

**SIGN LEARNING AND ITS RELATIONSHIP TO WORD
LEARNING IN HEARING ADULTS**

A Dissertation
Presented to
The Academic Faculty

by

David Martinez

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Psychology

Georgia Institute of Technology
December 2018

COPYRIGHT © 2018 BY DAVID MARTINEZ

**SIGN LEARNING AND ITS RELATIONSHIP TO WORD
LEARNING IN HEARING ADULTS**

Approved by:

Dr. Jenny L. Singleton, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Daniel H. Spieler
School of Psychology
Georgia Institute of Technology

Dr. Susan E. Embretson
School of Psychology
Georgia Institute of Technology

Dr. Michael F. Bunting
Center for Advanced Study of
Language
University of Maryland

Dr. Randall W. Engle
School of Psychology
Georgia Institute of Technology

Date Approved: August 15, 2018

To Fortune.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Jenny Singleton, for providing me with the opportunity and guidance to complete my PhD. I would also like to thank my committee members who provided valuable feedback and resources and the undergraduate research assistants who helped complete this study. Finally, I would like to acknowledge the monetary support received from the National Science Foundation, Doctoral Dissertation Research Improvement Grant, award number 165115.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vi
LIST OF FIGURES	vii
SUMMARY	viii
CHAPTER 1. Introduction	1
1.1 Phonological Short-Term Memory	3
1.2 Crystallized and Fluid Intelligence	4
1.3 Working Memory Capacity	7
1.4 The Present Study	9
1.4.1 Predictions	10
CHAPTER 2. Methods	11
2.1 Participants	11
2.2 Procedure	12
2.3 Tasks	14
2.3.1 Lexical Learning	14
2.3.2 Phonological Short-Term Memory	21
2.3.3 Intelligence	29
2.3.4 Working Memory Capacity	33
2.3.5 Summary of Tasks	35
2.4 Statistical Analyses	36
2.4.1 Data Screening	36
2.4.2 Statistical Procedure	36
CHAPTER 3. Results	38
3.1 Observed Variable Analyses: Descriptive Statistics, Reliability, and Correlations	38
3.2 Latent Variable Analyses	40
3.2.1 Predicting Sign and Word Learning	43
3.2.2 Modelling the Relationship Between Sign and Word Learning	47
CHAPTER 4. Discussion	50
4.1 Crystallized and Fluid Intelligence and Relation-Construction	50
4.2 Working Memory Capacity and Phonological Short-Term Memory	53
4.3 Phonological Short-Term Memory: The Specific and the General	54
4.4 Lexical Learning: The General and the Partially Specific	57
4.5 Summary and Implications	59
APPENDIX A. List of English Words used in this study	61
REFERENCES	62

LIST OF TABLES

Table 1	Task administration order	13
Table 2	Summary of tasks	35
Table 3	Descriptive statistics and reliabilities	39
Table 4	Zero-order correlations	40
Table 5	Correlated factors model fit statistics	41
Table 6	Latent variable correlations	42
Table 7	Correlations with Gf and WMCcs	42
Table 8	Fit statistics for predictive model	44
Table 9	Fit statistics for exploratory models	48

LIST OF FIGURES

Figure 1	Depiction of the ASL-SL Task. Panel A depicts a study trial. Panel B depicts a test trial. In both cases, the sign is shown first and is immediately followed by either the response word in the study phase or the response screen in the test phase.	16
Figure 2	Depiction of the PSL task. Panel A shows a study trial. Panel B shows a test trial.	18
Figure 3	Depiction of the 3TSL. Panel A depicts part of a study trial. Panel B depicts part of a test trial.	20
Figure 4	Depiction of the NSPT. In the NSPT, there were 28 target signs each with two reproductions, produced by different individuals	24
Figure 5	Example of two pseudosigns differing in orientation only. Note: this item is from the ProSign. It was chosen because it is amenable to a simple label.	25
Figure 6	Depiction of a ProSign item. The probe differs from the second pseudosign in the set.	26
Figure 7	Depiction of the SignCon. Participants were either tested on the number of pseudosigns (60% of trials) or letters (40% of trials) they could recall in order.	28
Figure 8	Depiction of a SLAT item. The correct answer choice is 1.	33
Figure 9	Latent variable and their indicators. Panel A shows the estimated path coefficients derived from analyzing Model Corr-LM. For reference, panel B shows the unresidualized Gf factor and its indicators as well as a WMC factor derived from the variance of complex span tasks only.	41
Figure 10	Model 1—the effect of intelligence	44
Figure 11	Model 2—accounting for WMC	45
Figure 12	Model 3—Accounting for spoken-PSTM	46
Figure 13	Model 4—Accounting for signed-PSTM	47
Figure 14	Model explaining the relationship between word learning and sign learning. Note: SL = sign learning, LL = lexical learning	49

SUMMARY

Typically, in associative lexical learning, familiar information is studied alongside unfamiliar information. For example, one may study a word and its definition (written in accessible language, e.g., *prosopagnosia*—face blindness). Although this is not the way we acquire most of the vocabulary in our lexicon, it does have its place, allowing individuals to select words for repeated study and to begin learning words in an unfamiliar language. Moreover, research has found that learning words in this fashion is related to more general language measures. Critically, however, the vast majority of this research has been conducted with spoken languages—we know very little about the cognitive processes involved in sign learning or whether sign learning is related to word learning. The present study was conducted to address this gap.

In this study, structural equation modeling was used to 1) extend individual differences research on second language word learning to sign learning in hearing non-signers and 2) to model the relationship between word and sign learning. Two-hundred thirty-six participants completed 25 tasks assessing word learning, sign learning, language modality specific phonological short-term memory, fluid intelligence, crystallized intelligence, and working memory capacity.

The results of this study indicated that fluid intelligence was predictive of both word and sign learning, however, after accounting for other variables, phonological short-term memory was only predictive of lexical learning within modality (e.g., short-term memory for signs was predictive of sign learning but not word learning). It was also observed that sign and word learning were strongly correlated. Exploratory analyses

revealed that all tasks loaded onto a general lexical learning factor but sign learning tasks additionally loaded onto a specific factor. As such, this study provides insight into the cognitive components that are common to lexical learning regardless of language modality and those that are unique to either signed or spoken languages.

CHAPTER 1. INTRODUCTION

We have all engaged in associative lexical learning at one point in our lives or another—perhaps while studying vocabulary for a test (e.g., *prosopagnosia*—“face blindness”) or to prepare for a trip to a country where a foreign language is spoken (e.g., *donde esta la biblioteca?*—“where is the library?”). Learning in this *decontextualized* manner is not how we have developed the bulk of our lexicon (Hulstijn, 2003; Krashen, 1989; Nation, 1980), but it does have its place. Associative lexical learning allows one to select items and study them repeatedly, leading to long term retention (Rohrer, Taylor, Pashler, Wixted, & Cepeda, 2005; Seibert, 1930; Thorndike, 1908) and fluent use (Elgort, 2011; Yang, 1997). It also allows individuals to begin learning verbal material in a second language (L2) before having mastered the phonology or grammar of the target language. Moreover, in the lab, associative lexical learning ability has been found to correlate moderately to strongly with other linguistic variables such as grammar learning (Cooper, 1964; Gardner & Lambert, 1965; K. I. Martin & Ellis, 2012; O'Brien, Segalowitz, Collentine, & Freed, 2006), L2 learning aptitude (Cooper, 1964; Li, 2015), and more generally with verbal ability (Hundal & Horn, 1977).

Given its place in our lives and its relation to other linguistic abilities, it is no wonder that researchers have been interested in individual differences in lexical learning, or more precisely, in *word* learning (e.g., Hundal & Horn, 1977; Kyllonen, Tirre, & Christal, 1991; Underwood, Boruch, & Malmi, 1978)—there have been few studies investigating individual differences in *sign* learning and, as a consequence, we know very little about the cognitive factors engaged while learning signs and whether they are

similar and relied upon to the same degree as those deployed during word learning. For example, crystallized intelligence (Gc), i.e., one's stock of knowledge, is a general cognitive factor implicated in word learning and the acquisition of other forms of information (Ackerman & Cianciolo, 2000; Hambrick, 2003; Kyllonen & Woltz, 1989), however, it is typically measured via linguistic tasks, e.g., vocabulary. In hearing individuals, the kind of knowledge assessed by typical Gc measures has been accumulated through *spoken* language, therefore it may be that Gc and word learning have more in common than Gc and sign learning, leading Gc to be more strongly related to word learning than to sign learning.

The present study was conducted to extend individual differences research from L2 word learning to sign learning and to examine the relationship between the two constructs. A prerequisite is the identification of a set of factors that are likely to be predictive of lexical learning in one or both language modalities. To that end, a literature review was conducted and a parsimonious account of the components involved in effective lexical learning was formulated. Effective lexical learning relies on encoding and maintaining verbal material, generating relationships to aid as cues¹, and maintaining target verbal material and relationships in the face of interference. As will be reviewed below, these components implicate phonological short-term memory, complementary action by fluid and crystallized intelligence, and working memory capacity.

¹ Rote rehearsal can be used but it is not the most effective means of learning material (Bradley & Glenberg, 1983; Nairne, 1983).

1.1 Phonological Short-Term Memory

Phonological short-term memory (PSTM) refers to the ability to encode verbal information and retain it in some form for a brief period of time. In hearing non-signers, it is often assessed via span tasks, in which individuals are asked to remember sets of verbal items (words, digits, pseudowords) and recall them in the order they were presented. A large body of research has found that *spoken*-PSTM tasks, utilizing spoken language material, are moderately to strongly correlated to word learning (Gupta, 2003; Hummel & French, 2016; K. I. Martin & Ellis, 2012; O'Brien et al., 2006; O'Brien, Segalowitz, Freed, & Collentine, 2007) and at least one study has found that *signed*-PSTM tasks, utilizing signed language material, are related to sign learning (Martinez & Singleton, 2018).

Given the scarcity of research investigating signed-PSTM and sign learning, it is worth noting that signed-PSTM tasks are related to other language outcomes. In Deaf children, signed-PSTM tasks have been used to discriminate between children with and without specific language impairment (Marshall et al., 2015; Mason et al., 2010). Additionally, studies investigating the role of signed-PSTM in hearing sign language interpreters have reported positive correlations between signed-PSTM and sign language ability (Gómez, Molina, Benítez, & de Torres, 2007; Shaw, 2011). These relationships are analogous to those observed in spoken language research (Daneman & Merikle, 1996; Gathercole & Baddeley, 1990).

Whether signed-PSTM and spoken-PSTM are best considered facets of a single construct or are entirely unrelated is still an empirical question. Gathercole (2006)

theorized that the relationship between lexical learning and PSTM is due to common phonological, perceptual, and motor processes. In theory, phonological processing is amodal, as the information being processed are abstract linguistic units (Baddeley, 2015; Baddeley, Gathercole, & Papagno, 1998; for a counter argument, see Jones, Hughes, & Macken, 2006). In fact, evidence from neuroimaging studies have shown that the same classic language areas that are activated by spoken language processing are active during signed language processing (Bavelier et al., 1998; Söderfeldt et al., 1997; J. T. Williams, Darcy, & Newman, 2015). There is evidence, however, that hearing non-signers do not immediately process signs linguistically, instead processing them as nonverbal movements (Martinez & Singleton, 2018; Newman-Norlund, Frey, Petitto, & Grafton, 2006; J. T. Williams, Darcy, & Newman, 2016b). The other two common processing components, perceptual and motor processes, are undoubtedly different across signed and spoken languages: signed languages are visuospatial languages articulated by the hands, face, and body; spoken languages on the other hand are aural-oral. The lack of phonological processing in non-signers learning signs along with differences in the perceptual and motor processes recruited to perceive and produce languages across modalities implies that in hearing non-signers, PSTM for signed language relies on processes that are at least partially distinct from those utilized to encode and maintain spoken language.

1.2 Crystallized and Fluid Intelligence

According to the relation-construction principle, “the strength of a bond between a pair of items (which governs the success of retrieval of that pair) is determined by the *quantity* and *quality* of the relations constructed between the items during study

(Kyllonen et al., 1991, p. 58).” The greater the number of relations formed—or the more *elaborative*—the greater the number of cues that can be used to retrieve the appropriate response. Of course, these relations are of little use if they do not uniquely index the items under study or if they have weak association values and are therefore unlikely to elicit the appropriate responses (Glaze, 1928; Jenkins, 1985; Noble, 1952). Thus, quality and quantity matter.

In terms of the simple task analysis presented above, crystallized intelligence (Gc) and fluid intelligence (Gf) are implicated in the construction of relationships. Gc refers to acquired knowledge and skills (Cattell, 1943)—it provides the “network of facts and associations into which new facts and associations might be interwoven (Kyllonen & Woltz, 1989, p. 246)”, or in other words, the material with which to construct relations. Gf refers to the ability to solve novel problems and reason in novel situations (Cattell, 1943). According to Shipstead, Harrison, and Engle (2016), Gf tasks place a premium on the ability to *disengage* from outdated information. When inducing a relationship between familiar and unfamiliar information, an individual must consider possible relations and be able to abandon those that are inadequate, lest they block one from constructing a more appropriate relation. Indeed, both Gc and Gf generally show moderate relationships with lexical learning, though Gc often shows a stronger relationship with lexical learning than Gf (e.g., Hundal & Horn, 1977; Kyllonen & Tirre, 1988).

To illustrate the impact of Gc and Gf, suppose one was studying a list of words and their meanings and one of the items was *gloaming-twilight*. One may note that *gloaming* and *twilight* both have 8 letters but this is not unique to this pair of words; this relation is

then abandoned in favor of one that relates *gloaming* and *twilight* via “glow,” which sounds similar to *gloaming* and relates to the level of light present at *twilight*. Assuming no other words in the list relate to dim lighting and/or sound similar to glow, than relating *gloaming* and *twilight* via *glow* will likely result in correct recall. Note, neither the word *glow* nor the concept of luminosity were explicit, rather, this information was drawn from prior knowledge and a relationship was induced. If it so happens that *gloomy* is another term in the list, then it would behoove one to abandon the previous relation (further implicating Gf) as *gloomy*, *gloaming*, and *glow* share sound similarities and all relate to dim lighting conditions, resulting in increased interference amongst the terms, and consequently affecting the likelihood of correct recall.

Another germane point related to relation-construction is the knowledge and use of mnemonic strategies. One can memorize a list of words by rote-rehearsal, that is, simply repeating the items, but this is an ineffective strategy compared to elaborative mnemonic techniques such as generating vivid imagery or relating the learning material to oneself (Bower & Winzenz, 1970; Bradley & Glenberg, 1983; Symons & Johnson, 1997). Still, there are individual differences in the use of mnemonics, possibly due to lack of awareness, experience, or knowledge of the efficaciousness of such techniques (Hertzog, Price, & Dunlosky, 2012; Shaughnessy, 1981), implicating Gc. Moreover, an individual may vary strategy use within or across tasks. This alternation in strategy requires one to disengage from a previously used strategy, implicating Gf.

To my knowledge, no study has investigated the relationship between Gf and sign learning in hearing individuals acquiring a sign language and only one study has investigated Gc. J. T. Williams, Darcy, and Newman (2016a) administered an English

vocabulary test (among other measures) to 25 individuals enrolled in an American Sign Language course. The English vocabulary test, an indicator of G_c , was significantly related to sign learning; caution, however, must be taken given the small number of participants.

1.3 Working Memory Capacity

Working memory capacity (WMC) is defined and operationalized in a variety of ways (Cowan, 2008; Oberauer et al., 2018). Here, WMC is defined as a domain-general ability that allows individuals to maintain a limited amount of information in a highly accessible state, even in the face of interference (Engle, 2002; Shipstead et al., 2016); it is best assessed by tasks that require short-term memory *and* prevent or disrupt motor rehearsal such as speech-motor (i.e., articulatory; Baddeley, Thomson, & Buchanan, 1975) or gaze-based (Tremblay, Saint-Aubin, & Jalbert, 2006) rehearsal, forcing individuals to rely on the control of attention, or *executive attention*, to maintain durable representations (Cowan, 2008; see also La Pointe & Engle, 1990, p. 1130).

To be sure, WMC, as defined here, is similar to short-term memory (STM) and therefore PSTM—both WMC and STM are defined in part by the ability to maintain information in memory for a brief period of time. In fact, modeling studies investigating the relationship between WMC and STM have observed correlations approaching unity (e.g., Colom, Shih, Flores-Mendoza, & Quiroga, 2006), however, researchers generally find correlations equal to or less than .80 (Cowan, 2008; Engle, Tuholski, Laughlin, & Conway, 1999; Kane et al., 2004; Unsworth & Engle, 2007).

The relationship between WMC and STM is at least partly due to the fact that both are supported by executive attention (Kane et al., 2004; Unsworth, 2010). The two are distinguished, however, by the fact that WMC depends on executive attention to a greater degree than STM and STM tends to depend on domain-specific processes to a greater degree than WMC (Kane et al., 2004). The distinction between WMC and STM is further supported by studies reporting independent contributions from WMC and STM to language-based outcomes (e.g., Cantor, Engle, & Hamilton, 1991; K. I. Martin & Ellis, 2012; Verhagen & Leseman, 2016).

Notably, researchers have found that WMC is predictive of associative word learning (Kaufman, DeYoung, Gray, Brown, & Mackintosh, 2009; K. I. Martin & Ellis, 2012; Tamez, Myerson, & Hale, 2008), though, to my knowledge, no study has explored WMC as a predictor of sign learning. As with STM, the relationship between word learning and WMC is at least partly due to the control of attention. To elaborate, one needs to control attention to stay focused on the task at hand and avoid attending to irrelevant information from the environment, our own thoughts, or from within the task itself; when our attention is pulled to irrelevant stimuli, then the encoding of target material is negatively affected and interference increases (Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999). Poor encoding and increased interference leads to a lower probability of a correct response, be that in a memory task, learning task, or some other task. Individuals with high WMC, however, are better able to use executive attention to prevent encoding failures and the accumulation of interference (Kane & Engle, 2000), resulting in easily accessible and durable memory representations (Shipstead & Engle, 2013; Unsworth, Brewer, & Spillers, 2013; Unsworth, Spillers, & Brewer, 2012).

1.4 The Present Study

The preceding review has provided evidence indicating that lexical learning plays a role in language learning and is related to a number of other abilities, namely: PSTM, Gc, Gf, and WMC. The vast majority of support for these claims, however, has come from research on spoken language, leading one to question how these constructs relate to sign learning and if and how word learning is related to sign learning.

The present study had two aims. The first aim was to extend the research on second language word learning to the sign domain. In order to accomplish the first aim, sign learning, signed-PSTM, Gc, Gf, and WMC were explored in hearing non-signers. The second aim was to investigate the relationship between sign learning and word learning. Accomplishing this aim required evaluating all of the previously mentioned constructs with the addition of word learning and spoken-PSTM. Including all of these constructs in the exploration of the second aim allowed for the control of potential mediating variables.

Given the multivariate nature of the two aims stated above, structural equation modeling (SEM) was used. SEM is a statistical modeling technique that allows for the simultaneous estimation of relationships between a number of observed and latent variables (Loehlin, 1998).

1.4.1 Predictions

Prediction 1: Regardless of modality, Gc and Gf will significantly predict lexical learning over and above other variables. Regardless of the language modality, lexical learning is best accomplished by constructing relationships between that which is known and that which is to be learned, and this implicates Gc and Gf.

Prediction 2: Regardless of modality, WMC will significantly predict lexical learning over and above other variables. WMC is a domain-general ability that allows individuals to maintain durable representations even in the face of interference. Word learning and sign learning require durable long-term memory representations and both should suffer from interference, thus the domain-general construct of WMC should be predictive of both.

Prediction 3: PSTM will only significantly predict lexical learning within modality. A significant proportion of the variance accounted for by STM tasks is domain-specific. Moreover, a prominent theory explaining the relationship between word learning and spoken-PSTM invokes modality-specific processes (Gathercole, 2006). As such, signed-PSTM should be predictive of sign learning but not word learning and spoken-PSTM should be predictive of word learning but not sign learning.

Prediction 4: sign learning and word learning will be distinguishable constructs. Sign learning and word learning are expected to rely on many of the same processes and therefore be highly correlated, however, due to differences in modality, the two should be distinguishable.

CHAPTER 2. METHODS

2.1 Participants

Participants were recruited from the Georgia Tech School of Psychology subject pool and surrounding community, including local colleges and universities. Georgia Tech students received course credit and an additional \$15 if they completed both sessions of the study. Community participants received \$30 for the first session and \$35 for the second session.

In order to participate in the study, participants had to be between the ages of 17-35, fluent in English, resided in the USA since at least the age of five, and have normal or corrected-to-normal hearing and vision. Due to the nature of the tasks and the aims of this study, participants were excluded if they indicated fluency in ASL or Turkish, were diagnosed with a language disorder, or if they possessed an upper-body injury or movement disorder affecting their arms or hands.

In total, 286 individuals consented to participate in the study. Of those individuals, 34 did not return for the second session of the study, 13 indicated poor English fluency, and three individuals were removed from the analysis because they were observed answering their cell phone, copying to-be-remembered items, or skipping task instructions—thus the final sample consisted of 236 participants. Additionally, it should be noted that one individual indicated having studied ASL as a child but reported very limited fluency and so was retained.

Within the final sample, 232 answered a demographic questionnaire, though not necessarily all questions. Based on the information provided, the mean age was 21.24 years ($SD = 3.57$); Approximately 62% of individuals (146/232) identified as female; all 232 individuals indicated that they had at least a high school diploma and nearly all (94.8%) indicated that they had at least some college education with 122 participants (52.6%) identifying as Georgia Tech students at the time of participation—the remaining 47.4% were community members, including students from local colleges and universities.

2.2 Procedure

The study consisted of two sessions, with nearly all tasks completed on a PC running E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) in a room with up to five participants; only a reading test and demographic questionnaire were completed on paper.

The first session lasted up to 2.5 hours and consisted of eight associative lexical learning tasks (four sign learning tasks and four word learning tasks) and six PSTM tasks (three signed-PSTM tasks and three spoken-PSTM tasks). The second session lasted up to 2 hours and consisted of eight intelligence tasks (four Gc tasks and four Gf tasks), three WMC tasks, an imagery questionnaire, the Object-Spatial Imagery Questionnaire (OSIQ; Blajenkova, Kozhevnikov, & Motes, 2006), and a language experience and demographics questionnaire. Task administration order was fixed and is presented in Table 1; task descriptions are provided in the next section. Note, the OSIQ and language

experience portion of the language experience and demographic questionnaire are not relevant to the present study and will not be discussed any further.

Table 1. Task administration order

	Session 1	Session 2
1	ASL-SL	Reading
2	PSL	Info
3	3TSL	Vocab
4	LetSpan	Gram
5	NWRec	Ospan
6	NWspan	SymSpan
7	DPSL	RoSpan
	[Optional 5min Break]	[Optional 5min Break]
8	TWL	Ravens
9	PWL	LetSets
10	3TWL	NumSeries
11	NSPT	SLAT
12	ProSign	OSIQ
13	SignCon	Questionnaire
14	DPWL	

Note: ASL-SL = ASL Sign Learning; PSL = pseudosign learning; 3TSL = three-term sign learning; LetSpan = letter span; NWRec = Nonword Recognition; NWspan = Nonword Span; DPSL = delayed pseudosign learning; TWL = Turkish word learning; PWL = pseudosign learning; 3TWL = three-term word learning; NSPT = nonsign paired task; ProSign = Probed Sign recognition task; SignCon = sign configuration task; DPWL = delayed word learning; Reading = test of reading comprehension; Info = information test; Vocab = extended range vocabulary test; Gram = grammar and usage test; Ravens = Raven’s Advanced Progressive Matrices, Set II; LetSets = letter sets; NumSeries = number series; SLAT = spatial learning ability test; Ospan = operation span; SymSpan = symmetry span; RoSpan = rotation span; OSIQ = object-spatial imagery questionnaire; Questionnaire = language experience and demographic questionnaire.

2.3 Tasks

Each task always began with instructions and at least one example item. Feedback was always provided during practice trials, however, the extent of the feedback ranged between simply stating whether the response was correct or giving a brief but detailed explanation. A research assistant was always present to observe participant as they completed each task and to answer any questions.

2.3.1 *Lexical Learning*

All lexical learning tasks utilized a similar associative learning paradigm with blocks consisting of study and test trials. First, each task began with instructions introducing the task, followed by a single example item. Next, participants were encouraged to use imagery or sentence generation to aid their learning. By suggesting that participants use strategies, the effect that Gc has on learning should be due to more general knowledge and experience and not to specific experience with associative learning tasks.

During the learning phase, a word or sign was presented aurally or visually—depending on the task—and immediately followed by its associate, a single English word, presented on screen in its written form for 1000ms for all tasks except the three-term tasks, which presented the word for 2000ms. After a number of pairs were presented, the testing phase would begin.

During the testing phase, participants viewed randomly selected stimulus items followed by a response screen with all the English words encountered in the task. The

participant was to click on the appropriate English word or guess. Once the participant made a response, the next item was presented, and so on. If the participant did not respond correctly to 100% of the items in a task then the task would continue with another block of trials until 100% of items were answered correctly or the maximum number of blocks (dependent on the task) was reached—whichever came first. Participants were never given explicit feedback or shown the correct associate during the test phase.

Scores were always calculated as the total number of correct responses across all trials, however, because participants vary in the number of trials necessary to reach the criterion, superfluous trials were awarded the maximum number of points. For example, in the Turkish word learning (TWL) task described below, a participant was charged with learning 15 pairs over a maximum of three trials. A participant who correctly identified all 15 pairs on the first trial would receive 15 points for that trial and 15 points for the two remaining trials for a total of 45 points. On the other hand, a participant who scored 10 on the first trial and 15 (i.e., 100%) on the second trial would receive 25 points for the attempted trials and an additional 15 points for the remaining trial—a total of 40 points.

All English words utilized in these tasks were selected from the SUBTLEX-US corpus (Brysbaert & New, 2009; Brysbaert, Warriner, & Kuperman, 2014; New, Brysbaert, Veronis, & Pallier, 2007) and were familiar concrete nouns ranging between 1-3 syllables and 4-6 characters in length. Familiar words were used to mimic what adults typically encounter when they first attempt to learn a new language. See the Appendix for a list of all English words used in this study.

All target lexical forms were either contrived or drawn from languages that are quite distinct from English, namely ASL and Turkish. This was done to limit the degree to which participants can rely on phonotactic and lexical knowledge, a strategy that would not generalize to all languages.

2.3.1.1 ASL Sign Learning (ASL-SL) Task

In the ASL-SL task, participants had up to two trials to learn 24 ASL signs and their associated English word pairs (Figure 1). ASL signs were selected from the ASL-LEX database (Caselli, Sehyr, Cohen-Goldberg, & Emmorey, 2016) such that 1) their English glosses conformed to the specifications listed above, 2) the ASL signs were low in iconicity (the mapping of form and meaning), and 3) signs were visually distinct, differing from each other in at least two of the following major phonological parameters: handshape, movement, or location (Brentari, 1998). A hearing native ASL signer performed all of the signs and the same video clips were used for both study and test trials. The maximum score was 48.

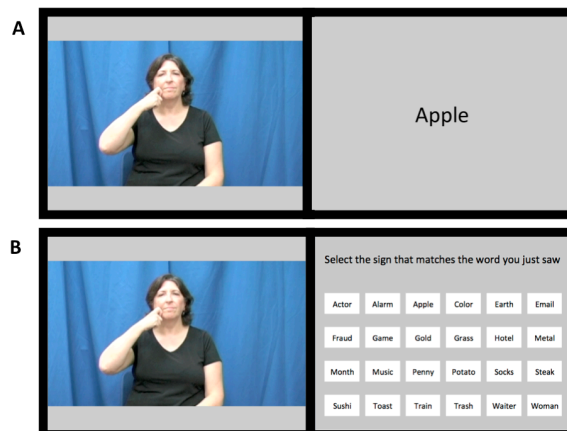


Figure 1. Depiction of the ASL-SL Task. Panel A depicts a study trial. Panel B depicts a test trial. In both cases, the sign is shown first and is immediately followed by either the response word in the study phase or the response screen in the test phase.

2.3.1.2 Pseudosign Learning (PSL) and Delayed Pseudosign Learning (DPSL) Tasks

Like the ASL-SL task, the PSL is a paired-associate task, however, it differs from the ASL-SL task in a number of ways. First, pseudosigns are used instead of real signs. Using pseudosigns confers greater control over such variables as iconicity and sign complexity. Second, the model performing the sign varied between the study and test phase of a block (see Figure 2). This reduced the possibility that participants could rely on extraneous details (i.e., the model slouching in one video while sitting straight in all others) and placed greater focus on the linguistic features of the signs. A native hearing signer performed all signs used during the study phase; test phase signs were reproductions of the study phase signs and were performed by a non-signer (the author). Third, a dropout procedure was used in which once a participant correctly identified a sign, it no longer appeared in any future block (i.e., study or test trial). This was done to control for the positive effect that overlearning (or the number of successful responses after achieving mastery) has on retention (Driskell, Willis, & Copper, 1992)—an important consideration for the Delayed Pseudosign Learning (DPSL) task.

The DPSL task consists of a single block of PSL test trials administered after four intervening tasks, approximately 30 min. after the PSL. As such, this task was intended to measure *retention* of lexical items, a construct that is substantially related to initial learning (Kyllonen & Tirre, 1988). Had a dropout procedure not been used in the PSL then individuals who mastered a list of PSL items on the first trial would have had three opportunities to overlearn while an individual who showed mastery of the entire list on the final trial could have anywhere between zero and three opportunities for overlearning.

As such, degree of overlearning would have confounded the relationship between initial learning (assessed by the PSL) and retention (assessed by the DPSL).

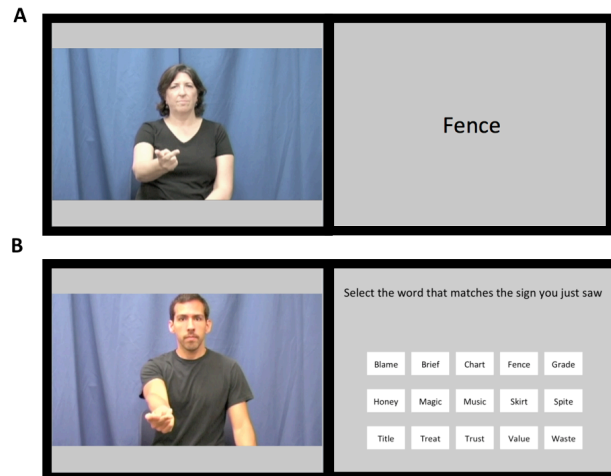


Figure 2. Depiction of the PSL task. Panel A shows a study trial. Panel B shows a test trial.

There were 15 PSL items and scores were calculated over a maximum of three trials for a possible score of 45. The DPSL consisted of a single test block of 15 items, however, in order to avoid penalizing participants for pairs they had not learned and to further remove the variance due to a participant's rate of learning, DPSL scores were calculated as a percentage of the number of pairs learned in the PSL. Thus the denominator used to calculate the DPSL score for an individual who correctly responded to 10/15 PSL items was 10. In the final analysis, only one individual had a score above 100% on the DPSL—this score was adjusted to 100%.

2.3.1.3 Three Term Sign Learning (3TSL) Task

The 3TSL is a complex associative learning task adapted from B. A. Williams and Pearlberg (2006) in which a stimulus is associated with three possible responses, contingent on a cue (see Figure 3). For example, during the study phase the stimulus sign, *S*, may be associated with *tree*, *bone*, and *fork*, and each response word is associated with the cues *1*, *2*, and *3*, respectively. The pseudosigns would be presented and immediately followed by instructions to press a particular key on the computer keyboard, for example, the 1-key. Once the button was pressed or after 2000ms had elapsed, the associated English word would be revealed for 2000ms. Next, the *same* pseudosign would be replayed, immediately followed by instructions to press the 2-key, and so on. During the test phase, a stimulus (e.g., *S*) and cue (e.g., *2*) would be presented followed by instructions to identify the associated English word (*bone* in this example). During the study phase, all English words associated with a particular pseudosign were presented sequentially, however, the order of the pseudosigns were randomized. During the test phase, pseudosign-cue combinations were presented randomly. Importantly, tasks of this type have been found to correlate with paired-associate tasks as well as Gf and WM (Kaufman et al., 2009; Tamez et al., 2008; B. A. Williams & Pearlberg, 2006).

In the 3TSL, there were 6 pseudosigns, each with three associated words and cues and scores were calculated based on performance over a maximum of three blocks. A non-signer performed all pseudosigns and the same movie clips were used

for the study and test phases. The maximum score was 54 (6 stimulus pseudosigns x 3 response words x 3 blocks).

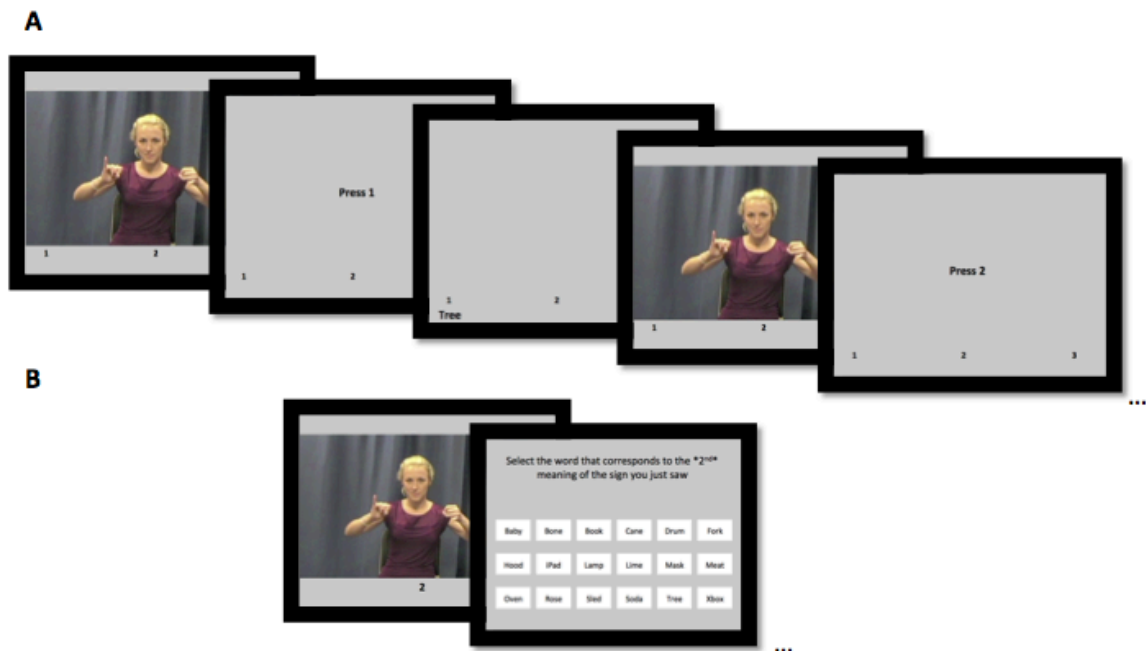


Figure 3. Depiction of the 3TSL. Panel A depicts part of a study trial. Panel B depicts part of a test trial.

2.3.1.4 Turkish Word Learning (TWL) Task

Participants attempted to learn 15 Turkish-English word pairs over a maximum of three blocks. The Turkish words were spoken by a native Turkish speaker from Istanbul and presented over headphones; the same audio clips were used for both study and test trials. The maximum possible score was 45. Note: Turkish was used because, as a Turkic language, it is part of the Turkic family of languages, a language family distinct from the Indo-European family, which includes English. As such, its lexicon is quite different from English and it possesses some phonemes that are not present in English.

2.3.1.5 Pseudoword Learning (PWL) and Delayed Pseudoword Learning (DPWL) Tasks

Like the PSL, the PWL employed a dropout procedure and two different people (in this case, two different female research assistants) produced the study and test items. All stimulus words were presented aurally over headphones. There were 15 pairs and participants' scores were calculated as the total correct over 3 blocks, for a maximum score of 45.

The DPWL learning consisted of a single test block of the PWL learning test items administered after four intervening tasks, approximately 40 min. after the PWL learning. The maximum possible score was a percentage of the total number of items a participant had learned in the PWL.

2.3.1.6 Three Term Word Learning (3TWL) Task

The 3TWL is similar to the 3TSL: six pseudowords (presented over headphones) were each associated with three English words and cues; participants had up to three trials to learn these items; the same audio clips were used during study and test trials; and the maximum possible score was 54.

2.3.2 *Phonological Short-Term Memory*

All PSTM tasks were either span tasks or discrimination (same-different) tasks. In a span task, a participant is presented with a set of items and is tasked with recalling the items in the order presented. Items were always selected from a limited pool of 9 to 12 items and the complete pool of items used in a task were always on display when participants responded. In order to reduce the role of WMC, an attempt was made to

reduce within-task item similarity (see Oberauer, Farrell, Jarrold, & Lewandowsky, 2016), either acoustically (Baddeley, 1966; Conrad & Hull, 1964) or visually (Wilson & Emmorey, 1997), depending on the variant of PSTM the task was intended to assess. In order to maximize individual differences in performance, sets varied in length and a partial credit scoring procedure with unit weighting was used (for details, see Conway et al., 2005, pp. 775-777). In partial credit unit scoring, participants receive credit for each item of a set recalled in its correct serial position, however, the amount awarded is equal to one over the total number of items in the set; thus correct recall of an entire set of five items merits 1 point while recall of 4/5 items merits .80 and recall of 7/9 items merits .78. Participants were not told all of the details of this scoring procedure; instead they were informed that they would receive one point for each item correctly recalled in its serial position and, to facilitate understanding, they were provided with feedback during practice trials. Feedback was never provided during the critical trials.

In the discrimination tasks, participants judged whether a target item or sets of items were the same or different from a reproduction of either a single target item or an entire set of target items, depending on the type of task. Relative to the span tasks, discrimination task items were drawn from larger pools (28 for the NSPT and NWRec and 16 for the ProSign) and, as such, it was difficult to limit within-task item similarity, though an effort was made to limit within-*set* item similarity.

During the response portion of a task, the response screen appeared simultaneously with the reproduction and participants were to use the computer mouse to click on buttons (i.e., text boxes) with the words “same” or “different” inscribed. The *same* button always appeared on the right hand side and the *different* button on the left.

Participants were able to make their judgments as soon as they recognized a difference and were warned that they should not make a *same* judgment until the entire reproduction was presented. Finally, it should be noted that during *same* trials, the exact same stimuli were used for both the target and reproduction.

2.3.2.1 Nonsign Repetition Task (NSPT)

The NSPT used here is a shortened version of the original NSPT (Martinez & Singleton, 2018). In the NSPT, participants must judge whether target pseudosigns differ from their reproductions (see Figure 4). Martinez and Singleton (2018) observed moderate to strong correlations amongst it, a sign learning task, two visuospatial STM tasks, and another putative task of signed-PSTM, the Nonsign Repetition Task (Mann, Marshall, Mason, & Morgan, 2010), providing evidence of the validity of the NSPT as a measure of signed-PSTM. The decision to use a shortened version was based on time limitations as well as the fact that the original version, a 20 min discrimination task, was quite onerous; this version of the task, took about half the time to complete.

The NSPT begins with a 164 second instructional video. The video introduces participants to the task and three phonological parameters: handshape, orientation, and movement. It was explained that 50% of pairs would be faithful reproductions and should be classified as “same” while the other 50% of reproductions would differ on *one* of the forenamed parameters. In this way, a verbal mediation strategy in which a participant names a sign is (it is assumed) rendered ineffective and therefore the participant must rely on visuospatial and possibly motor processes. For example, the pseudosign depicted in Figure 5A consists of a downward motion of the arms, with palms facing away from the

body, and fingers wiggling. A participant may encode this as “rain” or simply “down” but if the reproduction is like that shown in Figure 5B where the pseudosign differs from the original in one parameter (here, the orientation of the palm), then neither “rain” nor “down” aids in discriminating between two.



Figure 4. Depiction of the NSPT. In the NSPT, there were 28 target signs each with two reproductions, produced by different individuals.

Next, participants were told that there would be two blocks. The same target pseudosigns would be used across both blocks and one individual would perform all target pseudosigns; the individual performing the reproductions, however, would differ across the two blocks, and pairs would be presented in a different order from one block to the other. Next, participants completed two blocks of three practice trials with automated feedback. The automated feedback either informed participants that they were correct, or if they were wrong, displayed a screen with a brief text description of the error as well as side by side static images of the target and reproduction with differences highlighted. The critical trials followed.

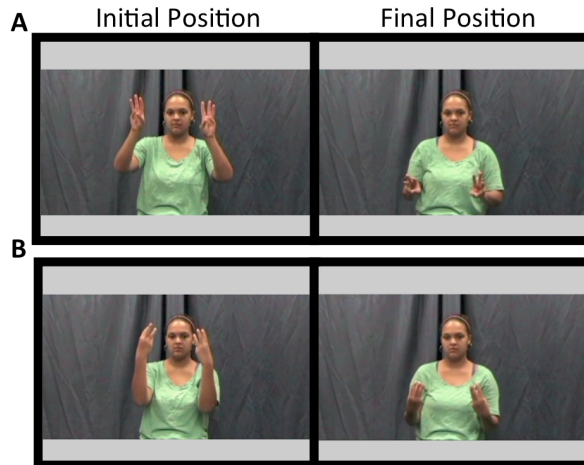


Figure 5. Example of two pseudosigns differing in orientation only. Note: this item is from the ProSign. It was chosen because it is amenable to a simple label.

The two critical blocks each consisted of 28 items for a maximum score of 56 points—feedback was never provided. This scoring scheme differs from that used by Martinez and Singleton (2018) in which items across the two blocks were paired according to the target sign and a point was awarded only if the responses to both reproductions were correct. This divergence was justified on the grounds that a reanalysis of the Martinez and Singleton (2018) data found that both scoring schemes produced similar results though there was a slight increase in reliability (indexed by Cronbach’s alpha) when items were left unpaired across blocks. Given the smaller number of items used in this version of the NSPT, reliability was at a premium.

2.3.2.2 Probed Sign (ProSign) Task

In the ProSign task, participants viewed sets of pseudosigns followed by a cue (500ms) and a probe; participants were to indicate whether the probed pseudosign was in the set just viewed or if it was different. If a probe was different, then, as in the NSPT, it differed from one of the pseudosigns in the set by one parameter: handshape, movement,

or orientation; if it was the same, then the pseudosign (and video clip) was exactly the same as a pseudosign in the set. To limit item similarity within-set, pseudosigns differed from each other on at least two of the aforementioned parameters.



Figure 6. Depiction of a ProSign item. The probe differs from the second pseudosign in the set.

There were 40 critical trials with 10 trials each at set lengths three through six. Half of all trials were *different* trials with six differing from the target in handshape, seven in orientation, and seven in movement. In an attempt to maximize individual differences in performance, the majority of the forty trials assessed memory for pseudosigns between the first and last pseudosigns in a set, as recall of items in the first and last positions of a set tend to be at or near ceiling (Jones, Farrand, Stuart, & Morris, 1995; Unsworth & Engle, 2007; Ward, Avons, & Melling, 2005; Wu & Coulson, 2014). In all, eight (20% of all trials) assessed memory for the first item, eight (20%) assessed memory for the final item, and 24 (60%) assessed memory for pseudosigns in between.

2.3.2.3 Sign Configuration Task (SignCon)

The SignCon is a dual-task in which participants completed two span tasks: a letter span (described in detail in section 2.3.2.4, below) and a pseudosign span (see Figure 7). The critical portion of the SignCon is the pseudosign span portion, however, to

limit the role of WMC, within-task item similarity was low, potentially enabling participants to effectively use a verbal mediation strategy (e.g., labeling) and articulatory rehearsal—the letter span portion of the task was meant to prevent the use of this strategy. Moreover, participants were explicitly told not to attempt to label any of the pseudosigns. To check for compliance, 40% of trials assessed only the letter span portion and 60% assessed only the pseudosign portion.

Every trial of the SignCon began with participants viewing sets of letters followed by one to four video clips of pseudosigns. The length of the set of letters was always equal to one minus the participant’s letter span—the maximum number of letters that could be perfectly recalled in serial order for three trials—calculated from the participant’s performance on the LetterSpan task completed earlier in the session. In this way, a participant’s ability to rehearse should be prevented and the memory load should be functionally equivalent across participants.

After the set of letters were presented for a length of time equal to 500 ms per letter, participants viewed one to four pseudosigns which, when set length was greater than one, differed from each other in at least two of the following parameters: handshape, movement, and/or location. Next, participants were tested on either the letters or the pseudosigns. On 40% of trials, participants recalled the set of letters in serial order. On the remaining 60% of trials, participants recalled the pseudosigns by clicking on static images—showing either the initial or final position of the sign—in order.

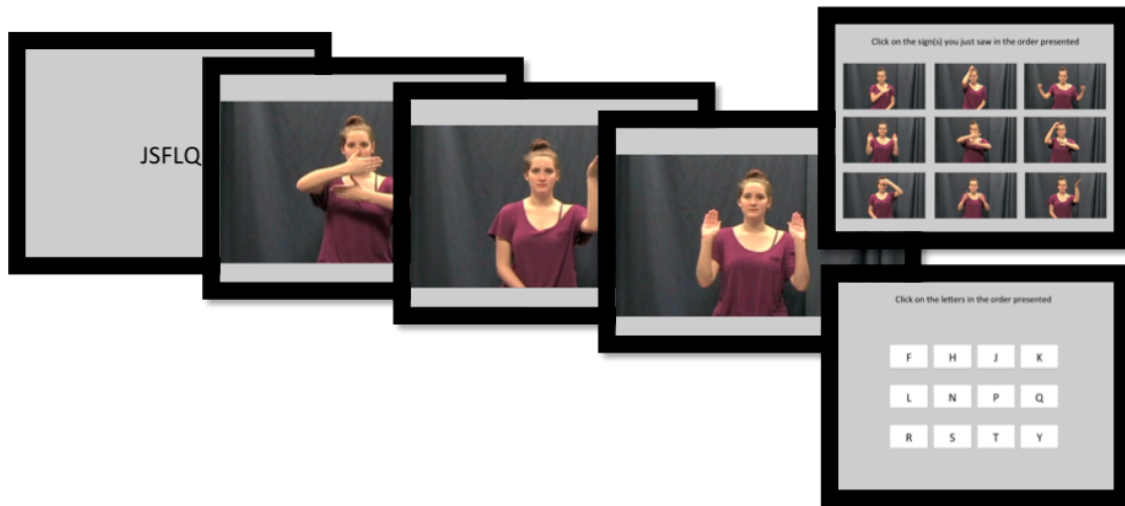


Figure 7. Depiction of the SignCon. Participants were either tested on the number of pseudosigns (60% of trials) or letters (40% of trials) they could recall in order.

Twelve of the 20 trials were pseudosign trials and there were three trials at each set length. Using the partial credit unit scoring procedure described above, the maximum possible score was 12.

2.3.2.4 Letter Span (LetSpan)

In this task, participants attempted to recall four to nine letters in serial order. The pool of items consisted of 12 letters: F, H, J, K, L, N, P, Q, R, S, T, Y. The entire set of letters was presented on screen for a length of time equal to the set length times 500ms (e.g., a set of 6 letters was presented for 3000ms). There were three trials at each set length for a total of 18 sets. Using partial credit unit scoring, the maximum was 18 points.

2.3.2.5 Nonword Recognition (NWRec) Task

The NWRec task was adapted from Gathercole, Pickering, Hall, and Peaker (2001) and similar tasks have been used by others (e.g., K. I. Martin & Ellis, 2012;

O'Brien et al., 2006). In the NWRec task, participants discriminate between two sequences of pseudowords presented aurally via headphones. If the sequences were different, then two neighboring pseudowords were transposed; if they were the same, then the exact same sequence of pseudowords was presented again. There were a total of 36 trials with four trials of set length three, six trials at set length four, and eight trials at set length five. Moreover, 1/3 of different trials contained a transposition of the first and second pseudowords, 1/3 were transpositions of the final and penultimate pseudowords, and the remaining were transpositions of pseudowords in between. Pseudowords were drawn from a pool of 28 items and were selected from Gathercole et al. (2001). The maximum score was 36.

2.3.2.6 Nonword Span (NWSpan) Task

In the NWSpan, participants heard a set of monosyllabic pseudowords over headphones and attempted to recall the pseudowords in the order presented by clicking on a response screen with the entire pool of words displayed. The pool of pseudowords consisted of 12 pseudowords drawn from Gathercole et al. (2001). Pseudowords were presented in sets ranging between two and six and there were three trials at each set length for a total of 15 trials. Using partial credit unit scoring, the maximum score was 15.

2.3.3 Intelligence

The following holds true for all intelligence tests used in this study: 1) test format was multiple-choice, 2) there was a time limit, 3) questions were generally ordered from

easiest to hardest, and 4) participants were told that they should work quickly but accurately and, when necessary, guess.

2.3.3.1 Test of Reading Comprehension (Reading)

Participants had up to 20 min. to read 5 passages (varying in length from 112 words to 739 words) and answer 17 questions. All passages and their corresponding questions were drawn from released SAT and GRE tests and were selected to provide a range in item difficulty. The test was administered in paper format and participants were encouraged to use whatever strategies they normally would use *except* answering questions out of order. The maximum score was 17.

2.3.3.2 Information (Info) Test

The Info test consisted of two parts. In part 1, participants had up to 7 min to answer 40 general knowledge questions from the Information subscale of the Multidimensional Aptitude Battery II (Jackson, 1998). In part 2, participants were allowed 2 min to answer an additional 11 questions. These questions were written by the present author and were added to broaden the domains of knowledge assessed and to increase the difficulty of the test to a level appropriate for a sample that, relative to the general population, would be disproportionally college educated. Performance across both parts were summed to form one score, thus the maximum score was 51.

2.3.3.3 Extended Range Vocabulary (Vocab) Test

In the vocab test (Ekstrom, French, Harman, & Dermen, 1976), participants are presented with a word and attempt to match it with one of five words that is closest in

meaning. There were two parts, each with 24 items, and a time limit of 6 min. The maximum was 48.

2.3.3.4 Grammar and Usage (Gram) Test

The Gram test consisted of 21 “improving sentences” items selected from sections 5 and 10 of official SAT practice tests released between 2004 and 2013. Each item consisted of a sentence with a portion underlined; the participant was to select the answer choice that best rephrased the underlined portion or, if the original phrasing was the best choice, select the first answer choice, which always repeated the original phrasing. Participants had up to 10 min to complete the test. The maximum score was 21.

2.3.3.5 Raven’s Advanced Progressive Matrices, Set II (Ravens)

In Ravens, Participants were presented with 18 3x3 matrices with all but the lower right cell of each matrix containing figures. The figures in each matrix were arranged according to a particular rule (see Carpenter, Just, & Shell, 1990) and it was up to participants to infer the rule and select which of eight figures presented below the matrix best completed the pattern. The 18 items used in this task were the odd items from set II of Raven’s Advanced Progressive Matrices (Raven, Raven, & Court, 1998). Participants had 10 min to complete the task and the maximum score was 18.

2.3.3.6 Letter Sets (LetSets)

In the LetSets task (Ekstrom et al., 1976), participants were presented with five sets of letters, each consisting of four letters. The participant was to identify the one set of

letters that did not obey the same rule as the others. There were 30 problems and participants were given up to 7 min to complete the task. The maximum score was 30.

2.3.3.7 Number Series (NumSeries)

In NumSeries (Thurstone, 1938), participants were presented with a series of numbers that obeyed a particular rule. The participant's task was to complete the series by selecting the one answer choice (out of five) that would continue to series. Participants had up to 5 minutes to complete 15 items. The maximum score was 15.

2.3.3.8 Spatial Learning Ability Test (SLAT)

The SLAT used here is an adaptation of the SLAT described by Embretson (1992). In this version of the SLAT (see Figure 8), participants were presented with a representation of an unfolded cube. The six faces of the target contained simple shapes such as arrows and pentagons. The participant was to choose which of four cubes matched the target by mentally rotating and folding the target to compare with the four choices. Tasks such as these tend to correlate moderately to strongly with putative measures of Gf (Lohman, 1996; Marshalek, Lohman, & Snow, 1983; Varriale, van der Molen, & De Pascalis, 2018) and so it is being used here an indicator of that construct. There were 20 items and participants had up to 15 min to complete them.

It should be noted that the original SLAT consisted of a pretest, an intervention, and a posttest and the dependent variable was an estimate of the learning that occurred due to the intervention (for further details, refer to Embretson, 1992). In this version of

the SLAT, there is only a single test and no intervention and so it does not measure learning in any appreciable way.

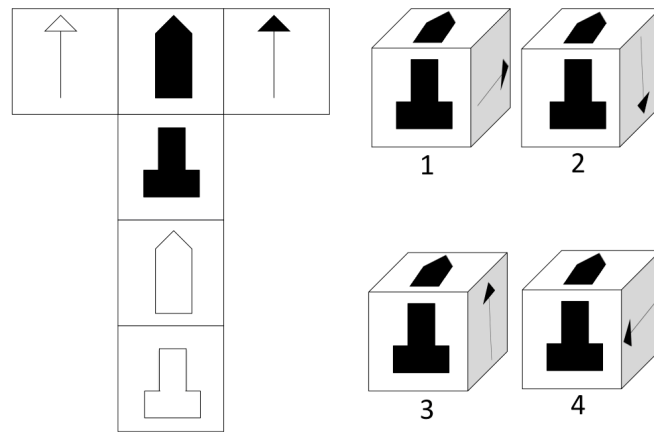


Figure 8. Depiction of a SLAT item. The correct answer choice is 1.

2.3.4 Working Memory Capacity

The WMC tasks used here were all shortened versions of the complex span tasks described by Foster et al. (2015). In a complex span task, participants complete a primary memory task and a secondary processing task. The dependent variable is the number of items from the primary task that the participant is able to remember in correct order. As with the PSTM span tasks described above, WMC tasks were scored using partial credit unit scoring.

Each task began with instructions and three blocks of practice. In the first block of practice, participants completed the memory portion of the task alone. In the second block of practice, participants completed the processing component alone. In the final block of practice, participants completed both the primary memory component and the

secondary processing component. After completing the practice block, participants completed one block of critical trials.

2.3.4.1 Operation Span (OSpan)

In the OSpan, participants were presented with a series of letters with math equations interleaved between letter presentations. Participants were to try to remember the letters in the order they were presented. There was one set at each set length of three through seven for a total of five trials. Using partial credit unit scoring, the maximum score was 5.

2.3.4.2 Symmetry Span (SymSpan)

In this task, the primary (memory) task was to remember the sequence of locations of a red square in a 4x4 matrix. The secondary task was to judge whether a figure composed of shaded squares on an 8x8 matrix was symmetrical along the vertical axis. The number of locations to be remembered varied from two to five per trial, for a total number of four trials. Using partial credit unit scoring, the maximum score was 4.

2.3.4.3 Rotation Span (RoSpan)

The primary task in the RoSpan was to remember a sequence of arrows varying in size and direction. The Secondary task was to judge whether a rotated letter, when mentally rotated to its upright position, is displayed correctly or is mirrored. The number of arrows to be remembered varied between two and five, for a total of four trials. Using partial credit unit scoring, the maximum score was 4.

2.3.5 Summary of Tasks

Table 2. Summary of tasks

Construct	Task	Brief Description
Sign Learning		
	ASL Sign Learning (ASL-SL)	Paired associate learning task using American Sign Language signs
	Pseudosign Learning (PSL)	Paired associate learning task utilizing pseudosigns and a dropout procedure
	Three-Term Sign Learning (3TSL)	Complex associate learning task using signs
	Delayed Pseudosign Learning (DPSL)	One test block of the PSL, administered about 30 min after the PSL
Word Learning		
	Turkish Word Learning (TWL)	Paired associate learning task using Turkish words
	Pseudoword Learning (PWL)	Paired associate task using pseudowords and a drop out procedure
	Three-Term Word Learning (3TWL)	Complex associate task using pseudowords
	Pseudoword Learning (PWL)	One test block of the PWL, administered about 40 min after the PWL
Signed-PSTM		
	Nonsign Paired Task (NSPT)	Paired discrimination task using pseudosigns
	Probed Sign (ProSign)	Probed discrimination task using pseudosigns
	Sign Configuration (SignCon)	Dual span task using letters and signs
Spoken-PSTM		
	Letter Span (LetSpan)	Span task using letters
	Nonword Recognition (NWRec)	Discrimination task using pseudowords
	Nonword Span (NWSpan)	Span task using pseudowords
Gc		
	Reading Comprehension (Reading)	Reading comprehension test
	Information (Info)	Test of general knowledge
	Vocabulary (Vocab)	Synonyms test
	Grammar and Usage (Gram)	Test of grammar and usage
Gf		
	Number Series (NumSeries)	Induction task using numbers
	Letter Sets (LetSets)	Induction task using letters
	Raven's Matrices (Ravens)	Figural reasoning (induction) test
	Spatial Learning Aptitude Test (SLAT)	Test requiring mental rotation and folding
WMC		
	Operation Span (OSpan)	Dual-task with a memory component requiring recall of letters and a processing component requiring the verification of simple math equations
	Symmetry Span (SymSpan)	Dual-task with a memory component requiring recall of the position of a red square and a processing component requiring a symmetry judgment
	Rotation Span (RoSpan)	Dual-task with a memory component requiring recall of arrows of different sizes and orientations and a processing component requiring mental rotation

2.4 Statistical Analyses

2.4.1 Data Screening

There were 61 missing values and 11 values were classified as outliers because they were 3.5 standard deviations from the mean. Missing values were deemed *missing at random* and so the expectation-maximization (EM) algorithm was used to impute those values (Little & Rubin, 2014; Rubin, 1976). Outliers were treated in one of two ways, either by imputing the values using the EM algorithm or by replacing the values with a score equal to 3.5 standard deviations from the mean, whichever value was smallest. Two scores (both for the OSpan) were still below the cutoff after imputation and so they were replaced with a value equal to a z-score of 3.5.

Multivariate normality was assessed using a normalized version of Mardia's coefficient (Mardia, 1970). Bentler (2001) suggests that a value above five is suggestive of non-normality. As will be observed below, all values of Mardia's normalized estimate were below five and so no actions were taken to correct for non-normality.

2.4.2 Statistical Procedure

Structural equation models were created and analyzed using EQS (Bentler, 2001). Because the data appeared to be normally distributed, model parameters were estimated using maximum likelihood, a method that yields the smallest errors when the data are normal (Ullman, 2006).

Model fit was assessed using several statistics recommended by Kline (2016): model chi-square (with associated degrees of freedom and p-value), Comparative Fit

Index (CFI), Standardized Root Mean Square Residual (SRMSR), and the Root Mean Square Error of Approximation (RMSEA). The chi-square test, SRMSR, and RMSEA are “badness-of-fit” tests—lower values indicate good fit; The CFI, on the other hand, is a *goodness-of-fit* test.

Finally, estimated parameters (e.g., path coefficients) were assessed using significance tests; a value of 0.05 was considered significant.

CHAPTER 3. RESULTS

3.1 Observed Variable Analyses: Descriptive Statistics, Reliability, and Correlations

Descriptive statistics and internal consistency coefficients (Cronbach's alpha) are provided in Table 3. All tasks were, on average, sufficiently difficult for individual differences research and the data were approximately normally distributed. Nearly all coefficient alphas were at or near .80, suggesting acceptable reliability (cf., Draheim, Mashburn, Martin, & Engle, 2018). Only four tasks, the NSPT, ProSign, SymSpan, and RoSpan had coefficients below .70, however, these tasks tended to show strong correlations with tasks measuring the same or similar constructs, indicating that they were valid measures of the intended constructs.

Bivariate correlations are provided in Table 4. All tasks were significantly correlated to each other at $p < .01$. More importantly, the correlation matrix shows evidence of discriminant and convergent validity. For example, the LetSpan correlates strongly with NWRec and NWSpan (.55 and .58, respectively) but correlations with other tasks range between .23 (with ASL-SL) and .44 (with Vocab).

Table 3. Descriptive statistics and reliabilities

Task	Mean (SD)	Range	Skew	Kurtosis	α
1. ASL-SL	.71 (.22)	.08-1.00	-.96	.27	.93
2. PSL	.71 (.20)	.07-1.00	-.78	-.02	.87
3. 3TSL	.63 (.27)	.04-1.00	-.53	-.79	.95
4. DPSL	.70 (.22)	.00-1.00	-.57	.01	.82
5. TWL	.54 (.24)	.02-1.00	-.24	-.88	.91
6. PWL	.64 (.25)	.07-1.00	-.46	-1.03	.91
7. 3TWL	.49 (.33)	.00-1.00	.03	-1.53	.98
8. DPWL	.58 (.25)	.00-1.00	-.40	-.42	.83
9. NSPT	.80 (.07)	.57-.98	-.53	.51	.66
10. ProSign	.67 (.11)	.40-.90	-.25	-.31	.57
11. SignCon	.61 (.17)	.06-.92	-.67	.50	.75
12. LetSpan	.81 (.09)	.49-1.00	-.65	.45	.78
13. NWRec	.80 (.12)	.42-1.00	-.61	-.12	.75
14. NWSpan	.70 (.11)	.40-.99	-.15	-.04	.76
15. Reading	.50 (.22)	.00-1.00	-.00	-.55	.79
16. Info	.60 (.13)	.14-.90	-.85	.98	.83
17. Vocab	.53 (.15)	.15-.85	-.03	-.36	.84
18. Gram	.45 (.20)	.00-.95	.15	-.52	.78
19. NumSeries	.67 (.20)	.13-1.00	-.44	-.60	.77
20. LetSets	.57 (.16)	.17-.90	-.39	-.41	.86
21. Ravens	.57 (.21)	.06-1.00	-.37	-.41	.80
22. SLAT	.54 (.25)	.00-.95	.11	-1.17	.86
23. OSpan	.82 (.18)	.18-1.00	-1.29	1.63	.74
24. SymSpan	.74 (.23)	.00-1.00	-.99	.87	.67
25. RoSpan	.61 (.22)	.00-1.00	-.75	.18	.64

Note: ASL-SL = ASL sign learning; PSL = pseudosign learning; 3TSL = three-term sign learning; DPSL = delayed pseudosign learning; LetSpan = letter span; NWRec = Nonword Recognition; NWSpan = Nonword Span; TWL = Turkish word learning; PWL = pseudosign learning; 3TWL = three-term word learning; DPWL = delayed word learning; NSPT = nonsign paired task; ProSign = Probed Sign recognition task; SignCon = sign configuration task; Reading = test of reading comprehension; Info = information test; Vocab = extended range vocabulary test; Gram = grammar and usage test; Ravens = Raven's Advanced Progressive Matrices, Set II; LetSets = letter sets; NumSeries = number series; SLAT = spatial learning ability test; OSpan = operation span; SymSpan = symmetry span; RoSpan = rotation span.

Table 4. Zero-order correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
1. ASL-SL																									
2. PSL	.70																								
3. 3TSL	.65	.63																							
4. DPSL	.73	.70	.61																						
5. TWL	.62	.66	.51	.72																					
6. PWL	.58	.70	.52	.67	.78																				
7. 3TWL	.52	.55	.53	.58	.64	.66																			
8. DPWL	.57	.65	.50	.76	.77	.75	.64																		
9. NSPT	.47	.41	.38	.53	.42	.46	.37	.48																	
10. ProSign	.49	.47	.44	.53	.50	.48	.44	.47	.62																
11. SignCon	.56	.51	.49	.59	.52	.54	.53	.57	.54	.58															
12. LetSpan	.23	.31	.23	.37	.42	.37	.29	.39	.36	.34	.27														
13. NWRec	.44	.48	.39	.54	.60	.54	.45	.55	.42	.48	.50	.55													
14. NWSpan	.39	.42	.39	.50	.60	.52	.46	.55	.47	.47	.52	.58	.63												
15. Reading	.41	.33	.31	.42	.40	.38	.31	.34	.45	.30	.29	.31	.32	.38											
16. Info	.48	.42	.33	.47	.55	.50	.42	.47	.51	.43	.38	.33	.40	.43	.66										
17. Vocab	.42	.33	.28	.44	.50	.47	.38	.45	.45	.40	.35	.44	.50	.43	.57	.62									
18. Gram	.41	.35	.34	.43	.44	.39	.39	.38	.51	.41	.34	.36	.45	.39	.55	.61	.58								
19. NumSeries	.42	.41	.33	.53	.50	.47	.42	.46	.43	.38	.38	.40	.43	.46	.55	.49	.50	.46							
20. LetSets	.48	.44	.37	.53	.52	.56	.44	.46	.43	.41	.46	.41	.46	.42	.43	.44	.55	.40	.66						
21. Ravens	.57	.51	.42	.62	.56	.54	.51	.55	.53	.48	.53	.26	.43	.46	.46	.50	.50	.43	.61	.57					
22. SLAT	.47	.45	.43	.53	.46	.44	.43	.47	.48	.45	.48	.27	.40	.42	.47	.50	.45	.43	.59	.55	.64				
23. OSpan	.32	.37	.34	.38	.39	.32	.25	.36	.42	.34	.34	.42	.43	.50	.30	.28	.36	.34	.47	.44	.37	.37			
24. SymSpan	.29	.35	.30	.38	.38	.33	.27	.38	.44	.37	.39	.33	.35	.36	.29	.28	.40	.31	.49	.41	.49	.46	.43		
25. RoSpan	.24	.27	.25	.38	.32	.25	.19	.36	.33	.31	.37	.27	.28	.32	.31	.20	.29	.28	.44	.33	.41	.33	.40	.54	

Note: See Table 3 for key.

3.2 Latent Variable Analyses

To assess the validity of the tasks used here, the manifest variables in this study were grouped into factors and SEM was used to model the relationships amongst the latent variables. Model fit was good (Table 5, *Corr* model), however, inspection of the results of the Lagrange Multiplier test offered by EQS (Bentler, 2001) revealed that two pairs of tasks shared a significant amount of variance: 1) 3TSL and 3TWL and 2) LetSets and NumSeries. These pairs of tasks are very similar in format and so it was deemed appropriate to account for this *method variance* by correlating their residuals. As can be seen (Table 5, *Corr-LM* model), these corrections resulted in a significantly better fitting model, $\Delta\chi^2(2) = 29.587, p < .001$, and so they were retained.

Table 5. Correlated factors model fit statistics

Model	Mardia's X^2	df	CFI	SRMR	RMSEA (95% CI)	
Corr	3.16	426.03	251	.953	.044	.054 (.045, .063)
Corr-LM	3.16	396.44	249	.961	.044	.050 (.041, .059)

Note: Corr = correlated factors model; Corr-LM = correlated factors model with corrections suggested by the Leverage Multiplier test.

Figure 9A, illustrates all latent variables (circles) and their corresponding observed variables (rectangles) as entered into the Corr-LM model, along with estimated path coefficients. For clarity, correlations are shown separately in Table 6.

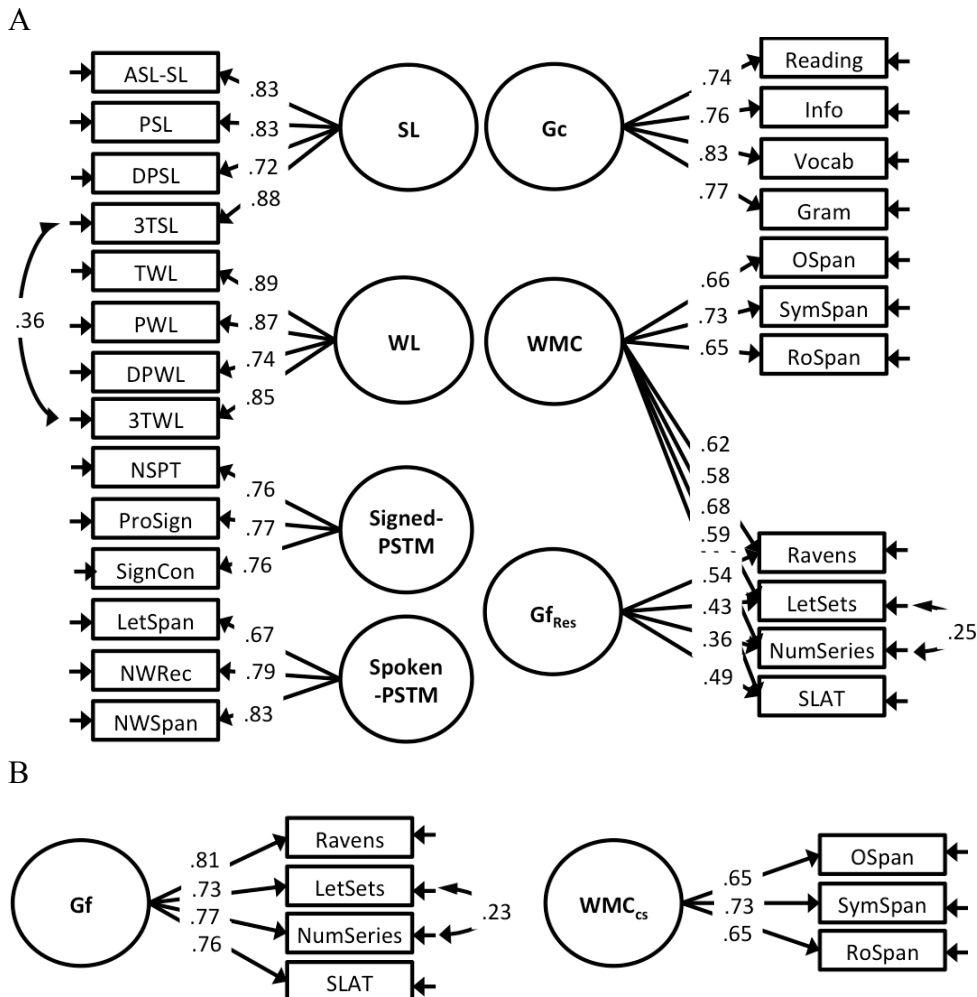


Figure 9. Latent variable and their indicators. Panel A shows the estimated path coefficients derived from analyzing Model Corr-LM. For reference, panel B shows the unresidualized Gf factor and its indicators as well as a WMC factor derived from the variance of complex span tasks only.

Table 6. Latent variable correlations

	SL	WL	Signed -PSTM	Spoken -PSTM	Gc	Gf _{Res}
1. SL						
2. WL	.88					
3. Signed-PSTM	.79	.74				
4. Spoken-PSTM	.65	.76	.75			
5. Gc	.62	.66	.68	.65		
6. Gf_{Res}	.54	.50	.40	.22	.54	
7. WMC	.57	.56	.69	.69	.57	---

Note: SL = sign learning; WL = word learning; PSTM = phonological short-term memory; Gc = crystallized intelligence; Gf_{Res} = fluid intelligence with variance accounted for by WMC partialled out (hence, residualized); WMC = working memory capacity.

Table 7. Correlations with Gf and WMC_{CS}

	SL	WL	Signed -PSTM	Spoken -PSTM	Gc	Gf
Gf	.77	.75	.78	.68	.79	
WMC_{CS}	.57	.56	.71	.68	.56	.80

Note: Note: SL = sign learning; WL = word learning; PSTM = phonological short-term memory; Gc = crystallized intelligence; Gf = fluid intelligence; WMC_{CS} = latent variable constructed from the variance due to complex span tasks only.

There are several things to note. First, in line with recent research (J. D. Martin et al., 2017), the Gf factor was residualized by partialing out the variance accounted for by WMC. This was done so that the Gf_{Res} factor would primarily represent individual differences in the ability to disengage from information while the WMC factor would represent a domain-general ability to maintain information in the face of interference. WMC and Gf also tend to be strongly related (Ackerman, Beier, & Boyle, 2005; Engle et al., 1999; Kane et al., 2004; Kyllonen & Christal, 1990)—and, in fact were here as well (see Table 7)—which can result in multicollinearity. Second, the fact that the Corr-LM model fits the data well and nearly all of the path coefficients between observed and latent variables were strong provides evidence of the validity of these tasks as measures of their intended constructs; only three path coefficients were below .50, however, this is

an outcome of the variance due to these tasks being split between the $G_{f_{Res}}$ and WMC factors. Third, the correlations amongst the latent variables and in particular those concerning the WMC and $G_{f_{Res}}$ factors speak to the appropriateness of modeling WMC and Gf as was done here and elsewhere (J. D. Martin et al., 2017). Specifically, the WMC factor is most strongly correlated with two other memory factors, Signed- and Spoken-PSTM, while the $G_{f_{Res}}$ factor correlates strongly with those factors that involve complex cognition. As further evidence, it should be noted that the path coefficients between the complex span tasks did not change substantially from what was observed when a WMC factor (WMC_{cs}) was constructed with only complex span tasks loading onto it (compare WMC in Figure 9A with WMC_{cs} in Figure 9B) nor did the correlations with other latent variables (compare Tables 6 and 7). Finally, it should be noted that the correlation between the sign learning (SL) and word learning (WL) factors was very strong (.88) but not perfect, suggesting that these latent variables are at least somewhat distinguishable (how to *best* model performance on the lexical learning tasks will be explored in section 3.2.2).

3.2.1 Predicting Sign and Word Learning

In this analysis, sign learning and word learning were set as outcome variables and the other latent variables were entered as predictors in a step-wise fashion. The first model was intended to assess the contribution of intelligence, indicated by Gf (unresidualized) and Gc. Model fit was good (Table 8, Model 1). As can be seen in Figure 10, Gf significantly predicted both sign learning and word learning but Gc did not make a significant contribution above and beyond Gf. By squaring the disturbance terms (inscribed in rectangles emanating from the outcome variables) and subtracting from 1,

we can calculate the proportion of variance accounted for by the predictors. In this case, the predictors accounted for 60% of sign learning variance and 57% percent of word learning variance. Moreover, the disturbance terms were significantly correlated (.72).

Table 8. Fit statistics for predictive model

Model	Mardia's χ^2	df	CFI	SRMR	RMSEA (95% CI)	
1	2.58	144.69	96	.98	.035	.046 (.030, .061)
2	3.65	208.41	137	.98	.039	.047 (.034, .059)
3	3.50	288.43	189	.97	.042	.047 (.036, .058)
4	3.16	396.44	249	.96	.044	.050 (.041, .059)

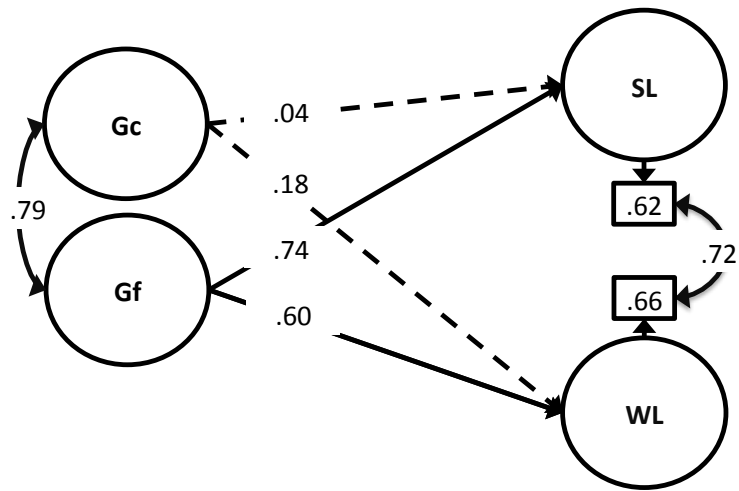


Figure 10. Model 1—the effect of intelligence

Next, WMC was added to the model and Gf was residualized. Model fit was good (Table 8, Model 2), however, it can be seen that, for the most part, this model is a redefined version of Model 1; that is, the variance in sign learning and word learning explained by WM is simply a portion of that which was already accounted for by the Gf variable in Model 1. This is supported by the fact that the proportions of variance accounted for in Model 1 and 2 are nearly identical (note the similarity in the disturbance

terms in Figure 10 and 11). Still, by partitioning the variance in this way, we can see that those processes that are common to complex span and Gf tasks and those that are unique to Gf are predictive of sign learning and word learning.

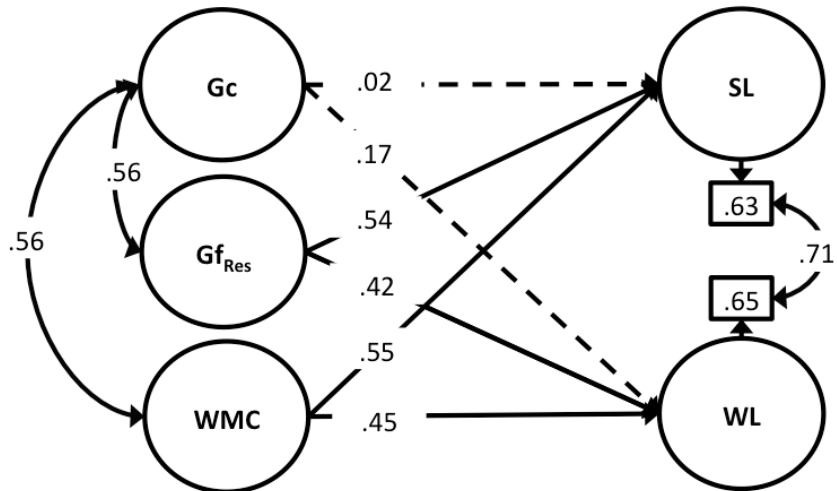


Figure 11. Model 2—accounting for WMC

In Model 3, Spoken-PSTM was added. Model fit was good (Table 8) and the proportion of sign learning and word learning variance accounted for increased to 66% and 72%, respectively, while the correlation between the disturbance terms dropped to .67. Importantly, the inclusion of the Spoken-PSTM factor resulted in the path between WMC and word learning becoming insignificant. This suggests that, in relation to word learning, Spoken-PSTM assesses very similar processes as WMC, however, Spoken-PSTM assess other relevant processes above and beyond those assessed by WMC.

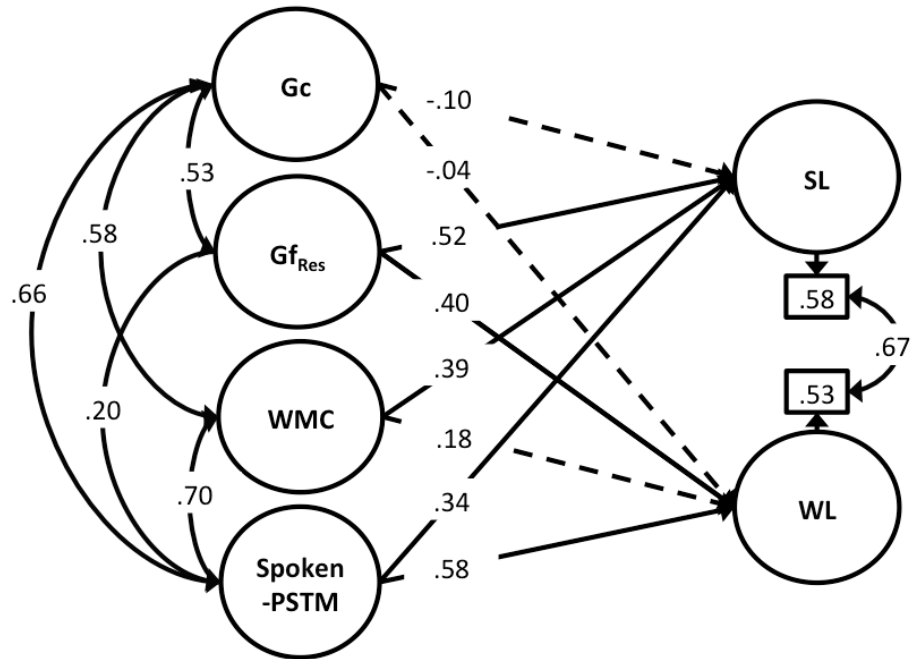


Figure 12. Model 3—Accounting for spoken-PSTM

Finally, in Model 4, Signed-PSTM was added. Model fit was good (Table 8); the proportion of sign learning and word learning variance accounted for were 71% and 72%, respectively; and the correlation between the disturbance terms was .67. Here, adding Signed-PSTM resulted in WMC and Spoken-PSTM no longer being significantly predictive of sign learning. Ultimately, it was only Gf and Signed-PSTM that significantly predicted sign learning while Gf and Spoken-PSTM were the only significant predictors of word learning (see Figure 13).

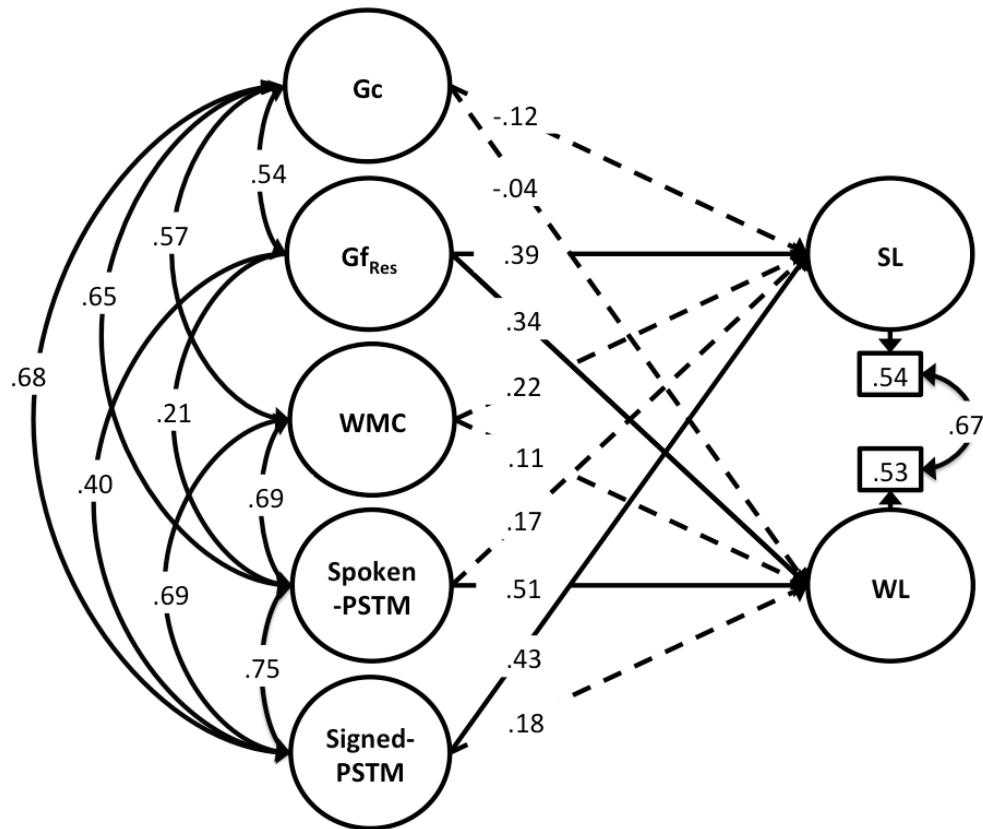


Figure 13. Model 4—Accounting for signed-PSTM

3.2.2 Modelling the Relationship Between Sign and Word Learning

Next, SEM was used to directly explore the relationship between sign learning and word learning. As was observed in Table 6, the correlation between the sign learning and word learning factors was quite strong, suggesting that a general lexical learning factor underlies performance on all lexical learning tasks used in this study. To investigate this possibility, a one-factor model was designated by loading all lexical learning tasks onto a single factor; next, this model was compared with a baseline model consisting of separate but correlated sign learning and word learning factors.

Table 9. Fit statistics for exploratory models

Model	Mardia's	X^2	<i>df</i>	CFI	SRMR	RMSEA (90% CI)
CF	4.31	39.06	18	.99	.027	.071 (.040, .101)
OF	4.31	117.37	19	.93	.047	.148 (.123, .174)
BF	4.31	20.23	11	.99	.018	.060 (.011, .100)
SS	4.31	21.57	15	1.0	.019	.043 (.000, .081)

Note: CF = correlated factors; OF = one-factor; BF = bifactor; SS = subset

As can be seen in Table 9, the one factor (OF) model had poor fit while the correlated factors model (CF) had adequate fit. These results suggested that sign learning and word learning are not completely independent factors nor are they fully determined by a single general lexical factor. Another possibility was that that a single factor contributed to individual differences on all lexical learning tasks but that specific factors *also* account for variance in performance. To test this possibility, a bifactor model was designated by loading all tasks onto a single lexical learning factor and loading sign learning and word learning tasks onto respective specific factors.

The bifactor (BF) model demonstrated good fit and fit the data better than the correlated factors model (See Table 9), however, an inspection of the path coefficients revealed a misspecification in the model. Specifically, the path coefficients between the word learning tasks and the specific word learning factor were insignificant or, in the case of the TWL task, ludicrously large and *negative*. What this suggested was that the specific word learning factor was unnecessary.

In the final model assessed, all lexical learning tasks were loaded onto a general factor and only the sign learning tasks were loaded onto a specific factor—because the tasks defining the sign learning factor are a subset of the tasks defining the general lexical

learning factor, this model was labeled Model SS (for subset). The model fit the data well (see Table 9) and was significantly better than the correlated factors model, $\Delta\chi^2(3) = 17.69, p < .001$.

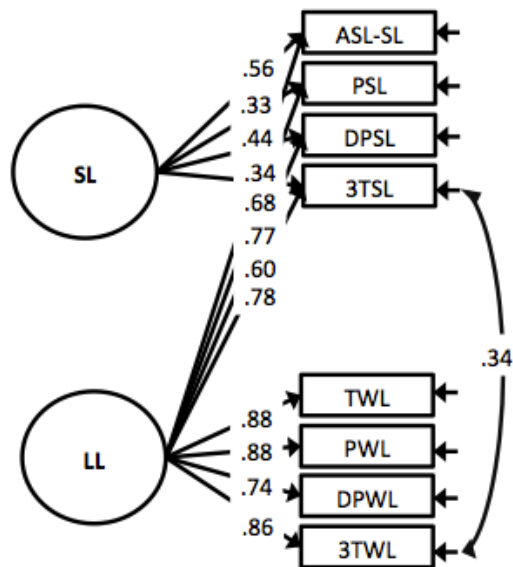


Figure 14. Model explaining the relationship between word learning and sign learning. Note: SL = sign learning, LL = lexical learning

CHAPTER 4. DISCUSSION

This study had two goals: first, to extend research on word learning to sign learning and second, to examine the relationship between these two constructs. Overall, the results of this study indicate that word learning and sign learning rely on similar processes, which can be partially accounted for by Gf and modality-specific PSTM. Accordingly, word learning and sign learning were highly related. An examination of this relationship revealed that individual differences on these variables could be accounted for by a general lexical learning factor and a specific sign learning factor.

4.1 Crystallized and Fluid Intelligence and Relation-Construction

It was expected that Gc and Gf would be predictive of both sign learning and word learning because, regardless of the language modality, participants can make use of the relation-construction principle which states that associative learning is partly determined by the quality and quantity of relationships between what is known and that which is to be learned (Kyllonen et al., 1991). It was presumed that, due to greater knowledge, high Gc individuals would be more likely to produce a greater number of high quality relationships; Gf, on the other hand, would come into play in the inducing of appropriate relationships and in discarding inappropriate ones. Indeed, previous studies have found that Gc and Gf are related to word learning (e.g., Hundal & Horn, 1977; Kyllonen & Tirre, 1988).

In this study, only Gf accounted for a significant proportion of variance in sign learning and word learning. Assuming that Gc and Gf are generally involved in the

construction of relations, then it is possible that the lack of an independent relationship between Gc and the lexical learning variables was due to item characteristics, presentation order, and/or the amount of time given to study items.

Consider the following two pairs of words: *electricity-banana* and *muz-banana*. Before continuing, it may be instructive for the reader to attempt to construct relationships between each pair of words. For the first pair, it should be fairly easy to generate associations for *both* words and to identify relationships. For example, the following quickly comes to the author's mind when thinking about electricity: yellow, the symbol for a lightning bolt, Thomas Edison (the namesake of an Electrical company), the light bulb, and Benjamin Franklin. For banana: yellow, mushy, fruit, breakfast, mealy, and republic. From here, relationships can be formed linking electricity and banana, perhaps as an image of Thomas Edison holding a glowing yellow banana (as if it were a light bulb) or as a sentence: "Thomas Edison loved mushy bananas." The Turkish word for banana, *muz*, however is unlikely to conjure up any associations independent of those that relate it to banana. So, for example, one may feel that *muz* sounds similar to *mushy*.

For the first pair of words, *electricity-banana*, one drew upon Gc to generate associates for both terms; for the second pair, the role of Gc was limited to the familiar English word. The wealth of information present in the first case can facilitate the construction of a number of unique relationships that link *electricity* and *banana* together. Though it is certainly possible to generate more relationships between *muz* and *banana*, it is likely that, all other things being equal, the quantity and quality of relationships that can be generated between a familiar word and a highly unfamiliar lexical form will be less than that which can be generated for two familiar words.

The role that Gc played in this study was likely further diminished by the fact that the unfamiliar lexical form was always presented first and the familiar word was presented second and only for 1000-2000ms. By presenting the unfamiliar lexical form first, participants were severely limited in the associations they could generate until they saw the familiar word. Once they saw the word, they only had a brief amount of time to attempt to form a relationship. This likely limited the amount of information that participants could draw from Gc and forced participants to quickly form a relationship with whatever came to mind first.

In fact, it may be that processing speed acted as a suppressor variable, meaning that had processing speed been accounted for, Gc would have been a significant predictor of lexical learning. To explain, numerous studies have found that processing speed is related to associative learning (e.g., Kyllonen & Tirre, 1988; Kyllonen et al., 1991; Park et al., 1996; Salthouse, 1994; Salthouse & Dunlosky, 1995). In particular, Kyllonen et al. (1991) observed that when study time was brief (500ms), fast processors tended to outperform slow processors independent of their verbal knowledge (a marker of Gc). Moreover, as study time increased (up to 8000ms), the effect of verbal knowledge on lexical learning tended to increase while the effect of processing speed attenuated. Kyllonen and colleagues interpreted these results as indicating that, when study time was brief, fast processors were able to produce a greater number of relations compared to slow processors, however, as study time increased, individuals with high verbal ability were able to use the time to continue elaborating while individuals with low verbal ability were unable to do so.

4.2 Working Memory Capacity and Phonological Short-Term Memory

It was expected that WMC and the domain-specific PSTM factors would make independent contributions to the prediction of lexical learning. According to theory, WMC and STM tasks assess executive attention, however, WMC is a domain-general construct that assesses executive attention to a greater degree than STM and STM additionally draws on domain-specific perceptual and motor processes (Kane et al., 2004). Thus it was presumed that the variance in lexical learning accounted for by WMC would be largely due to differences in executive attention while the PSTM factors would account for variance due to domain-specific processes.

The results were largely in accord with the above, with WMC and the PSTM factors all being strongly correlated with the lexical learning factors; however, when signed- and spoken-PSTM were included as predictors alongside WMC, the direct paths between WMC and the lexical learning factors were no longer significant (compare Figures 11-13). It appears that WMC that did not account for a significant proportion of variance in either sign learning or word learning above that which was accounted for by the PSTM factors.

This result may be due to 1) the use of shortened WM tasks which decreased the amount of variance accounted for, 2) the PSTM tasks used here relying more heavily on executive attention than typical (and Gf; see the next section), or 3) a combination of the above. To elaborate on the second point, the PSTM tasks in this study utilized highly unfamiliar material, which likely limited the supportive role that long-term memory plays in STM tasks, subsequently increasing the role of executive attention (Reder, Liu,

Keinath, & Popov, 2016; Shen, Popov, Delahay, & Reder, 2018). When STM tasks utilize familiar material, such as digits or words, individuals are able to use long-term memory to aid STM. For example, Woodward, Macken, and Jones (2008) found that as participants gained experience with producing sequences of pseudowords, their fluency also improved and this, they contended, largely explained gains in pseudoword span. More commonly, long-term memory is said to aid STM through the process of redintegration, referring to the process by which partial or degraded representations can serve as cues (Bower & Glass, 1976; Gathercole et al., 2001; Hulme, Maughan, & Brown, 1991; Hulme et al., 1997; Thorn, Gathercole, & Frankish, 2005). Without long-term memory to effectively support STM through efficient rehearsal processes or redintegration, memory traces for unfamiliar material are especially fragile and susceptible to competition from more entrenched representations—in this context, any shift in attention can be catastrophic, implicating executive attention. The fact that the WMC factor was composed of the variance in performance on *shortened* tasks and yet still exhibited correlations with the PSTM factors that were on par with what has been seen when full-length tasks are used offers some support for the above. In the future, it would be worthwhile to investigate relationships amongst tasks that vary in the familiarity of the stimuli, the use of secondary tasks (like complex span tasks), and tasks that inhibit motor rehearsal.

4.3 Phonological Short-Term Memory: The Specific and the General

As expected, the full model (Figure 13) revealed that after accounting for other relevant variables, PSTM was only predictive of lexical learning within modality, indicating a certain degree of domain-specificity. This outcome supports Gathercole's

(2006) theory that PSTM is related to lexical learning in part because of similarities in perceptual and motor processes.

The PSTM constructs also revealed a significant degree of domain-generalty: the PSTM factors were strongly related to WMC and apparently accounted for the same portion of variance in lexical learning accounted for by WMC (see discussion above). Interestingly, of the two PSTM factors, signed-PSTM was the most general, exhibiting slightly larger correlations with most other factors and a substantially larger correlation with the Gf and Gf_{Res} factors. This is inline with prior work which has typically demonstrated that, compared to auditory-verbal abilities, visuospatial abilities tend to exhibit greater correlations with general mental ability and Gf (Groeger, Field, & Hammond, 1999; Kane et al., 2004; Lohman, 1996; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). Theories have been put forth to address the relationship between Gf and performance on complex visuospatial tasks (e.g., the SLAT used in this study; see Lohman, 1996) but, to my knowledge, a theory has not been put forth to explain the relationship between Gf and comparatively simple visuospatial memory tasks. As with PSTM and WMC, familiarity may play a role.

Adult non-signers are quite adept at using speech-motor processes to aid in rehearsing auditory-verbal information; they are, however, unlikely to be skilled in rehearsing signs. Participants' experience with articulatory rehearsal may have biased them towards using this strategy during spoken-PSTM tasks and, as such, there were likely few individual differences due to strategy use. The novelty of signed-PSTM tasks, however, may have spurred variation in strategy use and it may be these differences that explain why signed-PSTM was more highly correlated with Gf and, in particular, with

the residualized factor, Gf_{Res} . In accordance with the idea that Gf_{Res} reflects individuals' ability to disengage from outdated information, individuals with greater Gf_{Res} may have optimized their performance by testing and discarding a number of strategies (Schunn & Reder, 2001). In fact, anecdotally, it was observed that some participants overtly rehearsed signs throughout the study, others initiated overt rehearsal at some point during the series of signed-PSTM tasks, and some of those participants who used overt rehearsal seemingly abandoned the strategy during or between signed-PSTM tasks.

There are of course other possible explanations for the relationship between signed-PSTM and Gf_{Res} . For example, recently Shipstead et al. (2016) and Engle (2018) have theorized that Gf or, more specifically, the ability to disengage is related to the ability to mitigate proactive interference—although attempts were made to reduce item similarity, the sign stimuli may not have been visually distinct enough. Clearly, more research is needed to investigate whether the magnitude of the correlations exhibited between the PSTM factors and Gf generalize and, if so, the cause.

One other thing to note is that spoken-PSTM was predictive of sign learning even after WMC was accounted for (Figure 12) and, although not shown, signed-PSTM was also predictive of word learning even when WMC was accounted for. In other words, even after accounting for executive attention, lexical learning and PSTM factors that bore no obvious perceptual or motor similarities were still related. One possibility is that the lexical learning and PSTM factors also enlisted amodal phonological processes (Baddeley, 2015; Baddeley et al., 1998). This is an interesting possibility but it is still necessary to explain why imaging studies have not found activity in classic language processing areas when non-signers are exposed to signs (Newman-Norlund et al., 2006; J.

T. Williams et al., 2016b). Moreover, this study was not intended to address this issue, thus with regard to amodal processing, the results are equivocal if beguiling.

4.4 Lexical Learning: The General and the Partially Specific

The second aim of this study was to investigate the relationship between sign learning and word learning. The two constructs were found to be highly correlated but not identical. Further analysis revealed that all lexical learning tasks loaded onto a general factor, however, sign learning tasks loaded onto an additional specific factor. What these results suggest is that in hearing non-signers, all lexical learning tasks rely on similar processes; however, sign learning tasks make additional demands. The analyses conducted in this study reveal some of those common processes (e.g., Gf, executive attention) but, as indicated by the correlated disturbance terms in Figure 13, a substantial proportion of common variance remains unexplained. Processing speed was already proposed as one possible mediating factor but others are likely.

Regarding the specific factor, it is not clear why it is needed. Evidently, part of what is being captured by this specific factor is domain-specific motor and perceptual processes however this is not enough to explain why sign learning merited its own specific factor but word learning did not. One may counter that non-signers used spoken language processing during sign learning, after all, signs were paired with spoken words however, based on this line of thinking, it would follow that spoken-PSTM should have been predictive of sign learning even after accounting for signed-PSTM, but it was not. A similar argument can be made for individuals using visual imagery during lexical learning, regardless of the modality, but it too is refuted by the fact that signed-PSTM

was not significantly predictive of word learning after accounting for spoken-PSTM. At present, all that can be said is that sign learning appears to be more complex than word learning, relying on largely the same processes as word learning with additional demands captured by a specific factor.

Future studies should investigate the conditions that affect the relationship between sign learning and word learning. For example, in this study, participants were instructed to use elaborative constructions to support learning and this likely increased the role of Gf and, consequently, the correlation between sign learning and word learning. In contrast, it is likely that instructing participants to use rote rehearsal would increase the role of domain-specific processes and therefore lower the relationship between sign learning and word learning. These constructs can also be investigated in individuals with varying degrees of experience with a signed language (e.g., college students enrolled in their first semester of a sign language course compared to those enrolled in their third semester). On the one hand, increased experience with signs will enable individuals to use effective rehearsal strategies which should increase domain-specific factors and therefore reduce the correlation between sign learning and word learning; on the other hand, experience may result in participants processing the signs linguistically (Newman-Norlund et al., 2006; J. T. Williams et al., 2016b), just as they do words, possibly increasing domain-generality and the correlation between sign learning and word learning. Quasi-experiments like these in conjunction with experimental, imaging, and computational studies can aid in explicating the relationship observed in this study.

4.5 Summary and Implications

The results of this study corroborate and extend prior research on lexical learning in spoken languages. Specifically, it was found that Gf and modality-specific PSTM were predictive of sign learning just as they have been found to be predictive of word learning (Baddeley et al., 1998). Interestingly, two predictors that were assumed to be important to lexical learning, Gc and WMC, were insignificant in this study. It is suggested that the effect of Gc may have been suppressed by the brief study period used here; WMC on the other hand, was likely accounted for by the PSTM and was therefore redundant. It was also shown that word learning and sign learning are highly correlated but partially distinct. Subsequent analyses revealed that all tasks loaded onto a general lexical learning factor but sign learning tasks additionally loaded onto a specific factor. As such, this study provides insight into the cognitive components that are common to lexical learning regardless of language modality and those that are unique to either signed or spoken languages. Future studies should continue to investigate the relationship between word learning and sign learning.

Importantly, this is the first study to investigate cognitive predictors of sign learning and, because lexical learning is related to other language learning measures (Cooper, 1964; Li, 2015; K. I. Martin & Ellis, 2012), possibly L2 learning in general. Researchers should investigate whether the predictors identified here are predictive of sign learning in the classroom and whether there is a relationship between these measures, sign learning, and mastery of a signed language. If so, then this research and future research can help guide and place students into appropriate second language courses. Moreover, many of the results were explained by recourse to familiarity,

suggesting that sign language instructors should spend a significant amount of time familiarizing their students with the phonology and phonotactics of the particular sign language (e.g., ASL).

**APPENDIX A. LIST OF ENGLISH WORDS USED IN THIS
STUDY**

ASL	PSL	3TSL	TWL	PWL	3TWL
actor	blame	baby	anvil	chair	ball
alarm	brief	bone	banana	dance	boat
apple	chart	book	birth	drama	bulb
color	fence	cane	chest	field	cage
earth	grade	drum	crown	money	coin
email	honey	fork	drawer	paper	dish
fraud	magic	hood	hairy	phone	hair
game	music	ipad	hinge	plane	knot
gold	skirt	lamp	jeans	pound	lard
grass	spite	lime	room	rifle	leaf
hotel	title	mask	skies	sauce	lint
metal	treat	meat	steel	shell	pail
month	trust	oven	storm	shirt	rain
music	value	rose	viper	taste	sand
penny	waste	sled	wrist	wheel	suit
potato		soda			tape
socks		tree			text
steak		xbox			wart
sushi					
toast					
train					
trash					
waiter					
woman					

REFERENCES

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working Memory and Intelligence: The Same or Different Constructs? *Psychological Bulletin*, *131*(1), 30-60. doi:10.1037/0033-2909.131.1.30
- Ackerman, P. L., & Cianciolo, A. T. (2000). Cognitive, perceptual-speed, and psychomotor determinants of individual differences during skill acquisition. *Journal of Experimental Psychology: Applied*, *6*(4), 259-290. doi:10.1037/1076-898X.6.4.259
- Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *The Quarterly Journal of Experimental Psychology*, *18*(4), 362-365. doi:10.1080/14640746608400055
- Baddeley, A. D. (2015). Working Memory in Second Language Learning. In Z. Wen, M. B. Mota, A. McNeill, M. Bunting, & R. Engle (Eds.), *Working Memory in Second Language Acquisition and Processing* (pp. 17-28). Bristol, England: Multilingual Matters.
- Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158-173. doi:10.1037/0033-295X.105.1.158
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, *14*(6), 575-589. doi:10.1016/S0022-5371(75)80045-4
- Bavelier, D., Corina, D., Jezzard, P., Clark, V., Karni, A., Lalwani, A., . . . Neville, H. J. (1998). Hemispheric specialization for English and ASL: Left invariance-right variability. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*, *9*(7), 1537-1542. doi:10.1097/00001756-199805110-00054
- Bentler, P. M. (2001). *EQS 6 structural equations program manual*. Encino, CA: Multivariate software.
- Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-spatial imagery: a new self-report imagery questionnaire. *Applied Cognitive Psychology*, *20*(2), 239-263. doi:10.1002/acp.1182
- Bower, G. H., & Glass, A. L. (1976). Structural units and the redintegrative power of picture fragments. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(4), 456-466. doi:10.1037/0278-7393.2.4.456

- Bower, G. H., & Winzenz, D. (1970). Comparison of associative learning strategies. *Psychonomic Science*, 20(2), 119-120. doi:10.3758/BF03335632
- Bradley, M. M., & Glenberg, A. M. (1983). Strengthening associations: Duration, attention, or relations? *Journal of Verbal Learning & Verbal Behavior*, 22(6), 650-666. doi:10.1016/S0022-5371(83)90385-7
- Brentari, D. (1998). *A prosodic model of sign language phonology*. Massachusetts: MIT Press.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, 41(4), 977-990.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, 46(3), 904-911.
- Cantor, J., Engle, R. W., & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence*, 15(2), 229-246. doi:10.1016/0160-2896(91)90032-9
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97(3), 404.
- Caselli, N. K., Sehyr, Z. S., Cohen-Goldberg, A. M., & Emmorey, K. (2016). ASL-LEX: A lexical database of American Sign Language. *Behavior Research Methods*, 1-18. doi:10.3758/s13428-016-0742-0
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40(3), 153-193. doi:10.1037/h0059973
- Colom, R., Shih, P. C., Flores-Mendoza, C., & Quiroga, M. Á. (2006). The real relationship between short-term memory and working memory. *Memory*, 14(7), 804-813. doi:10.1080/09658210600680020
- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, 55(4), 429-432. doi:10.1111/j.2044-8295.1964.tb00928.x
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786.

- Cooper, C. J. (1964). Some relationships between paired-associates learning and foreign-language aptitude. *Journal of Educational Psychology*, 55(3), 132-138. doi:10.1037/h0044223
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in brain research*, 169, 323-338. doi:10.1016/S0079-6123(07)00020-9
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3(4), 422-433. doi:10.3758/bf03214546
- Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2018). *The methodological problems of reaction time in differential research: A review*. Manuscript in preparation.
- Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology*, 77(5), 615-622. doi:10.1037/0021-9010.77.5.615
- Ekstrom, R., French, J., Harman, H., & Dermen, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, 61(2), 367-413. doi:10.1111/j.1467-9922.2010.00613.x
- Embretson, S. E. (1992). Measuring and Validating Cognitive Modifiability as an Ability: A Study in the Spatial Domain. *Journal of Educational Measurement*, 29(1), 25-50.
- Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, 11(1), 19-23. doi:10.1111/1467-8721.00160
- Engle, R. W. (2018). Working Memory and Executive Attention: A Revisit. *Perspectives on Psychological Science*, 13(2), 190-193. doi:10.1177/1745691617720478
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309-331. doi:10.1037/0096-3445.128.3.309
- Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, 43(2), 226-236. doi:10.3758/s13421-014-0461-7
- Gardner, R. C., & Lambert, W. E. (1965). Language aptitude, intelligence, and second-language achievement. *Journal of Educational Psychology*, 56(4), 191-199. doi:10.1037/h0022400

- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27(4), 513-543. doi:10.1017/S0142716406060383
- Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, 29(3), 336-360. doi:10.1016/0749-596X(90)90004-J
- Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 54A(1), 1-30. doi:10.1080/02724980042000002
- Glaze, J. A. (1928). The association value of non-sense syllables. *The Pedagogical Seminary and Journal of Genetic Psychology*, 35, 255-269. doi:10.1080/08856559.1928.10532156
- Gómez, M. J. L., Molina, T. B., Benítez, P. P., & de Torres, J. S. (2007). Predicting proficiency in signed language interpreting: A preliminary study. *Interpreting*, 9(1), 71-93.
- Groeger, J. A., Field, D., & Hammond, S. M. (1999). Measuring Memory Span. *International Journal of Psychology*, 34(5-6), 359-363. doi:10.1080/002075999399693
- Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 56A(7), 1213-1236. doi:10.1080/02724980343000071
- Hambrick, D. (2003). Why are some people more knowledgeable than others? A longitudinal study of knowledge acquisition. *Memory & Cognition*, 31(6), 902-917. doi:10.3758/BF03196444
- Hasher, L., & Zacks, R. T. (1988). Working Memory, Comprehension, and Aging: A Review and a New View. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 22, pp. 193-225): Academic Press.
- Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher, A. Koriat, D. Gopher, & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application*. (pp. 653-675). Cambridge, MA, US: The MIT Press.
- Hertzog, C., Price, J., & Dunlosky, J. (2012). Age Differences in the Effects of Experimenter-Instructed Versus Self-Generated Strategy Use. *Experimental aging research*, 38(1), 42-62. doi:10.1080/0361073X.2012.637005

- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for Familiar and Unfamiliar Words: Evidence for a Long-Term Memory Contribution to Short-Term Memory Span. *Journal of Memory and Language*, 30(6), 685-701.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(5), 1217-1232. doi:10.1037/0278-7393.23.5.1217
- Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty and M. H. Long (Eds.). *The handbook of second language acquisition* (pp. 349-381), Malden, MA: Blackwell.
- Hummel, K. M., & French, L. M. (2016). Phonological memory and aptitude components: Contributions to second language proficiency. *Learning and Individual Differences*, 51(Supplement C), 249-255. doi:https://doi.org/10.1016/j.lindif.2016.08.016
- Hundal, P. S., & Horn, J. L. (1977). On the relationships between short-term learning and fluid and crystallized intelligence. *Applied Psychological Measurement*, 1(1), 11-21. doi:10.1177/014662167700100104
- Jackson, D. N. (1998). *Multidimensional Aptitude Battery-II*. London, Ontario, Canada: Sigma Assessment Systems.
- Jenkins, J. J. (1985). Nonsense syllables: Comprehending the 'almost incomprehensible variation'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(3), 455-460. doi:10.1037/0278-7393.11.3.455
- Jones, D. M., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 1008-1018. doi:10.1037/0278-7393.21.4.1008
- Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and Language*, 54(2), 265-281. doi:10.1016/j.jml.2005.10.006
- Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(2), 336-358. doi:10.1037/0278-7393.26.2.336
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span and Reasoning. *Journal of*

Experimental Psychology: General, 133(2), 189-217. doi:10.1037/0096-3445.133.2.189

- Kaufman, S. B., DeYoung, C. G., Gray, J. R., Brown, J., & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*, 37(4), 374-382. doi:10.1016/j.intell.2009.03.004
- Kline, R. B. (2016). *Principles and practice of structural equation modeling (4th ed.)*. New York, NY, US: Guilford Press.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal*, 73(4), 440-464. doi:10.2307/326879
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14(4), 389-433. doi:10.1016/S0160-2896(05)80012-1
- Kyllonen, P. C., & Tirre, W. C. (1988). Individual differences in associative learning and forgetting. *Intelligence*, 12(4), 393-421. doi:10.1016/0160-2896(88)90004-9
- Kyllonen, P. C., Tirre, W. C., & Christal, R. E. (1991). Knowledge and processing speed as determinants of associative learning. *Journal of Experimental Psychology: General*, 120(1), 57-79. doi:10.1037/0096-3445.120.1.57
- Kyllonen, P. C., & Woltz, D. J. (1989). Role of cognitive factors in the acquisition of cognitive skill. In R. Kanfer, R. Cudeck, R. Kanfer, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on Learning and Individual Differences*. (pp. 239-280). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- La Pointe, L. B., & Engle, R. W. (1990). Simple and Complex Word Spans as Measures of Working Memory Capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(6), 1118-1133. doi:10.1037/0278-7393.16.6.1118
- Li, S. (2015). The construct validity of language aptitude. *Studies in Second Language Acquisition*, 1-42. doi:10.1017/S027226311500042X
- Little, R. J. A., & Rubin, D. B. (2014). *Statistical analysis with missing data (2nd ed., ed.)*: Hoboken, New Jersey : John Wiley & Sons, Inc.
- Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis*: Lawrence Erlbaum Associates Publishers.
- Lohman, D. F. (1996). Spatial ability and g. In I. Dennis, P. Tapsfield, I. Dennis, & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement*. (pp. 97-116). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

- Mann, W., Marshall, C., Mason, K., & Morgan, G. (2010). The acquisition of sign language: The impact of phonetic complexity on phonology. *Language Learning and Development, 6*(1), 60-86. doi:10.1080/15475440903245951
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*(3), 519-530.
- Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence, 7*(2), 107-127. doi:10.1016/0160-2896(83)90023-5
- Marshall, C., Mason, K., Rowley, K., Herman, R., Atkinson, J., Woll, B., & Morgan, G. (2015). Sentence repetition in deaf children with specific language impairment in British Sign Language. *Language Learning and Development, 11*(3), 237-251. doi:10.1080/15475441.2014.917557
- Martin, J. D., Shipstead, Z., Harrison, T., Redick, T. S., Bunting, M., & Engle, R. W. (2017). *The role of maintenance and disengagement in predicting reading comprehension and vocabulary learning*. Manuscript submitted for publication.
- Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*(3), 379-413. doi:10.1017/S0272263112000125
- Martinez, D., & Singleton, J. L. (2018). Predicting sign learning in hearing adults: The role of perceptual-motor (and phonological?) processes. *Applied Psycholinguistics, 1*-27. doi:10.1017/S0142716418000048
- Mason, K., Rowley, K., Marshall, C. R., Atkinson, J. R., Herman, R., Woll, B., & Morgan, G. (2010). Identifying specific language impairment in deaf children acquiring British Sign Language: Implications for theory and practice. *British Journal of Developmental Psychology, 28*(1), 33-49. doi:10.1348/026151009X484190
- Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130*(4), 621-640. doi:10.1037/0096-3445.130.4.621
- Nairne, J. S. (1983). Associative processing during rote rehearsal. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(1), 3-20. doi:10.1037/0278-7393.9.1.3
- Nation, I. S. P. (1980). Strategies for receptive vocabulary learning. *Guidelines, 3*(1), 23.
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics, 28*(04), 661-677.

- Newman-Norlund, R. D., Frey, S. H., Petitto, L.-A., & Grafton, S. T. (2006). Anatomical Substrates of Visual and Auditory Miniature Second-language Learning. *Journal Of Cognitive Neuroscience*, 18(12), 1984-1997.
- Noble, C. E. (1952). An analysis of meaning. *Psychological Review*, 59(6), 421-430. doi:10.1037/h0054087
- O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, 27(3), 377-402. doi:10.1017/S0142716406060322
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, 29(4), 557-582. doi:10.1017/S027226310707043X
- Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, 142(7), 758-799. doi:10.1037/bul0000046
- Oberauer, K., Lewandowsky, S., Avh, E., Brown, G. D., Conway, A., Covan, N., . . . Hurlstone, M. (2018). Benchmarks for models of short term and working memory. *Psychological Bulletin*.
- Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., & Gaines, C. L. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging*, 11(4), 621-637. doi:10.1037/0882-7974.11.4.621
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Advanced Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.
- Reder, L. M., Liu, X. L., Keinath, A., & Popov, V. (2016). Building knowledge requires bricks, not sand: The critical role of familiar constituents in learning. *Psychonomic Bulletin & Review*, 23(1), 271-277. doi:10.3758/s13423-015-0889-1
- Rohrer, D., Taylor, K., Pashler, H., Wixted, J. T., & Cepeda, N. J. (2005). The Effect of Overlearning on Long-Term Retention. *Applied Cognitive Psychology*, 19(3), 361-374. doi:10.1002/acp.1083
- Rubin, D. B. (1976). Inference and Missing Data. *Biometrika*, 63(3), 581-592. doi:10.2307/2335739
- Salthouse, T. A. (1994). Aging associations: Influence of speed on adult age differences in associative learning. *Journal of Experimental Psychology*, 20(6), 1486.
- Salthouse, T. A., & Dunlosky, J. (1995). Analyses of adult age differences in associative learning. *Zeitschrift Für Psychologie Mit Zeitschrift Für Angewandte Psychologie*, 203(4), 351-360.

- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. . Pittsburgh, PA: Psychology Software Tools Inc.
- Schunn, C. D., & Reder, L. M. (2001). Another source of individual differences: Strategy adaptivity to changing rates of success. *Journal of Experimental Psychology: General*, *130*(1), 59-76. doi:10.1037/0096-3445.130.1.59
- Seibert, L. C. (1930). An experiment on the relative efficiency of studying French vocabulary in associated pairs versus studying French vocabulary in context. *Journal of Educational Psychology*, *21*(4), 297-314. doi:10.1037/h0070517
- Shaughnessy, J. J. (1981). Memory monitoring accuracy and modification of rehearsal strategies. *Journal of Verbal Learning and Verbal Behavior*, *20*(2), 216-230. doi:https://doi.org/10.1016/S0022-5371(81)90389-3
- Shaw, S. (2011). Cognitive and motivational contributors to aptitude: A study of spoken and signed language interpreting students. *Interpreting*, *13*(1), 70-84. doi:10.1075/intp.13.1.05sha
- Shen, Z., Popov, V., Delahay, A. B., & Reder, L. M. (2018). Item strength affects working memory capacity. *Memory & Cognition*, *46*(2), 204-215. doi:10.3758/s13421-017-0758-4
- Shipstead, Z., & Engle, R. W. (2013). Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 277-289. doi:10.1037/a0028467
- Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, *11*(6), 771-799. doi:doi:10.1177/1745691616650647
- Söderfeldt, B., Ingvar, M., Rönnerberg, J., Eriksson, L., Serrander, M., & Stone-Elander, S. (1997). Signed and spoken language perception studied by positron emission tomography. *Neurology*, *49*(1), 82-87. doi:10.1159/000053093
- Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371-394. doi:10.1037/0033-2909.121.3.371
- Tamez, E., Myerson, J., & Hale, S. (2008). Learning, working memory, and intelligence revisited. *Behavioural Processes*, *78*(2), 240-245. doi:10.1016/j.beproc.2008.01.008
- Thorn, A. S. C., Gathercole, S. E., & Frankish, C. R. (2005). Redintegration and the benefits of long-term knowledge in verbal short-term memory: An evaluation of Schweickert's (1993) multinomial processing tree model. *Cognitive Psychology*, *50*(2), 133-158. doi:http://dx.doi.org/10.1016/j.cogpsych.2004.07.001

- Thorndike, E. L. (1908). Memory for paired associates. *Psychological Review*, 15(2), 122-138. doi:10.1037/h0073570
- Thurstone, L. L. (1938). *Primary Mental Abilities*: Psychometric Monographs.
- Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin & Review*, 13(3), 452-457. doi:10.3758/BF03193869
- Ullman, J. B. (2006). Structural Equation Modeling: Reviewing the Basics and Moving Forward. *Journal of Personality Assessment*, 87(1), 35-50. doi:10.1207/s15327752jpa8701_03
- Underwood, B. J., Boruch, R. F., & Malmi, R. A. (1978). Composition of episodic memory. *Journal of Experimental Psychology: General*, 107(4), 393-419. doi:10.1037/0096-3445.107.4.393
- Unsworth, N. (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable approach. *Acta Psychologica*, 134(1), 16-28. doi:https://doi.org/10.1016/j.actpsy.2009.11.010
- Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: The role of controlled search. *Memory & Cognition*, 41(2), 242-254. doi:10.3758/s13421-012-0261-x
- Unsworth, N., & Engle, R. W. (2007). On the Division of Short-Term and Working Memory: An Examination of Simple and Complex Span and Their Relation to Higher Order Abilities. *Psychological Bulletin*, 133(6), 1038-1066. doi:10.1037/0033-2909.133.6.1038
- Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility. *The Quarterly Journal of Experimental Psychology*, 65(12), 2397-2410. doi:10.1080/17470218.2012.690438
- Varriale, V., van der Molen, M. W., & De Pascalis, V. (2018). Mental rotation and fluid intelligence: A brain potential analysis. *Intelligence*, 69, 146-157. doi:https://doi.org/10.1016/j.intell.2018.05.007
- Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology*, 141, 65-82. doi:10.1016/j.jecp.2015.06.015
- Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory: Functional equivalence across modalities. *Memory*, 13(3-4), 308-317. doi:10.1080/09658210344000279

- Williams, B. A., & Pearlberg, S. L. (2006). Learning of three-term contingencies correlates with Raven scores, but not with measures of cognitive processing. *Intelligence*, 34(2), 177-191. doi:10.1016/j.intell.2005.03.007
- Williams, J. T., Darcy, I., & Newman, S. D. (2015). Modality-independent neural mechanisms for novel phonetic processing. *Brain Research*, 1620, 107-115. doi:10.1016/j.brainres.2015.05.014
- Williams, J. T., Darcy, I., & Newman, S. D. (2016a). The beneficial role of L1 spoken language skills on initial L2 sign language learning: Cognitive and linguistic predictors of M2L2 acquisition. *Studies in Second Language Acquisition*, 1-18. doi:10.1017/S0272263116000322
- Williams, J. T., Darcy, I., & Newman, S. D. (2016b). Modality-specific processing precedes amodal linguistic processing during L2 sign language acquisition: A longitudinal study. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 75, 56-67. doi:10.1016/j.cortex.2015.11.015
- Wilson, M., & Emmorey, K. (1997). A visuospatial 'phonological loop' in working memory: Evidence from American Sign Language. *Memory & Cognition*, 25(3), 313-320. doi:10.3758/BF03211287
- Woodward, A. J., Macken, W. J., & Jones, D. M. (2008). Linguistic familiarity in short-term memory: A role for (co-)articulatory fluency? *Journal of Memory and Language*, 58(1), 48-65. doi:10.1016/j.jml.2007.07.002
- Wu, Y. C., & Coulson, S. (2014). A psychometric measure of working memory capacity for configured body movement. *PLOS ONE*, 9(1).
- Yang, L. (1997). Tracking the acquisition of L2 vocabulary: The Keki language experiment. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 125-156). Cambridge: Cambridge University Press.