Review

Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance

Jarrad A.G. Lum [a], Gina Conti-Ramsden [b,], Angela T. Morgan [c] and Michael T. Ullman [d]*

[a] *School of Psychology, Deakin University, Melbourne, Australia*
[b] *School of Psychological Sciences, The University of Manchester, Manchester, UK*
[c] *Murdoch Children's Research Institute, Melbourne, Australia*
[d] *Department of Neuroscience, Georgetown University, Washington, D.C., USA*

ARTICLE INFO

ABSTRACT

Meta-analysis and meta-regression were used to evaluate whether evidence to date demonstrates deficits in procedural memory in individuals with specific language impairment (SLI), and to examine reasons for inconsistencies of findings across studies. The Procedural Deficit Hypothesis (PDH) proposes that SLI is largely explained by abnormal functioning of the frontal-basal ganglia circuits that support procedural memory. It has also been suggested that declarative memory can compensate for at least some of the problems observed in individuals with SLI. A number of studies have used Serial Reaction Time (SRT) tasks to investigate procedural learning in SLI. In this report, results from eight studies that collectively examined 186 participants with SLI and 203 typically-developing peers were submitted to a meta-analysis. The average mean effect size was .328 (CI_{95}: .071, .584) and was significant. This suggests SLI is associated with impairments of procedural learning as measured by the SRT task. Differences among individual study effect sizes, examined with meta-regression, indicated that smaller effect sizes were found in studies with older participants, and in studies that had a larger number of trials on the SRT task. The contributions of age and SRT task characteristics to learning are discussed with respect to impaired and compensatory neural mechanisms in SLI.

© 2013 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license.

* *Corresponding author.* Human Communication and Deafness (HCD), School of Psychological Sciences, The University of Manchester, Ellen Wilkinson Building, Oxford Road, Manchester M13 9PL, UK.
E-mail address: gina.conti-ramsden@manchester.ac.uk (G. Conti-Ramsden).

1. Introduction

Specific language impairment (SLI) is a neurodevelopmental disorder characterized by impaired or delayed language skills that occur in the absence of intellectual, sensory or medical problems (American Psychiatric Association, 2000; World Health Organization, 1996). Substantial research suggests an association between SLI and a range of cognitive and motor impairments (for reviews see Hill, 2001; Leonard, 2000; Ullman & Pierpont, 2005). In some cases, such non-language problems are thought to either cause or exacerbate the difficulties that affected individuals have in understanding and using language (e.g., Gathercole & Baddeley, 1990; Montgomery, Magimairaj, & Finney, 2010; Tallal, 2004).

The Procedural Deficit Hypothesis (PDH), proposed by Ullman and Pierpont (2005), holds that a number of the language difficulties in SLI, in particular the grammatical deficits, may be largely explained by procedural memory impairments. The procedural memory system underlies the implicit learning and representation of skills and knowledge, as well as their automatic and rapid execution (Gabrieli, 1998; Ullman, 2004). The learning and memory functions of the system are said to be implicit because they do not require awareness. Learning via the procedural memory system is often slow, with substantial repetition or practice required in order for skills or knowledge to be processed rapidly and automatically. According to Ullman and Pierpont (2005), the procedural memory impairments in SLI are likely to be caused by neural abnormalities of one or more structures that underlie the procedural memory system, in particular the basal ganglia and frontal cortex, especially the caudate nucleus and Broca's region.

Ullman and colleagues (Ullman & Pierpont, 2005; Ullman and Pullman, submitted for publication) further suggest that the presence or severity of cognitive and language impairments in SLI will depend not only on procedural memory deficits but also on the extent to which declarative memory, which is proposed to remain largely intact in SLI, can compensate for the procedural deficits. Thus, in principle, if declarative memory could fully compensate for such underlying problems, impairments in procedural memory might not be evident.

Despite the possibility of such compensation, the PDH predicts that individuals with SLI should generally perform worse than typically-developing individuals on tasks assessing the learning and memory functions of the procedural memory system. To date, procedural memory in SLI has been explored using a range of different paradigms, including artificial grammar learning (Plante, Gomez, & Gerken, 2002), probabilistic classification (Kemény & Lukács, 2010; Mayor-Dubois, Zesiger, van der Linden, & Roulet-Perez, 2013), implicit statistical auditory learning (Evans, Saffran, & Robe-Torres, 2009; Mayor-Dubois et al., 2013), and Serial Reaction Time (SRT) tasks (Gabriel et al., 2013a; Hedenius et al., 2011; Lum, Conti-Ramsden, Page, & Ullman, 2012; Mayor-Dubois et al., 2013; Tomblin, Mainela-Arnold, & Zhang, 2007). A number of studies have reported procedural learning impairments in SLI (Adi-Japha, Strulovich-Schwartz, & Julius, 2011; Evans et al., 2009; Kemény & Lukács, 2010; Lum et al., 2012;

Lum, Gelgec, Conti-Ramsden, 2010; for phonotactic information only Mayor-Dubois et al., 2013; Tomblin et al., 2007). However, these results have not always been replicated (Gabriel, Maillart, Guillaume, Stefaniak, & Meulemans, 2011, Gabriel et al., 2013b, Gabriel, Stefaniak, Maillart, Schmitz, & Meulemans, 2012; Lum & Bleses, 2012; Mayor-Dubois et al., 2013). Thus, it is not yet clear whether procedural memory impairments constitute a core deficit of SLI.

The heterogeneity of study findings calls for a systematic assessment of the evidence in order to test whether or not SLI is indeed associated with overall procedural memory impairments, and to identify potential sources of variability between studies. To achieve this aim, we performed a systematic search of the literature and then used meta-analysis to pool results from studies and compute an overall result. Meta-analysis enables results from studies using similar methodologies to be combined, allowing population parameters to be estimated with greater precision (Borenstein, 2009; Hunter & Schmidt, 1990). Given inconsistent findings in past research, we also used meta-regression to investigate whether participant and study level variables predicted differences between study findings. Importantly, these quantitative approaches to reviewing past research overcome limitations with traditional qualitative narrative reviews in which it is difficult to pool results from studies, whilst simultaneously taking into account study-specific features such as effect size, sample size, and task-related methodological differences.

1.1. The SRT task

In our analyses we focused on the SRT task, because it has been widely used to investigate procedural memory in SLI. In the SRT task a visual stimulus repeatedly appears in one of four predefined spatial locations on a computer display. Participants are provided with a four-button response box. The topographic positioning of the four buttons matches the spatial locations where the stimulus appears on the display. Participants are instructed to press the button that matches the location of the visual stimulus. Reaction times (RTs) that measure how fast participants press the button following the appearance of the visual stimulus constitute the main dependent variable of interest. Presentation of the visual stimulus is divided into blocks. In the implicit version of the task, unknown to participants, stimulus presentations in most blocks follows a predefined sequence. This sequence repeats multiples times within these 'Sequenced Blocks'. Following one or more 'Sequenced Blocks', a 'Random Block' is then presented, in which the visual stimulus appears randomly, or in some studies a new sequence is introduced (e.g., Gabriel et al., 2011).

In participant groups that do not have procedural memory impairments, RTs become faster across the Sequenced Blocks, but then slow down in the Random Block (e.g., Lum, Kidd, Davis, & Conti-Ramsden, 2010; Thomas et al., 2004). This increase in RTs in the Random Block is taken to indicate that information about the sequence has been learnt (Robertson, 2007). However, in participant groups with neurodegenerative diseases or lesions affecting parts of the brain supporting the procedural memory system, the change in RTs between

Sequenced and Random Blocks is absent (for review see Siegert, Taylor, Weatherall, & Abernethy, 2006), or is smaller than in neurologically intact control participants (Knopman & Nissen, 1991; Pascual-Leone et al., 1993; Siegert et al., 2006).

1.2. Studies examining SRT task performance in SLI

Research investigating procedural memory in SLI using the SRT task has produced mixed results. In an early report, Tomblin et al. (2007) found that adolescents with SLI evidenced slower procedural learning of the sequence compared to typically developing (TD) age-matched controls. In studies by Lum and colleagues (Lum et al., 2012; Lum, Gelgec, et al., 2010; Lum, Kidd, et al., 2010), the TD control group showed a larger difference between sequenced and random RTs than the SLI group, also suggesting procedural memory deficits. However, studies by Gabriel and colleagues (Gabriel et al., 2013b, 2012) have not replicated these findings. In these studies, the SLI and control groups have shown comparable changes between sequenced and random RTs. Furthermore, in one study on implicit learning (Gabriel et al., 2011), a non-significant trend was found whereby children with SLI showed a larger increase in RTs from sequenced to random blocks than the control group.

Several explanations might account for these inconsistent findings. First, contrary to the predictions of Ullman and Pierpont (2005), procedural memory and thus SRT task performance might in fact remain intact in SLI. If this is the case, the differences between study findings most likely reflect sampling error, and thus a meta-analysis of SRT studies should not reveal a reliable impairment.

A second possibility is that differences in study findings may be explained by issues relating to statistical power (e.g., small sample sizes) in some studies. In this case, pooling study findings using meta-analysis, and thereby increasing statistical power, may show a deficit on the task.

Third, the age of the participants may impact on study findings. The mean age of participants with and without SLI ranges from 7 years to about 15 years of age (Lum & Bleses, 2012; Tomblin et al., 2007). One possibility is that the development of the procedural memory system might be delayed in SLI, as has been observed in other non-linguistic domains in affected children (Bishop & Edmundson, 1987; McArthur & Bishop, 2005). If this were the case, differences between individuals with SLI and their age-matched peers on the SRT task might decrease as participants become older. Another possibility is that in SLI, age may moderate the relationship between declarative memory-based compensatory processes and SRT task performance. Declarative memory has been shown to compensate for procedural memory impairments in adults with neuro-developmental or neurodegenerative conditions affecting the parts of the brain that support the procedural memory system (Beauchamp, Dagher, Panisset, & Doyon, 2008; Dagher, Owen, Boecker, & Brooks, 2001; Moody, Bookheimer, Vanek, & Knowlton, 2004; Rauch et al., 2007). Compensation may be less likely in younger children with SLI because declarative memory is still developing in childhood (Giedd et al., 1999; Lum, Kidd, et al., 2010; Ullman, 2005; Ullman & Pierpont, 2005). On this account, it would also be

predicted that SRT impairments in SLI may be negatively correlated with age: that is, the older the participants, the greater likelihood of compensation.

Fourth, SRT task characteristics may impact study findings. In particular, the amount of training in the task, as measured by the number of times participants are exposed to the sequence, could affect outcomes. In the SLI/SRT task literature, training has varied considerably between studies, with the number of times participants are exposed to the sequence ranging from 20 to 96 exposures (Gabriel et al., 2011; Lum, Gelgec, et al., 2010). This aspect of the task is likely to impact on study findings since learning by the procedural memory system requires repetition or repeated exposures to information. If the procedural memory system is compromised in SLI, affected individuals may require more training or exposures to the sequence. Research into implicit statistical learning of auditory information, which depends upon procedural memory brain structures (Karuza et al., 2013) supports this claim. Evans et al. (2009) found that children with SLI evidenced poorer implicit learning of auditory information following a 24 min exposure period, but not when the exposure period was increased to 48 min. This pattern of results suggests that procedural learning can occur in SLI, but that affected individuals may require increased exposure to the information. In the case of the SRT task, we would predict that differences between SLI and age-matched peers would be smaller in studies that have provided more exposures to the repeating sequence.

Finally, the participant response method used in the SRT task may also account for conflicting findings in the literature. The most common method involves having participants press one of the four buttons on a button box or computer keyboard in response to the visual stimulus (e.g., Lum & Bleses, 2012; Lum et al., 2012; Lum, Gelgec et al., 2010; Lum, Kidd et al., 2010; Tomblin et al., 2007). However, Gabriel et al. (2011) suggest that since individuals with SLI often have fine motor problems (Hill, 2001), the standard response format of the task may disproportionally disadvantage these participants. In support of this proposal, Gabriel et al. (2012) found no significant differences between children with SLI and age-matched controls on an SRT task that required participants to use a touchscreen to respond to visual stimuli. If the response method contributes to different findings in the literature, differences between participants with and without SLI are likely to be smaller in those studies that use a touchscreen compared to those that use a button box or keyboard.

In this report we used meta-analysis to systematically review and integrate the evidence relating to performance of individuals with SLI on SRT tasks. This synthesis provides key information relevant to the claims of the PDH, and thus strengthens our understanding of the potential underlying causes of the disorder. Specifically, our study asks two questions: First, to what extent do individuals with SLI show poor impaired performance in procedural learning on the SRT task? Second, what factors − in particular response method, participant age, and number of exposures to the sequence − may explain variability in findings. Are participant age and amount of training on the SRT Task key factors in explaining variability among the across studies findings?

2. Methods

2.1. Study design

A systematic search for articles was undertaken using searches in CINAHL (EbscoHost), EMBASE, ERIC (EbscoHost), MEDLINE (OvidSP) and PsycInfo (EbscoHost) electronic databases up to June 2013. The search strategy aimed to identify studies that presented a version of Nissen and Bullemer (1987) SRT task to participants with SLI. Details of all keywords, fields search, Boolean operators and syntax used for each database are presented in Appendix A of the online supplemental material.

2.1.1. Study inclusion criteria

An inclusionary criteria based on previous meta-analyses of SRT investigations (Siegert et al., 2006; Siegert, Weatherall, & Bell, 2008) was used to identify studies that could be included in the meta-analysis. First, only studies published later than 1986 were included, since the SRT task used to assess implicit learning had not been described before this date (i.e., Nissen & Bullemer, 1987). Second, only investigations that reported on an original piece of research, which had been published or had been accepted for publication were included. Third, the SRT task used in the study needed to be a version of Nissen and Bullemer (1987) task. Specifically, visual stimulus presentations were required to be presented in blocks comprising either sequenced or random presentations. This criterion led to the exclusion of studies that interspersed sequenced and random trials (Hedenius et al., 2011; Mayor-Dubois et al., 2013). Fourth, participants in the study had to include one group of individuals with SLI and one age-matched typically-developing control group who did not have language impairments. Fig. 1 summarizes studies removed following application of each criterion according to PRISMA guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009).

2.1.2. Study selection

After the removal of duplicate entries, one reviewer assessed all the abstracts. A random sample of 10% of all abstracts was assessed by a second reviewer. Any disagreements were resolved by discussion. Finally, the reviewers independently retrieved full-text articles and screened them accordingly to the eligibility criteria. There was 100% agreement on the selection of these articles. A total of seven studies were identified. A hand search of the reference list of these studies led to the identification of one study that was in press that also met the aforementioned criteria. Thus a total of eight articles were included and data extracted for use in the meta-analysis and meta-regression. A summary of each study's participants and SRT task methodology is summarized in Table 1.

Overall the studies identified for inclusion in the meta-analysis had comparable parameters with respect to presentation of the sequence and then the random blocks. However, the structure of the sequence presented to participants was found to be different for one study. In seven studies participants were presented with a deterministic sequence (Gabriel et al., 2012, 2013a, 2013b; Lum & Bleses, 2012; Lum et al., 2012; Lum, Gelgec, et al., 2010; Tomblin et al., 2007), however, the study by Gabriel et al. (2011) used a probabilistic sequence instead. Probabilistic sequences permit deviations in the order the visual stimulus can appear during Sequenced Blocks. For example, if the sequence is 31432412 and deterministic, then 3 will always be followed by 1 or 2. But if the sequence is probabilistic, it is permissible to have 4 appear after 3 on some trials (e.g., Gabriel et al., 2011). Preliminary analyses using meta-regression revealed the use of probabilistic sequence was not systematically influencing study findings after controlling for differences in study ages and exposures to the sequence ($z = -.642, p = .521$). Consequently, all studies identified by our search, including the study by Gabriel et al. (2011), were included in the main analyses.

2.1.3. Effect size calculations and data extraction procedures

The most commonly used method of comparing two groups on SRT tasks involves determining whether the difference in RTs between the final block, comprising random stimulus presentations, and preceding block, comprising sequenced stimulus presentations, is different between groups (e.g., Nissen & Bullemer, 1987). More specifically, the main result of interest is whether a significant Group (i.e., SLI *vs* Control) × Block (i.e., Random Block *vs* Sequence) interaction is observed. Data was extracted from the results reported in each study to allow the effect size for the interaction to be computed and its variance. The effect size measure computed was a standardized mean difference (SMD), which quantified differences in groups on the SRT task in standard deviation units. The value for SMD was calculated so that positive values indicated that the control group in each study demonstrated higher levels of implicit learning on the task. That is, positive values indicate that children in the SLI group performed poorly on the SRT Task relative to children in the control group. The general formula used to compute SMD is shown in (1) and its variance in (2). This approach has been used previously in a meta-analysis of SRT studies Siegert et al. (2006).

$$\text{SMD} = \frac{\overline{x}_{\text{control}} - \overline{x}_{\text{SLI}}}{\text{SD}_{\text{pooled}}} \tag{1}$$

$$\text{var(SMD)} = \frac{n_{\text{control}} + n_{\text{SLI}}}{n_{\text{control}} \times n_{\text{SLI}}} + \frac{\text{SMD}^2}{2(n_{\text{control}} + n_{\text{SLI}})} \tag{2}$$

where:

$\overline{x}$ = Mean difference in RTs between the final random block and preceding sequence block.

$\text{SD}_{\text{pooled}}$ = within-group standard deviation of the difference between the final random block and preceding block, pooled across the control and study group.

A single effect size was extracted from seven of the eight studies. For one study it was necessary to combine two sets of effect sizes reported. Specifically, in the study by Gabriel et al. (2012) two effect sizes that compared children with SLI and age-matched children in the control group on two different SRT tasks were averaged. In that study one task required children to use a response pad as an input device and in the second a touchscreen as an input device.
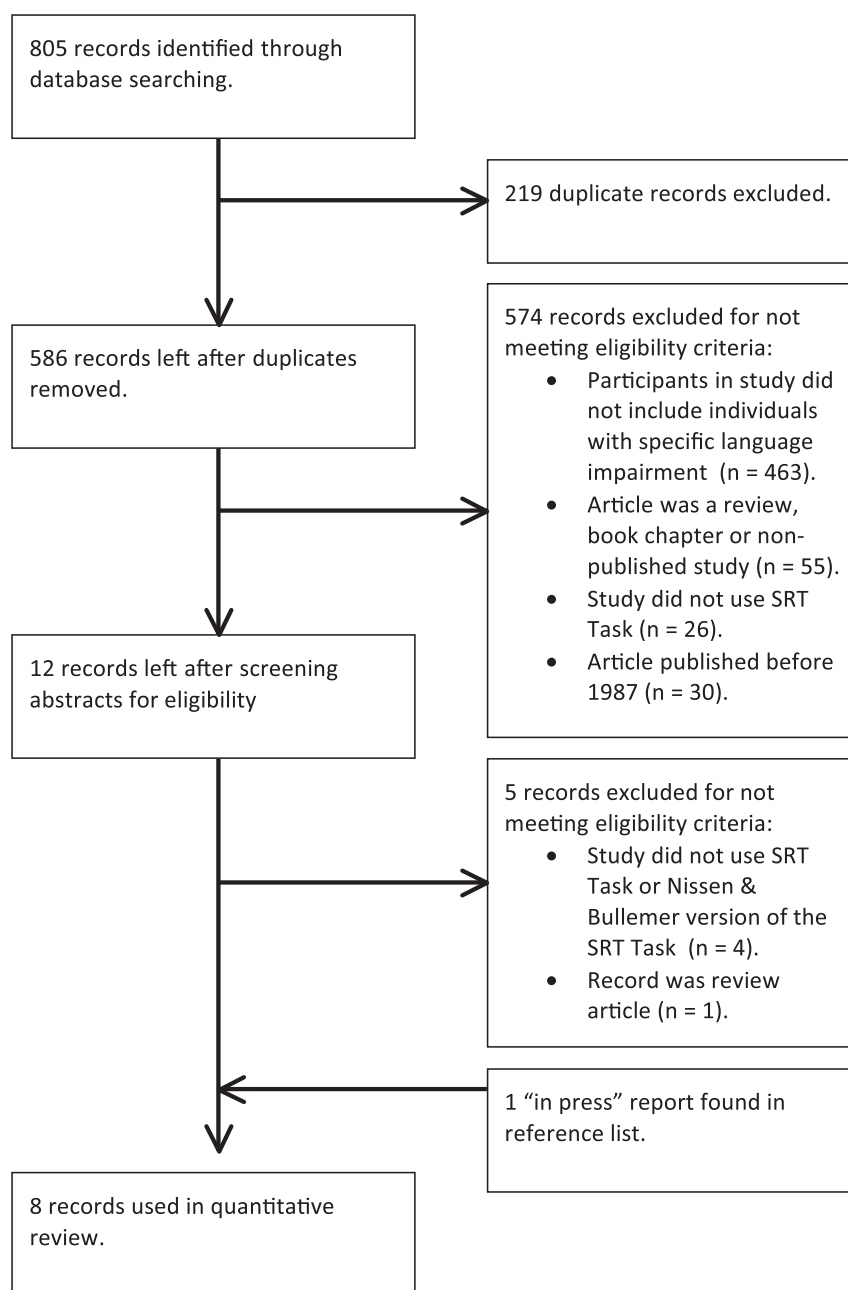
Fig. 1 — PRISMA flowchart showing the process of article identification.

The data extracted from each study to compute (1) and (2) varied depending on the results presented. Conversion of results to SMD and its variance was undertaken using Comprehensive Meta-Analysis Software Package (Borenstein, Rothstein, & Cohen, 1999). The data extracted from individual studies to compute SMD and *var* (SMD) is presented in Appendix B of the online supplemental material. The specific conversion used in the Comprehensive Meta-Analysis Software Package to obtain (1) and (2) is also described in Appendix B.

2.1.4. Meta-analytic procedures

To quantify the overall difference between SLI and controls on the SRT task, individual study effect sizes were pooled, and a

weighted averaged effect size was computed using a random effects model (Hedges & Olkin, 1985). By using a random effects model we are assuming that heterogeneity or differences in study level effect sizes are the sum of within-study error (e.g., sampling error) and between-study error (e.g., systematic influences on effect sizes). To evaluate the statistical significance of computed effect sizes an alpha level of .05 was used.

The I^2 statistic (Higgins & Thompson, 2002) was used to measure heterogeneity between effect sizes. This index expresses the amount of between-study error as a percentage. Alternatively stated, the index measures the heterogeneity in effect sizes not attributable to within-study error/sampling

Table 1 — Methodological characteristics of studies included in the meta-analysis.

| Study	Sample size		Mean age (years)		Exposures to sequence	Input method
	SLI (n_{SLI})	Control ($n_{control}$)	SLI group	Control group		
Gabriel et al. (2011)	16	16	10.2	10.3	96	Touchscreen
Gabriel et al. (2012)	15	15	10.3	10.4	48	Keyboard (Exp. 1), Touchscreen (Exp. 2)
Gabriel et al. (2013a)	23	23	9.7	9.6	48	Touchscreen
Gabriel et al. (2013b)	16	16	9.9	9.8	48	Touchscreen
Lum and Bleses (2012)	13	20	7.7	7.9	24	Button box
Lum et al. (2012)	51	51	9.8	9.9	36	Button box
Lum, Gelgec, et al. (2010)	14	15	7.1	7.0	36	Button box
Tomblin et al. (2007)	38	47	15	14.8	20	Button box

error. As a guideline, Higgins, Thompson, Deeks, and Altman (2003) suggest that values of 25%, 50% and 75% correspond to low, moderate and high levels of between-study error respectively.

Mixed-effects subgroup analyses (Borenstein, 2009) were used to investigate whether studies' response mode was related to effect sizes. Specifically, we tested whether the effect sizes for keyboard or button box studies were different than the effect sizes of touchscreen studies. Finally, multivariable meta-regression (Greenland, 1987) was used to investigate the contribution of participants' age and SRT task characteristics to heterogeneity in study level effect sizes.

3. Results

3.1. Evaluation of publication bias

Fig. 2 presents preliminary analyses investigating selection/publication bias using a funnel plot (Egger, Smith, Schneider, & Minder, 1997). Funnel plots show publication bias if individual effect sizes are asymmetrically distributed around the weighted average effect size for those studies that have low
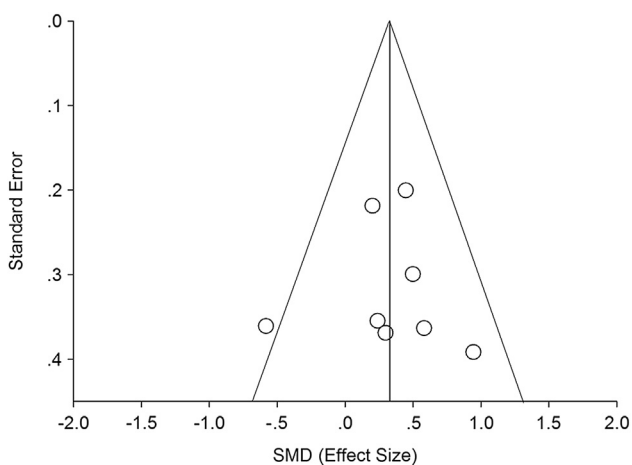
Fig. 2 — Funnel plot showing SMD plotted against standard error (which measures study precision). Note that effect sizes are symmetrically distributed when standard errors are high (i.e., study precision is low). Less variability in effect sizes is observed at higher levels of study precision.

precision. Precision in this context refers to the accuracy of a study's findings. Quantitatively, this is captured by the standard error computed for individual effect sizes. Thus, a study with relatively low precision will have a larger standard error than a study with relatively high precision. Egger's test of asymmetry was not significant [Intercept = −.126, t (6) = .069, p = .947]. Non-significant asymmetry indicates that publication/selection bias was not found in our systematic search.
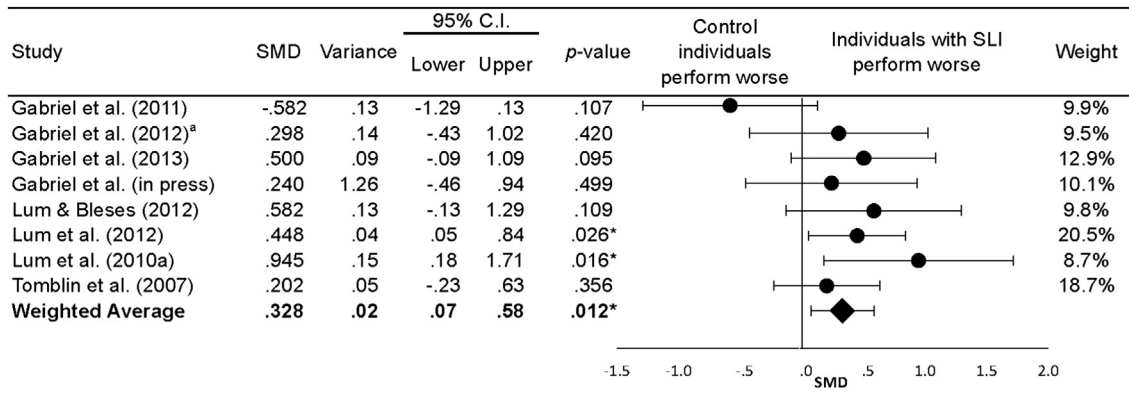
3.2. Differences between participants with and without SLI on SRT tasks

A forest plot showing study effect sizes and the weighted average is presented in Fig. 3. Positive SMD values in the forest plot show that the control group had a larger difference in RTs between sequenced and random blocks compared to the SLI group.

The weighted average SMD for the studies was observed to be .328 and was statistically significant (p = .012). This weighted average value represents a small to medium effect size according to Cohen's (1988) taxonomy. This result can be interpreted to suggest that the difference in RTs between sequenced and random blocks is around .328 standard deviations larger in TD control participants than in participants with SLI, suggesting worse procedural learning in children with SLI. However, Fig. 3 shows variability in study level effect sizes ranging from .945 to −.582. The observed value of the I^2 statistic for the studies in Fig. 3 was 32.8% (i.e., 32.8% of the heterogeneity in study level effect sizes reflects between-study error). This value indicates small to medium levels of heterogeneity using the guidelines outlined by Higgins et al. (2003).

The next set of analyses investigated the source (or sources) of between-study heterogeneity using Mixed-effects subgroup analyses and multivariable random-effects meta-regression.

Mixed-effects subgroup analyses were used to investigate whether effect sizes for touchscreen studies were different from effect sizes for button boxes or keyboard studies. The effect sizes used in this analysis are SMDs presented in Fig. 3. There was one exception. Gabriel et al. (2012) used both a touchscreen and a button box in separate experiments. To increase the number of data points for touchscreen effect sizes, only touchscreen results from the Gabriel et al. (2012) were used in this analysis. Note that including both touchscreen and keyboard effect sizes for the Gabriel et al. study would bias the results by treating dependent data points

Study	SMD	Variance	95% C.I.		p-value	Control individuals perform worse	Individuals with SLI perform worse	Weight
			Lower	Upper				
Gabriel et al. (2011)	-.582	.13	-1.29	.13	.107			9.9%
Gabriel et al. (2012)[a]	.298	.14	-.43	1.02	.420			9.5%
Gabriel et al. (2013)	.500	.09	-.09	1.09	.095			12.9%
Gabriel et al. (in press)	.240	1.26	-.46	.94	.499			10.1%
Lum & Bleses (2012)	.582	.13	-.13	1.29	.109			9.8%
Lum et al. (2012)	.448	.04	.05	.84	.026*			20.5%
Lum et al. (2010a)	.945	.15	.18	1.71	.016*			8.7%
Tomblin et al. (2007)	.202	.05	-.23	.63	.356			18.7%
Weighted Average	**.328**	**.02**	**.07**	**.58**	**.012***			

Notes: [a]Effect size is the average of results over two conditions. In one condition participants completed the SRT Task using response pad as an input device and in the second participants used a touchscreen.
*p < .05; **p < .001

Fig. 3 – Forest plot showing study level and average weighted effect sizes for individuals with SLI and control individuals.

as independent (see Tramèr, Reynolds, Moore, & McQuay, 1997). The average effect size for studies using keyboard/response boxes versus touchscreens is presented in Fig. 4. Results from the analyses revealed no significant differences [Q (1) = .666, p = .415].
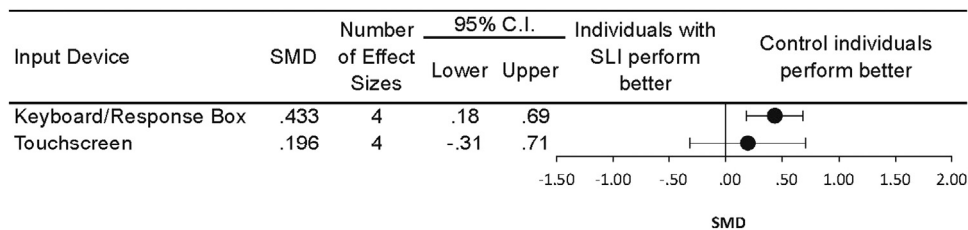
The final analyses used multivariable meta-regression to test whether participants' age and number of exposures to the sequence predicted the effect sizes presented in Fig. 3. The predictor variables in the analyses were participants' age and number of exposures as presented in Table 1. Because the covariates – age and number of exposures to the sequence – were predicted to decrease study effect sizes, a one-tailed significance test was used. The one-tailed test also protected against the probability of making a Type II error, as it is recommended that the ratio of predictor variables to studies used is 1:10 (Borenstein, 2009).

A summary of the model coefficients is presented in Table 2. For illustrative purposes, Fig. 5 plots effect sizes predicted by the model against observed effect sizes. Overall, the model was found to be a significant predictor of effect sizes [Q (1) = 7.138, p = .014, R^2 = .929]. Of particular interest was that the model accounted for 92.9% of heterogeneity between effect sizes. It should be noted that in meta-regression, only the between-study heterogeneity is modelled. Thus the R^2 value corresponds to the amount of variance captured by the I^2 statistic. Both age and number of exposures to the sequence were found to significantly predict study level effect sizes. A significant negative association was observed for both

predictors. That is, older participants and increased exposures to the sequence were significant predictors of small effect sizes. That is, when participants were younger or had fewer exposures to the sequence, the observed study effect sizes were larger, demonstrating bigger differences in performance between groups.

4. Discussion

In this paper we used meta-analysis and meta-regression to investigate and evaluate available evidence regarding procedural learning abilities in SLI, as indexed by SRT task performance. The first goal of our synthesis was to estimate the magnitude of the difference between participants with and without SLI on implicit sequence learning as measured by SRT tasks. The average effect size computed from eight studies, representing 186 participants with SLI and 203 typically-developing peers was found to be .328, and was statistically significant. Our second goal was to investigate the sources of heterogeneity among study findings. Consistent with expectations, the age of participants and the number of exposures to the sequence were found to predict variability across the studies, i.e., study effect sizes. Overall, the results indicate a significant difference between participants with and without SLI on SRT tasks. However, the magnitude of the effect appears to vary as a function of the age of participants and characteristics of the SRT task.

Input Device	SMD	Number of Effect Sizes	95% C.I.		Individuals with SLI perform better	Control individuals perform better
			Lower	Upper		
Keyboard/Response Box	.433	4	.18	.69		
Touchscreen	.196	4	-.31	.71		

Fig. 4 – Forest plot showing average effect sizes for study's using Keyboard/Response Boxes and Touchscreens as the input device on the SRT Task.

Table 2 − Summary of variables in the meta-regression model.

Variables in the model	Coefficient summary			
	β	B	95% CI for B	p-value[a]
Constant		2.03	.588, 3.470	
Age	−.65	−.10	−.212, .013	.042[*,a]
No. exposures to sequence	−.91	−.02	−.029, −.003	.007[*,a]

[*]p < .05.
[a] One-tailed test.

Overall, the findings are consistent with the prediction of the PDH (Ullman, 2004; Ullman & Pierpont, 2005). The pooled results from all studies showed significantly impaired procedural learning in SLI. The average weighted effect size was found to be significantly different from zero. The magnitude of the difference between individuals with and without SLI represented a small to medium effect size. Note that this result is obtained with the classic version of the SRT task (e.g., Nissen & Bullemer, 1987), a task that taps visuo-perceptual-motor procedural learning rather than procedural learning in the verbal domain. In contrast, children with SLI show no deficits whatsoever in tasks tapping visual non-verbal learning in declarative memory (Lum et al., 2012).

In this report we also investigated the sources of heterogeneity between study findings. We found that sampling error was the largest component, accounting for 64% of the observed heterogeneity. The remaining 36% of heterogeneity indicated the presence of one or more systematic influences, referred to as between-study error. Subgroup analyses revealed that the method for collecting responses on the SRT task did not account for differences in effect sizes. Specifically, there was no significant difference in effect sizes between studies that used a keyboard/response box versus those using a touchscreen. In contrast, meta-regression analysis showed that nearly all of the between-study error could be predicted
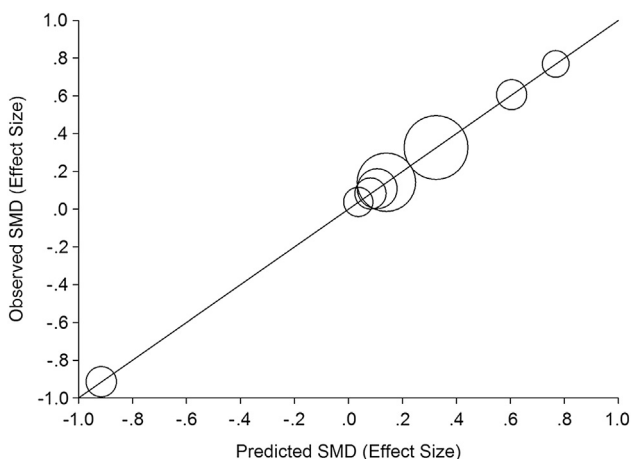
Fig. 5 − Predicted study effect sizes from model reported in Table 2, plotted against observed effect sizes. Departures from the diagonal line are residuals. Data points are proportionally sized according to their weight in the model.

by a model comprising participants' age and the number of exposures to the sequence. Here we briefly discuss each of these factors.

First of all, the meta-regression showed that the number of exposures to the sequence in the SRT task was a significant predictor of study effect sizes. Specifically, differences between participants with and without SLI on SRT tasks were smaller when studies provided participants with more exposures to the sequence. This association is consistent with the idea that, as a consequence of procedural memory impairments in SLI, affected individuals require more training or exposures to the information in order to demonstrate implicit learning that is comparable to unaffected individuals. As noted earlier, a similar trend has been observed in implicit learning of auditory information (Evans et al., 2009). The results from this meta-analysis suggest slower learning in the visuo-spatial domain for individuals with SLI. This proposal can be directly tested experimentally in future research. Based on the findings of this meta-analysis, we would predict that individuals with SLI should show higher levels of implicit sequence learning if presented with more training trials on the SRT task.

Meta-regression analyses also showed that age was a significant predictor of differences in study findings. As expected, studies with older participants reported smaller effect sizes between participants with and without SLI. In line with the PDH, we suggest that one interpretation of this association is that the results reflect compensatory processes of the declarative memory system due to increased involvement of this memory system during childhood. This interpretation is further supported by evidence suggesting that children with SLI can rely on declarative memory rather than procedural memory in tasks involving language skills, in particular grammar. Lum et al. (2012) found significant correlations between language tasks involving grammar and declarative memory, but no association between these tasks and procedural memory in 10 year old children with SLI. In contrast, this pattern of associations was not found in typically-developing children of the same age, for whom the strongest associations were between procedural memory and language tasks involving grammar, consistent with the PDH.

The effect of age on study effect sizes could instead or additionally be an indicator of delayed maturation of the procedural memory system in SLI. On this view, differences in performance between individuals with and without SLI on SRT tasks would be expected to become smaller as participants with SLI grow older and their frontal/subcortical neural networks develop more fully. Future neuroimaging studies of SRT in SLI should be particularly revealing in disentangling maturational versus compensatory explanations. Specifically, if compensatory mechanisms underlie SRT task performance in SLI, we would expect to see greater activation in the medial temporal lobe structures that underlie declarative memory in individuals with SLI as compared to their unaffected peers.

4.1. Limitations and directions for future research

Meta-analysis and meta-regression are useful techniques for highlighting associations between variables studied in past research. However, their limitations need to be taken into

account to avoid over-interpretation of the findings. First, all studies entered into the meta-analysis used a correlational design. Consequently, these data do not test whether procedural learning problems in SLI are causally related to these individuals' language problems. The average study effect size we reported indicates an association between poor performance on the SRT task and SLI. These findings are consistent with Ullman and Pierpont (2005) PDH of SLI, i.e., these individuals appear to have an impaired procedural memory system. Future longitudinal studies, including cross-lagged research, are needed to examine potential causal relationship between procedural learning deficits and language problems.

Second, a limitation with meta-regression is that predictors used in the model are most likely to be correlated with other measured and non-measured variables (see Thompson & Higgins, 2002). Thus, while we found that age and the number of exposures to the sequence predicted effect sizes, it is possible that these associations may be better explained with reference to another correlated variable. Given this, we suggest that the time is ripe for experimental studies specifically designed to examine the claims made by the findings from this meta-regression. Future studies that directly manipulate participant age and number of exposures, as well as that directly assess the declarative memory system in order to examine its potential compensatory role, are warranted.

Acknowledgement

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cortex.2013.10.011.

REFERENCES [*]

Adi-Japha, E., Strulovich-Schwartz, O., & Julius, M. (2011). Delayed motor skill acquisition in kindergarten children with language impairment. *Research in Developmental Disabilities, 32*(6), 2963–2971.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (4th ed.). Washington, DC: American Psychiatric Association.

Beauchamp, M., Dagher, A., Panisset, M., & Doyon, J. (2008). Neural substrates of cognitive skill learning in Parkinson's disease. *Brain and Cognition, 68*(2), 134–143.

Bishop, D. V. M., & Edmundson, A. (1987). Specific language impairment as a maturational lag: evidence from longitudinal data on language and motor development. *Developmental Medicine & Child Neurology, 29*(4), 442–459.

Borenstein, M. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons.

Borenstein, M., Rothstein, H., & Cohen, J. (1999). *Comprehensive meta-analysis: A computer program for research synthesis* [Computer software]. Englewood: Biostat.

Cohen, J. (1988). *Statistical poower analysis for the behavioral sciences*. Lawrence Erlbaum.

Dagher, A., Owen, A. M., Boecker, H., & Brooks, D. J. (2001). The role of the striatum and hippocampus in planning A PET activation study in Parkinson's disease. *Brain, 124*(5), 1020–1032.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language and Hearing Research, 52*(2), 321–335.

**Gabriel, A., Maillart, C., Guillaume, M., Stefaniak, N., & Meulemans, T. (2011). Exploration of serial structure procedural learning in children with language impairment. *Journal of the International Neuropsychological Society, 17*(2), 336–343.

**Gabriel, A., Maillart, C., Stefaniak, N., Lejeune, C., Desmottes, L., & Meulemans, T. (2013a). Procedural learning in specific language impairment: effects of sequence complexity. *Journal of the International Neuropsychological Society, 19*(3), 264–271.

**Gabriel, A., Maillart, C., Stefaniak, N., Lejeune, C., Desmottes, L., & Meulemans, T. (2013b). Procedural learning across modalities in French speaking children with specific language impairment. *Applied Psycholinguistics*. in press.

**Gabriel, A., Stefaniak, N., Maillart, C., Schmitz, X., & Meulemans, T. (2012). Procedural visual learning in children with specific language impairment. *American Journal of Speech Language Pathology, 21*(4), 329–341.

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49*(1), 87–115.

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.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience, 2*(10), 861–862.

Greenland, S. (1987). Quantitative methods in the review of epidemiologic literature. *Epidemiologic Reviews, 9*(1), 1–30.

Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Veríssimo, J., Dye, C. D., et al. (2011). Grammar predicts procedural learning and consolidation deficits in children with Specific Language Impairment. *Research in Developmental Disabilities, 32*(6), 2362–2375.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. FL: Orlando: Academic Press Orlando.

Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558.

Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560.

Hill, E. L. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders, 36*(2), 149–171.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park: Sage Publications.

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain and Language, 127*(1), 46–54.

Kemény, F., & Lukács, Á. (2010). Impaired procedural learning in language impairment: results from probabilistic

---

* Studies included in the meta-analysis.

categorization. *Journal of Clinical and Experimental Neuropsychology, 32*(3), 249—258.

Knopman, D., & Nissen, M. J. (1991). Procedural learning is impaired in Huntington's disease: evidence from the serial reaction time task. *Neuropsychologia, 29*(3), 245—254.

Leonard, L. B. (2000). *Children with specific language impairment*. MIT Press.

**Lum, J. A. G., & Bleses, D. (2012). Declarative and procedural memory in Danish speaking children with specific language impairment. *Journal of Communication Disorders, 45*(1), 46—58.

**Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex, 48*(9), 1138—1154.

**Lum, J. A. G., Gelgec, C., & Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders, 45*(1), 96—107.

Lum, J. A. G., Kidd, E., Davis, S., & Conti-Ramsden, G. (2010). Longitudinal study of declarative and procedural memory in primary school-aged children. *Australian Journal of Psychology, 62*(3), 139—148.

Mayor-Dubois, C., Zesiger, P., van der Linden, M., & Roulet-Perez, E. (2013). Nondeclarative learning in children with specific language impairment: predicting regularities in the visuomotor, phonological, and cognitive domains. *Child Neuropsychology*. in press.

McArthur, G. M., & Bishop, D. V. M. (2005). Speech and non-speech processing in people with specific language impairment: a behavioural and electrophysiological study. *Brain and Language, 94*(3), 260—273.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med, 6*(7), e1000097.

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: an update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology, 19*(1), 78—94.

Moody, T. D., Bookheimer, S. Y., Vanek, Z., & Knowlton, B. J. (2004). An implicit learning task activates medial temporal lobe in patients with Parkinson's disease. *Behavioral Neuroscience, 118*(2), 438—442.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cognitive Psychology, 19*(1), 1—32.

Pascual-Leone, A., Grafman, J., Clark, K., Stewart, M., Massaquoi, S., Lou, J. S., et al. (1993). Procedural learning in Parkinson's disease and cerebellar degeneration. *Annals of Neurology, 34*(4), 594—602.

Plante, E., Gomez, R., & Gerken, L. A. (2002). Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders, 35*(5), 453—462.

Rauch, S. L., Wedig, M. M., Wright, C. I., Martis, B., McMullin, K. G., Shin, L. M., et al. (2007). Functional magnetic resonance imaging study of regional brain activation during implicit sequence learning in obsessive-compulsive disorder. *Biological Psychiatry, 61*(3), 330—336.

Robertson, E. M. (2007). The serial reaction time task: implicit motor skill learning? *The Journal of Neuroscience, 27*(38), 10073—10075.

Siegert, R. J., Taylor, K. D., Weatherall, M., & Abernethy, D. A. (2006). Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis. *Neuropsychology, 20*(4), 490—495.

Siegert, R. J., Weatherall, M., & Bell, E. M. (2008). Is implicit sequence learning impaired in schizophrenia? A meta-analysis. *Brain and Cognition, 67*(3), 351—359.

Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience, 5*(9), 721—728.

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., et al. (2004). Evidence of developmental differences in implicit sequence learning: an fMRI study of children and adults. *Journal of Cognitive Neuroscience, 16*(8), 1339—1351.

Thompson, S. G., & Higgins, J. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine, 21*(11), 1559—1573.

**Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development, 3*(4), 269—293.

Tramèr, M. R., Reynolds, D., Moore, R. A., & McQuay, H. J. (1997). Impact of covert duplicate publication on meta-analysis: a case study. *British Medical Journal, 315*(7109), 635—640.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition, 92*(1), 231—270.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: the declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp. 141—178). Washington, DC: Georgetown University Press.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex, 41*(3), 399—433.

Ullman, M. T. and Pullman, M. Y. A compensatory role for declarative memory in neurodevelopmental disorders, submitted for publication.

World Health Organization. (1996). *Multiaxial classification of child and adolescent psychiatric disorder: The ICD-10 classification of mental and behavioural disorders in children and adolescents*. Cambridge: Cambridge University Press.