating of reading comp. (subjective) | 0.44 | 2.33 | 22 | p<.05 * |
| ACC of responses to TF comp. questions (objective) | 0.46 | 2.46 | 22 | p<.05 * |
| Mean time on task (comp. q-ns) | **-0.60** | -3.55 | 22 | p<.01 ** |

# New word learning: Study 1

**Incidental learning from reading a long connected text:**
**Knowledge of meaning (meaning generation task:** What does ***afuse*** mean?**)**

| Predictors | Coef.*β* | SE(*β*) | *z* | *p* | |
|---|---|---|---|---|---|
| WordCV | -0.865 | 0.243 | -3.57 | 3.6E-04 | *** |
| VST | 0.345 | 0.265 | 1.30 | 0.194 | |
| **Interactions** | | | | | |
| No.Occur:VST | 0.365 | 0.166 | 2.21 | 0.027 | * |

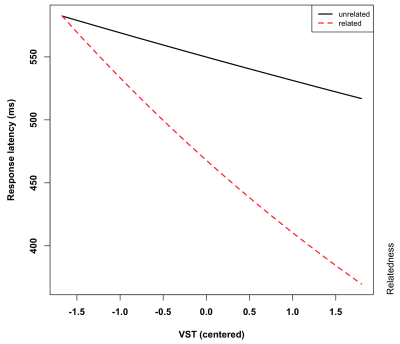



# New word learning: Study 2

**Incidental learning from sentence context:**
**Knowledge of meaning (semantic relatedness task):**
***commondious – spacious* vs. *commondious – fabulous***

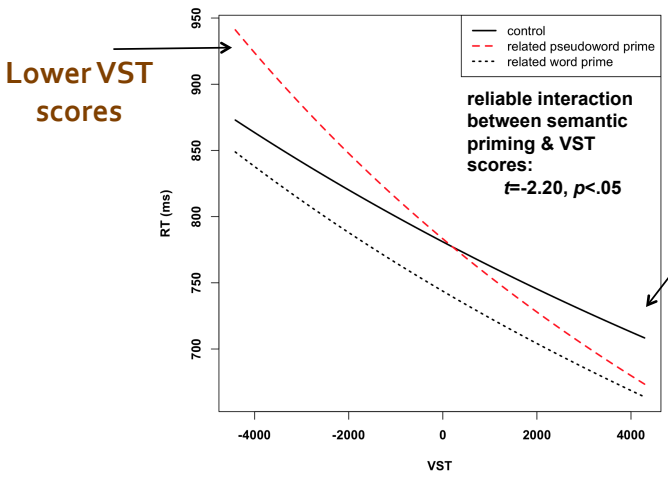
**RT analysis - LME**

| Predictors | Coef.*β* | MCMC mean | *HPD95 lower* | *HPD95 upper* | *p* | |
|---|---|---|---|---|---|---|
| Related:VST | -0.0967 | -0.0979 | -0.1909 | 3.80e-13 | 0.0494 | * |



# New word learning: Study 3

**Deliberate vocabulary learning (flashcards)**
**Knowledge of meaning (semantic priming task):**
***surmit*** *(meaning bulldozer)* ***- excavator***




# Conclusions

"the VST primarily measures a single latent variable, which is ***presumably*** written receptive vocabulary knowledge" (Beglar, 2010: 112)

- VST scores correlate as expected with behavioural measures of L2 lexical development;
- VST scores correlate as expected with L2 reading comprehension;
- The VST predicts new L2 vocabulary learning, in line with the Matthew effect observed in L1.

irina.elgort@vuw.ac.nz