



## UvA-DARE (Digital Academic Repository)

### Visual artificial grammar learning in dyslexia

*A meta-analysis*

van Witteloostuijn, M.; Boersma, P.; Wijnen, F.; Rispens, J.

**DOI**

[10.1016/j.ridd.2017.09.006](https://doi.org/10.1016/j.ridd.2017.09.006)

**Publication date**

2017

**Document Version**

Final published version

**Published in**

Research in Developmental Disabilities

**License**

CC BY

[Link to publication](#)

**Citation for published version (APA):**

van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J. (2017). Visual artificial grammar learning in dyslexia: A meta-analysis. *Research in Developmental Disabilities*, 70, 126-137. <https://doi.org/10.1016/j.ridd.2017.09.006>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Contents lists available at ScienceDirect

# Research in Developmental Disabilities

journal homepage: [www.elsevier.com/locate/redevdis](http://www.elsevier.com/locate/redevdis)

Research Paper

## Visual artificial grammar learning in dyslexia: A meta-analysis

Merel van Witteloostuijn<sup>a,\*</sup>, Paul Boersma<sup>a</sup>, Frank Wijnen<sup>b</sup>, Judith Rispens<sup>a</sup><sup>a</sup> University of Amsterdam, Spuistraat 134, 1012 VB, Amsterdam, The Netherlands<sup>b</sup> Utrecht University, Trans 10, 3512 JK, Utrecht, The Netherlands

### ARTICLE INFO

#### Keywords:

Implicit learning  
Artificial grammar learning  
Dyslexia  
Meta-analysis

### ABSTRACT

**Background:** Literacy impairments in dyslexia have been hypothesized to be (partly) due to an implicit learning deficit. However, studies of implicit visual artificial grammar learning (AGL) have often yielded null results.

**Aims:** The aim of this study is to weigh the evidence collected thus far by performing a meta-analysis of studies on implicit visual AGL in dyslexia.

**Methods and procedures:** Thirteen studies were selected through a systematic literature search, representing data from 255 participants with dyslexia and 292 control participants (mean age range: 8.5–36.8 years old).

**Results:** If the 13 selected studies constitute a random sample, individuals with dyslexia perform worse on average than non-dyslexic individuals (average weighted effect size = 0.46, 95% CI [0.14 ... 0.77],  $p = 0.008$ ), with a larger effect in children than in adults ( $p = 0.041$ ; average weighted effect sizes 0.71 [sig.] versus 0.16 [non-sig.]). However, the presence of a publication bias indicates the existence of missing studies that may well null the effect.

**Conclusions and implications:** While the studies under investigation demonstrate that implicit visual AGL is impaired in dyslexia (more so in children than in adults, if in adults at all), the detected publication bias suggests that the effect might in fact be zero.

### 1. Introduction

Individuals with dyslexia have severe and persistent difficulties with learning to read and spell. These difficulties occur despite normal intelligence, adequate educational and socio-economic opportunities, and in absence of sensory or neurological impairment<sup>1</sup> (DSM-IV; American Psychiatric Association, 2000). A generally accepted hypothesis is that the persistent difficulties with written language result from a core deficit in phonological processing and, specifically, phonological awareness (see Melby-Lervåg, Lyster, & Hulme, 2012 for a meta-analysis). Phonological awareness is the ability to detect and manipulate phonological segments of words (Shankweiler et al., 1995) and is related to the ability to map letters to sounds, which in turn affects the ability to learn to read and spell. Individuals with dyslexia also experience difficulties in other areas of language. Subtle problems have been reported in the area of inflectional morphology (e.g. pluralization and tense marking: Joanisse, Manis, Keating, & Seidenberg, 2000; subject-verb agreement: Rispens & Been, 2007; Rispens, Roeleven, & Koster, 2004) and syntax (relative clauses: Mann, Shankweiler, & Smith, 1984; Stein, Cairns, & Zurif, 1984, passive sentences: Stein et al., 1984; binding: Waltzman & Cairns, 2000). Additionally, dyslexia is associated with a range of non-linguistic cognitive dysfunctions, including impairments in visual and auditory processing

\* Corresponding author.

E-mail addresses: [M.T.G.vanWitteloostuijn@uva.nl](mailto:M.T.G.vanWitteloostuijn@uva.nl) (M. van Witteloostuijn), [Paul.Boersma@uva.nl](mailto:Paul.Boersma@uva.nl) (P. Boersma), [F.N.K.Wijnen@uu.nl](mailto:F.N.K.Wijnen@uu.nl) (F. Wijnen), [J.E.Rispens@uva.nl](mailto:J.E.Rispens@uva.nl) (J. Rispens).

<sup>1</sup> Note that in the DSM-V (American Psychiatric Association, 2013), dyslexia is included under the umbrella term ‘Specific Learning Disorder’.

<http://dx.doi.org/10.1016/j.ridd.2017.09.006>

Received 3 October 2016; Received in revised form 9 August 2017; Accepted 9 September 2017

Available online 18 September 2017

0891-4222/ © 2017 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

(Stein & Walsh, 1997; Tallal, 2004), attention (Facoetti & Paganoni & Lorusso, 2000), motor functioning (Ramus, Pidgeon, & Frith, 2003), and verbal working memory (Gathercole & Alloway, 2006; Gathercole & Baddeley, 1990; Swanson & Jerman, 2007).

Several theories have attempted to define the underlying deficit that accounts for the range of problems experienced by individuals with dyslexia. One recent approach is explaining dyslexia as the result of a problem with implicit learning (see Nicolson & Fawcett, 2007; Ullman & Pierpont, 2005). The term implicit learning refers to the process through which humans extract rules and regularities from visual and auditory sequences available in the environment. Importantly, this happens in absence of awareness.

### 1.1. Implicit learning and literacy acquisition

Many studies have related implicit learning abilities to different aspects of language acquisition: the ability to segment words from continuous speech (Saffran, Aslin, & Newport, 1996), the acquisition of phonological categories and phonotactics (Nicolson & Fawcett, 2007; Wijnen, 2013), vocabulary acquisition (Evans, Saffran, & Robe-Torres, 2009; Yu, 2008), and more general language processing (e.g. passives: Kidd, 2012; relative clauses: Misyak, Christiansen, & Tomblin, 2010). Most important to the present discussion is the relationship between implicit learning and the acquisition of literacy skills, as these are the skills most affected in individuals with dyslexia. Learning to read and spell involves the mapping between letters and sounds (grapheme-to-phoneme mapping), which requires phonological awareness and knowledge of the orthographic system. This mapping, and the writing system in general, comprises many regularities. For example, a single letter (e.g. 'c') can map onto several phonemes (e.g. /k/ /s/). Whether the letter 'c' is realized as a/k/or an/s/, depends on co-occurring letters (e.g. the letter 'c' followed by the letter 'a' generally results in the realization of the phoneme/k/as in *can't*, but in the phoneme/s/when followed by an 'e' as in *cent*). In other words, the writing system consists of a "set of correlations that determine the possible co-occurrences of letter sequences, which eventually result in establishing orthographic representations" (Frost, Siegelman, Narkiss, & Afek, 2013, p. 2). Although some of these regularities in written language are taught explicitly, it seems plausible that children's literacy acquisition is aided by implicit learning through exposure to written language.

Previous research has suggested a link between implicit learning and literacy skills in the typically developing population (e.g. Apfelbaum, Hazeltine, & McMurray, 2013; Arciuli & Cupples, 2006; Arciuli & Simpson, 2012; Frost et al., 2013; Pacton, Fayol, & Perruchet, 2005; Spencer, Kaschak, Jones, & Lonigan, 2014). For example, typically developing children apply orthographic regularities in pseudo-word spelling (e.g. in French, /et/ is more often written as < ette > after -v than after -f), which reflects their implicit knowledge of single letters and letter combinations (Pacton et al., 2005). Similarly, Pacton, Perruchet, Fayol, and Cleeremans (2001) show that French-speaking typically developing children are sensitive to the orthographic constraints of the positions of double consonants (e.g. *xevvu* is more acceptable than *xxevu*). Additionally, correlational studies have established a link between performance on implicit learning tasks and reading in English (Arciuli & Simpson, 2012), reading in Hebrew as a second language (Frost et al., 2013), and a variety of literacy-related skills including oral language, vocabulary and phonological processing (Spencer et al., 2014). Using a linear regression analysis, Ise, Arnoldi, Bartling, and Schulte-Körne (2012) showed that children's performance on a visual artificial grammar learning (AGL) task, a measure of implicit learning which will be explained in more detail below, predicts their performance on a spelling task. Together, the abovementioned studies suggest there is a relationship between implicit learning and (the acquisition of) literacy skills in typical populations.

### 1.2. Implicit learning in dyslexia

A number of studies have investigated the hypothesis that individuals with dyslexia have problems with implicit learning, which affect their literacy skills. Several tasks have been deployed to investigate implicit learning skills in dyslexia. Examples include the serial reaction time (SRT) task (e.g. Deroost et al., 2010; Menghini et al., 2010; Vicari et al., 2005), the alternating SRT task (Hedenius et al., 2013), as well as visual AGL tasks (e.g. Ise et al., 2012; Pothos & Kirk, 2004; Rüsseler, Gerth, & Münte, 2006). Although both the SRT and AGL paradigm are methods used to investigate implicit learning, the type of structure learned in each paradigm differs (greatly). Whereas the SRT measures a motoric response to visual sequences and is stimulus-bound (i.e. no generalization rule can be abstracted from the sequence), the visual AGL measures rule learning from visual input. While numerous studies report implicit learning difficulties in individuals with dyslexia (e.g. Du & Kelly, 2013; Ise et al., 2012; Jiménez-Fernández, Vaquero, Jiménez, & Defior, 2011; Vicari et al., 2005), others do not find evidence for such a deficit (e.g. Deroost et al., 2010; Menghini et al., 2010; Pothos & Kirk, 2004; Rüsseler et al., 2006). Because of these mixed results, Lum, Ullman, and Conti-Ramsden (2013) performed a meta-analysis on 14 studies that investigated implicit learning in individuals with dyslexia using the SRT paradigm. Their results show that implicit sequence learning, as measured by the SRT task, is significantly poorer in people with dyslexia than in non-dyslexic controls (average weighted effect size 0.45,  $p < 0.001$ ). Thus, these results indicate a deficit in implicit visuo-motor learning in dyslexia. In the current study we investigate whether individuals with developmental dyslexia are also affected in visual artificial grammar learning. If individuals with dyslexia have difficulties with implicit learning across the board, group differences should be found using both the SRT and AGL paradigms. However, it could also be the case that poor performance by individuals with dyslexia on the SRT task is due to a specific *motor* learning deficiency, as dyslexia has previously been associated with motor problems (e.g. Fawcett & Nicolson, 1995; Ramus, 2003; Ramus et al., 2003). In that case, one would not necessarily also expect difficulties in the area of visual AGL learning.

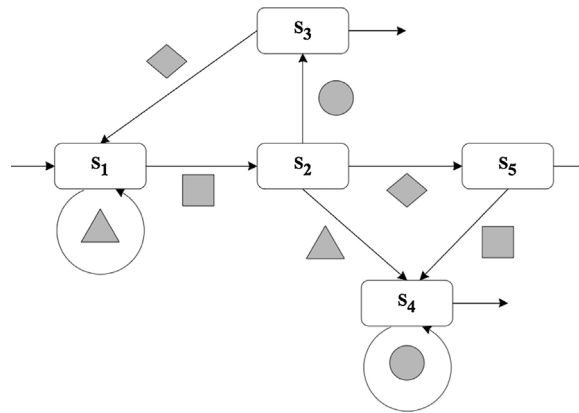


Fig. 1. An illustration of a grammar used in AGL experiments (after Laasonen et al., 2014; original grammar Abrams & Reber, 1988) to generate grammatical sequences (e.g. triangle, square, triangle, circle or square, diamond, square) and ungrammatical sequences (e.g. triangle, diamond, circle, diamond or square, triangle, square).

### 1.3. Visual AGL in dyslexia

Visual AGL refers to an experimental design that investigates participants' ability to implicitly learn rules from mere exposure to sequences of visual stimuli generated by these rules. First introduced by Reber (1967), the visual AGL paradigm involves structured sequences that can be presented as letters or abstract shapes. In visual AGL tasks, sequences are generated on the basis of a (finite state) grammar that determines which stimuli can and cannot succeed one another (Fig. 1). In the example depicted in Fig. 1, from the node  $S_2$ , the sequence can proceed either to  $S_4$  (a triangle) or  $S_5$  (a diamond), but not back to  $S_1$ .

The AGL task typically consists of two phases: a training and a test phase. In the training phase, participants are exposed to a set of structured sequences. Importantly, in the implicit version of the AGL task that is explored in the current meta-analysis, participants are not informed about the presence of the structural rules in the input. The exposure during the training phase can be either passive (i.e. participants are merely exposed to stimuli) or active (i.e. participants are instructed to memorize strings of stimuli and repeat them afterwards).

At the beginning of the test phase, participants are often informed that certain rules guided the presentation of stimuli during the training phase. Subsequently, they are tested on their ability to distinguish sequences that adhere to the artificial grammar (grammatical strings) from sequences that do not (ungrammatical strings). Typically, recognition of grammatical strings is tested within a grammaticality judgment task in which participants are requested to specify whether single strings are grammatical or ungrammatical. Other studies adopt a two-alternative forced choice paradigm, where participants are presented with two strings, one grammatical and one ungrammatical, and have to indicate which of the two strings belongs to the grammar. Performance above chance level (50%) during the test phase is taken as evidence that participants have learned the rules of the underlying grammar.

Several studies using the visual AGL paradigm have reported learning deficits in dyslexia among adults (Kahta & Schiff, 2016; Laasonen et al., 2014) or children (Ise et al., 2012; Pavlidou, Williams, & Kelly, 2009; Pavlidou & Williams, 2014). In each of these studies, this deficit is reflected by significantly lower accuracy scores in the group of individuals with dyslexia as compared to a control group. Several other studies failed to show a significant effect of dyslexia in children (Nigro, Jiménez-Fernández, Simpson, & Defior, 2016) or adults (Pothos & Kirk, 2004; Rüsseler et al., 2006). These differences in degrees of significance might be due to chance (i.e. sampling error), because no direct statistical comparisons were ever made between the studies. However, differences in group effects might also reflect genuine differences between the studies. Here we will speculate on several factors that may help explain such genuine differences between individual studies.

Firstly, the age of the participants may influence the results of individual studies, as several studies have reported that implicit learning improved with age in typical populations (e.g. Arciuli & Simpson, 2011; Maybery, Taylor, & O'Brien-Malone, 1995, but see Jost, Conway, Purdy, & Hendricks, 2011). In a meta-analysis investigating SRT performance, Lum et al. (2013) found smaller differences between participants with and without dyslexia for studies involving adult as opposed to child participants when certain sequences of stimuli were used (second-order sequences) or when the exposure phase was longer. However, no previous studies have examined the developmental trajectory of AGL in individuals with dyslexia.

Secondly, the use of either linguistic or non-linguistic stimuli may influence the difficulty of the task, especially for participants with dyslexia. Linguistic stimuli include visually presented letters (e.g. Ise et al., 2012; Nigro et al., 2016; Rüsseler et al., 2006), whereas non-linguistic experiments have used abstract shapes (e.g. Laasonen et al., 2014; Nigro et al., 2016; Pavlidou et al., 2009; Pothos & Kirk, 2004). The results are mixed: several studies report impaired learning within a AGL task involving linguistic stimuli (i.e., letters, e.g. Ise et al., 2012; Samara, 2013), while others do not find evidence for an effect of dyslexia (e.g. Nigro et al., 2016; Rüsseler et al., 2006). Similarly, studies have yielded mixed results in AGL tasks with non-linguistic stimuli such as abstract shapes (evidence for learning deficits: Laasonen et al., 2014; Pavlidou, Kelly, & Williams, 2010; Pavlidou & Williams, 2014, no evidence for learning deficits: Nigro et al., 2016; Pothos & Kirk, 2004).

Thirdly, the training method potentially affects participants' performance. As mentioned, the training phase generally includes one of two possible methods: passive exposure (Du, 2013; Laasonen et al., 2014; Nigro et al., 2016) or active memorization (e.g. Ise et al., 2012; Rüsseler et al., 2006; Samara, 2013). Active training may lead to better learning, as participants are more focused on the stimuli. Whether the observed differences in results between the studies are genuine or due to chance is one of the questions that the present paper tries to address.

Thus, mixed results exist for the visual AGL paradigm: whereas several studies report significant differences between participants with and without dyslexia (e.g. Ise et al., 2012; Laasonen et al., 2014), others do not (Nigro et al., 2016; Rüsseler et al., 2006). Schmalz, Altoè, and Mulatti (2016) conducted a meta-analysis on a subset of studies investigating visual AGL in dyslexia. They report significantly poorer performance by participants with dyslexia (average weighted effect size 0.47). However, at the same time they are careful in their interpretation and state "[...] publication bias and questionable research practices result in an inflated effect size" (p. 9). As no meta-regression analysis was performed, the authors could not quantitatively explain the differences in effect size between studies.

The primary aim of the present meta-analysis is to extend the findings by Schmalz et al. (2016) to a larger set of (unpublished) studies and determine whether the accumulated evidence indicates a difference in performance on visual AGL between individuals with and without dyslexia. By doing a systematic literature search and by including a number of unpublished studies, we want to provide a more complete update on the strength of the evidence regarding the association between dyslexia and a deficiency in visual artificial grammar learning. Additionally, we aim to investigate the effect of certain methodological variables through a meta-regression analysis. These variables include the age of participants and the nature and complexity of the task used, which potentially help explain heterogeneity in results of individual studies. Factors included in the analysis are (a) age (adult or child participants), (b) stimulus type (letters or abstract shapes), and (c) type of training method (passive exposure or active memorization).

## 2. Method

### 2.1. Literature search

We identified studies published up until September 2016 through searches in PubMed, PsycInfo, ERIC, MEDLINE, CINAHL, and LLBA databases. Additionally, the OATD database was searched for unpublished work in the form of theses and dissertations. A complete overview of keywords used for each of the databases can be found in Appendix A in Supplementary data. In addition to database searches, references of included articles were reviewed. Finally, the CogDevSoc and LinguistList mailing lists were used to inquire whether subscribers knew of unpublished data (deadline response: September 2016).

### 2.2. Study selection

Fig. 2 depicts the selection of studies according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines in a flow diagram (Moher, Liberati, Tetzlaff, & Altman, 2009). Out of all 229 records found, 143 duplicates were removed. Subsequently, one researcher examined the abstracts of 86 unique studies. Studies had to fulfill several selection criteria for inclusion in the present meta-analysis. First, only studies that had administered a visual AGL task were considered. The main reason for a focus on visual AGL studies is to eliminate modality as a possible cause of heterogeneity in results. Second, the experiment had to address implicit learning, i.e., participants were not to be informed of the presence of rules in the input. Third, studies had to include two groups of participants, one group of individuals with dyslexia and one group of non-dyslexic controls.

Fifty-six records were removed after screening the title and abstract because they did not meet the abovementioned selection criteria. An additional 19 records were removed from the sample on the basis of full-article screening, thus leaving eleven records for inclusion in the present review and meta-analysis. Two out of the eleven records (Ise et al., 2012; Nigro et al., 2016) involved two experiments with distinct participant groups that were included separately in the present meta-analysis, resulting in 13 individual effect size calculations. For the remainder of the present meta-analysis, we will refer to the number of individual effect size calculations as the number of studies included ( $N = 13$ ). A second researcher performed identical database searches and assessed all abstracts and full texts. For 28 out of 30 full-text studies the reviewers independently came to the same conclusions regarding inclusion in the present meta-analysis (high inter-rater reliability: Cohen's kappa = 0.851). Consensus on the remaining two records was reached through discussion of the contents.

Note that articles did not have to have been published in peer-reviewed journals in order to be included in our meta-analysis. This means that conference papers or posters, unpublished results and dissertations could be included in the final sample (under the category "other" in Fig. 2). This was done to minimize the possibility of a publication bias. 10 out of 12 records in this category were found through the OATD database (Open Access Theses and Dissertations), of which 2 are included in the final sample (Du, 2013; Samara, 2013). The other two were discovered through personal communication with authors or were presented at the *Interdisciplinary Advances in Statistical Learning* conference (2015, San Sebastian<sup>2</sup>). At the time of analysis, two out of thirteen individual effect sizes included in the present meta-analysis were unpublished.

<sup>2</sup> This resulted in the inclusion of a poster presentation by Kahta and Schiff, which was later published in 2016 and was thus later also found through the systematic literature search.

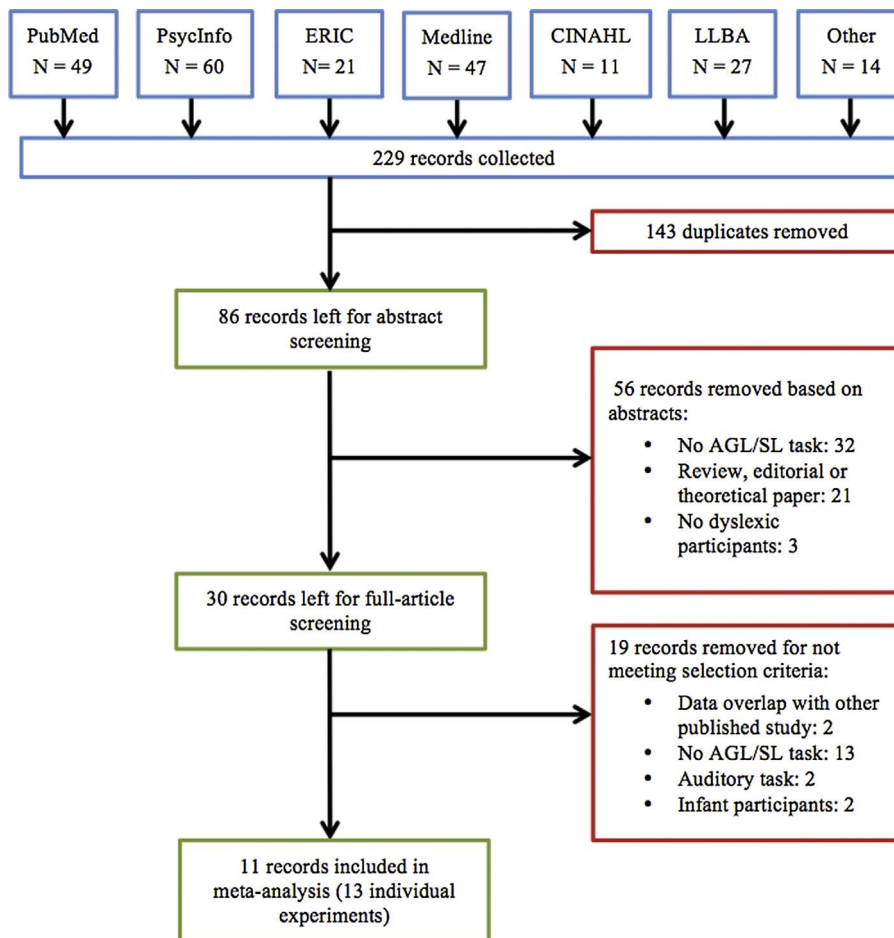


Fig. 2. PRISMA flow diagram depicting the selection of studies.

Note. The category 'other' includes the OATD dissertation database, presentations at conferences, and personal communication. Data overlap: 2 studies that overlapped in data included one bachelor's thesis that contained the same data as a second bachelor's thesis, while the other was Pavlidou's (2010) dissertation of which data was published elsewhere and included in the present meta-analysis

### 2.3. Data extraction and effect size calculations

The standard measure of learning in an AGL task is the percentage of correct responses (i.e. overall accuracy) during the test phase of the experiment. Therefore, the method for comparing the performance of two groups on an AGL task is to test whether the overall accuracy differs between the study and the control group. In order to calculate a single effect size for each of the included individual studies, the mean, standard deviation (*SD*) and sample size of each of the study groups were extracted from the article. If these data were not available from studies themselves, we asked the authors to supply these. Authors provided these data in three cases (Ise et al., 2012; Laasonen et al., 2014; Samara, 2013), which allowed us to calculate single effect sizes for each individual study.

Calculations were made based on raw data for Rüsseler et al. (2006). For the study by Kahta and Schiff (2016) the mean and 95% confidence interval had to be gleaned from figures. This was done using Digitizeit digitizer software (available from <http://www.digitizeit.de/>). Next, 95% confidence intervals were converted into *SD*s according to Eq. (1), which assumes that authors had computed the confidence intervals with the help of the *t*-distribution.<sup>3</sup> Additionally, the studies by Du (2013), Kahta and Schiff (2016) and Samara (2013) did not report the means and *SD* of the participants' average accuracy scores, but separately the means and *SD*s (or 95% confidence intervals in the case of Kahta & Schiff, 2016) of the participants' percentages of correctly accepted and incorrectly accepted test items (i.e. endorsement rates), which we used to calculate the means and *SD*s of the average accuracy scores. In the absence of data on the correlation (over the participants) between percentage of correctly accepted and correctly rejected test items, and in the absence of good evidence from the literature about what a typical correlation could be, we had to make a conservative estimate of the *SD* of the participants' average percentages. If the correlation is 0, then the variance of the average is smaller than each of the reported variances. If the correlation is 1, then the variance of the average is a weighted average of the two reported

<sup>3</sup> For large *n*, *t*<sub>crit</sub> [*n*-1] is close to 1.96.

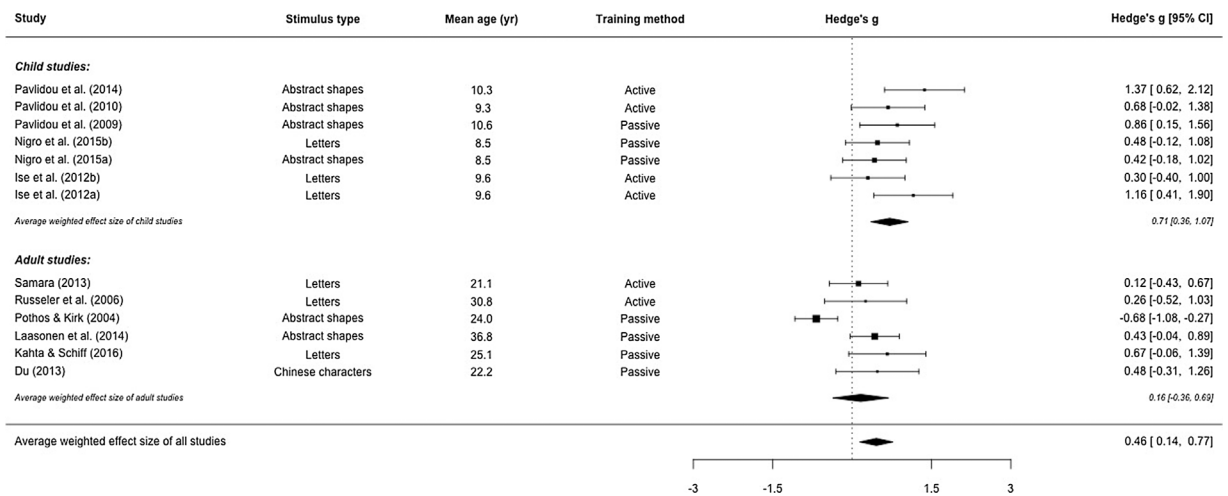


Fig. 3. Forest plot showing the results of 13 studies investigating AGL performance in individuals with and without dyslexia. The average weighted effect size (and its 95% confidence interval) is indicated at the bottom of the figure.

variances. The conservative choice for the estimation is therefore to assume that the correlation is 1, so that our estimate of the SD of the average of the acceptance and rejection scores is given by (2), where  $n_1$  is the number of correct test items,  $n_2$  is the number of incorrect test items,  $SD_1$  is the observed standard deviation of the correctly accepted percentage, and  $SD_2$  is the observed standard deviation of the correctly rejected percentage.

$$SD = \sqrt{n} \times \frac{\text{upperlimit} - \text{lowerlimit}}{\text{tcrit}[n - 1]} \tag{1}$$

$$SD_{\text{average}} = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}} \tag{2}$$

For the study by Pothos and Kirk (2004), the SDs had to be gleaned from their Fig. 4 (p. 71). Additionally, the mean age of participants was not available, but since participants were (under)graduates, most aged between 18 and 30, this study was classified as a study involving adult participants. Appendix B in Supplementary data presents an overview of the extracted data that was used for effect size calculation for each included study. Tables 1 and 2 summarize characteristics of the sample and experimental design of the 13 studies included in the present meta-analysis.

Following data extraction procedures, a single effect size was computed for each individual study, using the “compute.es” package (Del Re, 2014) for R software (R Development Core Team, 2008). In the present meta-analysis, Hedges’ g effect size<sup>4</sup> and 95% confidence intervals summarize the results from each individual study. Positive Hedges’ g values indicate that the control group reached higher accuracy levels on the AGL task compared to the group of individuals with dyslexia, whereas negative values indicate the opposite. The 95% confidence interval provides an estimate of the precision of the study’s effect size: the larger the confidence interval, the poorer the precision. A combination of the “metafor” (Viechtbauer, 2010) and “meta” packages (Schwarzer, 2012) for R software was used to convert the computed individual effect sizes and variances to an average weighted effect size<sup>5</sup> and variance across studies.

### 3. Results

#### 3.1. AGL in dyslexia

Our first goal was to elucidate whether, combining the results from 13 previous studies, individuals with dyslexia perform different from their TD peers on visual AGL tasks. To this end, the effect sizes of all 13 individual studies were combined into a single average weighted effect size using a random-effects model (Hedges & Olkin, 1985). Random-effects models, as opposed to fixed-effect models, allow for variation in true effect sizes between independent studies (Borenstein, Higgins, & Rothstein, 2009). The model was run using the rma.uni function in the “metafor” package with the restricted maximum likelihood (REML) method and the adjustment by Knapp and Hartung (2003) for finite numbers of degrees of freedom. Effect sizes for individual studies and the overall average weighted effect size are presented in Fig. 3. Performance was measured as the overall accuracy score in the test phase of the AGL experiment. Effect sizes ranged from -0.68 to 1.37, with only one effect size in the negative direction (Pothos & Kirk, 2004). All

<sup>4</sup> Hedges’ g is a variation of Cohen’s d that corrects for biases due to small sample sizes (Hedges, 1981).

<sup>5</sup> A standardized effect size was used as opposed to a raw mean difference score, because the raw mean difference scores and pooled standard deviations of individual studies showed large deviations from the overall raw mean difference score (Bond, Wiitala, & Richard, 2003).

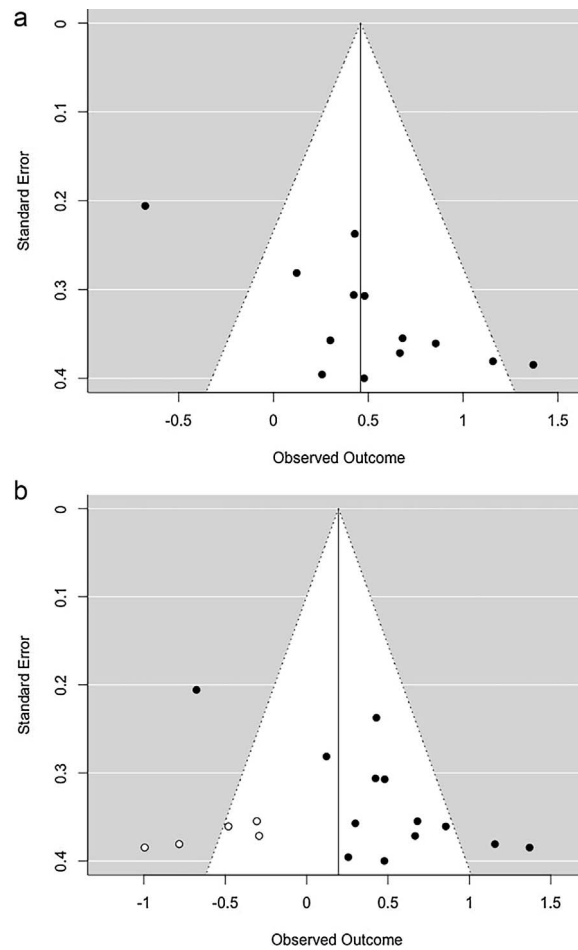


Fig. 4. a) Funnel plot showing the relationship between the standard error and the observed outcome (effect size) of individual studies. b) Funnel plot after controlling for publication bias using the trim and fill method. Solid circles represent included studies, while open circles are imputed missing studies due to publication bias.

**Table 1**  
Overview of study sample characteristics per individual study.

| Study                   | Native language | Sample size |         | Mean age |         | Matching within individual studies |
|-------------------------|-----------------|-------------|---------|----------|---------|------------------------------------|
|                         |                 | Dyslexia    | Control | Dyslexia | Control |                                    |
| Du (2013)               | English         | 12          | 12      | 22.1     | 22.3    | Age, gender, IQ                    |
| Ise et al. (2012)       | German          | 14          | 17      | 9.4      | 9.8     | Age, IQ                            |
| Ise et al. (2012)       | German          | 15          | 15      | 9.4      | 9.8     | Age, IQ                            |
| Kahta and Schiff (2016) | Hebrew          | 14          | 15      | 25.1     |         | MD                                 |
| Laasonen et al. (2014)  | Finnish         | 36          | 35      | 36.1     | 37.5    | Age, gender, IQ, handedness        |
| Nigro et al. (2016)     | Spanish         | 21          | 21      | 8–9      | 8–9     | Age, gender, IQ                    |
| Nigro et al. (2016)     | Spanish         | 21          | 21      | 8–9      | 8–9     | Age, gender, IQ                    |
| Pavlidou et al. (2009)  | English         | 16          | 16      | 10.6     | 10.6    | Age, gender, classroom             |
| Pavlidou et al. (2010)  | English         | 16          | 16      | 9.3      | 9.3     | Age, gender, classroom             |
| Pavlidou et al. (2014)  | English         | 16          | 16      | 10.3     | 10.3    | Age, gender, classroom             |
| Pothos and Kirk (2004)  | English         | 37          | 72      | MD       | MD      | MD                                 |
| Rüsseler et al. (2006)  | German          | 12          | 12      | 28.8     | 32.8    | Gender, IQ, handedness             |
| Samara (2013)           | English         | 25          | 24      | 21.6     | 20.5    | Age, IQ                            |

Note. Both Ise et al. (2012) and Nigro et al. (2016) reported two separate experiments with distinct participant groups; MD = missing data.

other studies report a lower accuracy level for the group of participants with dyslexia than for the control group. Importantly, as mentioned, some of the individual studies report significant differences, whereas others do not. The meta-analysis reveals that, grouping over 13 studies and despite the negative-estimate study, participants with dyslexia performed significantly worse than control participants (average weighted effect size = 0.46, 95% CI [0.14 ... 0.77],  $p = 0.008$ ). Looking at studies involving either



**Table 2**  
Overview of AGL task design per individual study.

| Study                     | Stimulus type      | Sequence length | Training phase | No. of training sequences | Test phase | No. of test sequences |
|---------------------------|--------------------|-----------------|----------------|---------------------------|------------|-----------------------|
| <b>Child participants</b> |                    |                 |                |                           |            |                       |
| Ise et al. (2012)         | Letters (CVCV)     | 5               | Active         | 15                        | GJ         | 24                    |
| Ise et al. (2012)         | Letters (CCCC)     | 5               | Active         | 15                        | GJ         | 24                    |
| Nigro et al. (2016)       | Abstract shapes    | 4               | Passive        | 36                        | 2-AFC      | 32                    |
| Nigro et al. (2016)       | Letters (CVCV)     | 4               | Passive        | 36                        | 2-AFC      | 32                    |
| Pavlidou et al. (2009)    | Abstract shapes    | 2–6             | Passive        | 23                        | GJ         | 32                    |
| Pavlidou et al. (2010)    | Abstract shapes    | 2–6             | Active         | 8                         | GJ         | 20                    |
| Pavlidou et al. (2014)    | Abstract shapes    | 2–6             | Active         | 8                         | GJ         | 20                    |
| <b>Adult participants</b> |                    |                 |                |                           |            |                       |
| Du (2013)                 | Chinese characters | 6–8             | Passive        | 20                        | GJ         | 60                    |
| Kahta and Schiff (2016)   | Letters (CCCC)     | 5–7             | Passive        | 20                        | GJ         | 40                    |
| Laasonen et al. (2014)    | Abstract shapes    | 2–6             | Passive        | 23                        | GJ         | 32                    |
| Pothos and Kirk (2004)    | Abstract shapes    | 2–6             | Passive        | 23                        | GJ         | 32                    |
| Rüsseler et al. (2006)    | Letters (CCCC)     | 4–7             | Active         | 20                        | GJ         | 24                    |
| Samara (2013)             | Letters (CCCC)     | 7               | Active         | 12                        | GJ         | 48                    |

*Note.* Stimulus type: several studies used strings that alternated consonants (C) and vowels (V) (CVCV), whereas others used consonant strings (CCCC); Training phase: Active = memorization, Passive = exposure; Test phase: GJ = grammaticality judgment, 2-AFC = two-alternative forced choice.

child or adult participants separately, we find that the average weighted effect size for child studies is significant ( $N = 7$ , average weighted effect size = 0.71, 95% CI [0.36 ... 1.07],  $p < 0.001$ ), whereas it is not for adult studies ( $N = 6$ , average weighted effect size = 0.16, 95% CI [−0.36 ... 0.69],  $p = 0.461$ ). Before investigating whether the observed difference between the adult effect (0.16) and the child effect (0.71) reflects a genuine decreasing difference between dyslexic and non-dyslexic people as a function of age, we inspect the possibility of publication bias.

### 3.2. Publication bias

To verify the interpretability of the abovementioned findings, we examined the possibility of publication bias in our collected sample of studies. This was initially done through examining a standard funnel plot, which plots the standard error (a measure of study precision) against the effect sizes of the individual studies (Fig. 4a). Generally speaking, in the absence of publication bias, studies should be symmetrically distributed around the average weighted effect size. This distribution takes a funnel shape configuration: studies with high precision are closer to the average weighted effect size, whereas lower precision studies are symmetrically scattered around the average weighted effect size. A linear regression analysis (Egger et al., 1997), using the *metabias* function (Schwarzer, 2012), formally tested the presence of publication bias. It turned out that effect sizes were significantly asymmetrically distributed, skewing to the lower right corner, indicating the presence of a publication bias in our sample ( $t[11] = 4.014$ ,  $p = 0.002$ ).

To evaluate the effect of the publication bias in our sample we approximated what the effect size might be in absence of this bias, using Duval and Tweedie (2000) trim and fill method (*trimfill* function in the “metafor” package using the “L0” estimator for the number of missing studies). Importantly, the trim and fill method can be used to investigate how sensitive the observed effect is to the presence of potential missing studies, but is not meant as a way to calculate the actual values of missing studies (Duval & Tweedie, 2000; Duval, 2005). By using small studies on the positive side of the funnel plot to impute missing studies on the negative side, the trim and fill method estimated that five studies reporting negative findings are missing in our present sample (Fig. 4b). When these five imputed missing studies are added to our dataset of 13 studies, the estimated effect size is considerably reduced and is no longer significantly different from zero (average weighted effect size = 0.20, 95% CI [−0.11 ... 0.50],  $p = 0.205$ ). Note, however, that the trim and fill analysis is known to be a sometimes conservative method for adjusting for publication bias (Peters et al., 2007; Schwarzer, Carpenter, & Rücker, 2010) and the creation of imputed studies can be heavily influenced by a single deviant study, such as the study by Pothos and Kirk (2004) in our sample (e.g. Borenstein et al., 2009, p. 286).<sup>6</sup> Additionally, this method of adjusting results for publication bias makes the assumption that the asymmetry observed in the funnel plot is caused exclusively by publication bias, while another possible cause for funnel plot asymmetry is heterogeneity between studies (Mavridis & Salanti, 2014). Finally, we cannot be certain that the computed missing studies would indeed have been found in the absence of such a bias (Mavridis & Salanti, 2014). Nonetheless, the results of the present meta-analysis on our selected 13 studies are likely to be overly optimistic in the direction of the existence of the main effect, as the effect can well be nulled by unpublished findings.

<sup>6</sup> There exist alternatives to using the L0 method. Using the “R0” estimator instead of “L0”, we find zero missing studies in the present sample, while applying the Copas selection model (Copas, 1999; Copas & Shi, 2000) converges to a fully negative confidence interval. Both of these alternative results are due to the presence of the single large study that reports a negative effect (Pothos & Kirk, 2004). The R0 estimate must be incorrect given the significance of the linear regression analysis, and the Copas result must be incorrect because the other 12 included studies show effects in the positive direction.

**Table 3**

Results from meta-regression analyses exploring the effect of participant and methodological factors on effect sizes of individual studies.

| Moderator                                 | $R^2$ | $Q_{\text{model}}$ | $df$ | $p$    |
|---|-------|--------------------|------|--------|
| 1: Age                                    | 35%   | 5.35               | 1    | 0.041* |
| 2: Training method                        | 1%    | 1.02               | 1    | 0.333  |
| 3: Stimulus type                          | 0%    | 0.02               | 1    | 0.903  |
| 4: Age $\times$ training method           | 21%   | 1.75               | 3    | 0.659  |
| 5: Age $\times$ stimulus type             | 27%   | 1.94               | 3    | 0.364  |
| 6: Training method $\times$ stimulus type | 3%    | 1.04               | 3    | 0.192  |

Note.  $R^2$  = the proportion of the total heterogeneity between studies accounted for by the moderator;  $Q_{\text{model}}$  is the statistic for testing whether the moderator accounts for some of the heterogeneity between studies;  $p$  is the significance for  $Q_{\text{model}}$  being greater than  $df$ .

\*  $p < 0.05$ .

### 3.3. Heterogeneity in findings and meta-regression

The second aim of the present study was to explore several factors that may help account for the heterogeneity (between-studies variability) that appears to be present across different studies investigating AGL in participants with dyslexia. Although the main outcome of the present meta-analysis is probably influenced by the observed publication bias, such a bias is less likely to affect meta-regression analyses, which consider secondary effects.

Cochran's Q-test for heterogeneity was significant ( $Q[12] = 41.07, p < 0.001, I^2 = 71\% [0.49 \dots 0.83]$ ). This result allows us to reject the null hypothesis that all the studies share a common true effect size. As can be seen in Fig. 3, it appears that some factors may influence the effect size of individual studies. As mentioned, the average weighted effect size for child studies is larger than for adult studies (0.71 for child studies vs. 0.16 for adult studies). Thus, we decided to explore the effect of several potential moderator variables on the effect sizes of individual studies through meta-regression analysis.

In preparation for the meta-regression analysis, all three binary moderator variables were centered, i.e. coded as: (a) age  $-0.5$  (child) versus  $+0.5$  (adult), (b) type of stimulus  $-0.5$  (abstract shapes) versus  $+0.5$  (letters), and (c) type of training  $-0.5$  (passive exposure) versus  $+0.5$  (active memorization). Random-effects model meta-regression was used to explore the potential value of these factors in explaining variance in effect size between studies. Since the three moderators are correlated, we first tested each of the three main effects individually in a separate meta-regression model (Table 3). Additionally, we tested each of the three interaction effects individually, in a separate model that included the two relevant main effects (also in Table 3). None of the interaction effects turned out to significantly affect the effect sizes of individual studies, so we did not attempt to construct any more complicated models. As shown in Table 3, the only model that reaches significance in explaining variance between individual studies is model 1: the main effect of age. This model fits 35% of the heterogeneity (which is greater than 0% with  $p = 0.041$ ). When studies had adult participants, as opposed to child participants, effect sizes were smaller, reflecting a smaller difference between participants with and without dyslexia. None of the other main or interaction effects were found to significantly fit the heterogeneity between studies. To the extent to which a p-value of 0.041 can be considered statistically significant in this exploration of six possible effects (without correction for multiple testing), we can conclude that the difference between the observed adult and child effect sizes (0.16 and 0.71) indeed reflects a genuine difference between the two ages in the population.

## 4. Discussion

In the present study, we used meta-analysis and meta-regression to quantitatively review previous research on visual AGL in dyslexia. Our first goal was to elucidate whether the combined findings of thirteen previous studies provide evidence for a difference in visual AGL between individuals with and without dyslexia. The average weighted effect size computed from these individual visual AGL studies, reflecting results from 255 participants with dyslexia and 292 control participants, was found to be moderate and statistically significant. If our 13 selected studies were a sample randomly drawn from an imagined infinite set of possible studies, this finding would indicate that, overall, non-dyslexic people outperform people with dyslexia on visual AGL. Our results would then corroborate the earlier analysis in Schmalz et al. (2016) and strengthen these findings by involving a larger sample of studies (13 instead of 9). Taken together with the meta-analysis of SRT studies by Lum et al. (2013), these results would suggest a general implicit learning deficit in individuals with dyslexia.

Importantly, however, it seems plausible that these results have been influenced by a publication bias in the field of artificial grammar learning in dyslexia (see Schmalz et al., 2016). After conservatively controlling for publication bias, the computed effect size was no longer significant, and the results of the main effect of the present meta-analysis should therefore be regarded as unreliable. Large-scale future studies are needed to confirm the presence of a difference in performance on visual AGL between participants with and without dyslexia.

Extending the previously published meta-analysis by Schmalz et al. (2016) further, the present study aimed to explain the heterogeneity in results of individual studies by investigating the effect of certain methodological variables through a meta-regression analysis. This analysis revealed that the only moderator that (moderately, i.e. without correction for multiple tests) reached significance was the main effect of age: there were smaller differences between dyslexia and control groups for those studies that involved adult participants as opposed to child participants. This is an indication that the implicit learning deficit might be more

pronounced in children with dyslexia than in adults with dyslexia, since similar effects of age have been found in the meta-analysis investigating implicit learning in individuals with dyslexia using the SRT task (Lum et al., 2013). In line with the interpretation of their results, a possible explanation is that adults make use of compensatory processes (e.g. visual processing, pattern recognition, attentional resources, declarative memory) that enhance performance on visual AGL tasks. Another potential explanation for the age effect lies in the selection of participants. Whereas most studies with adult participants involved university students, child studies selected their participants from a broader population of primary school children. The performance of university students with dyslexia may not be representative of the whole population of adults with dyslexia, as these high-achieving individuals with dyslexia may have more developed compensatory mechanisms. This in turn may result in a smaller difference between the performance of adults with and without dyslexia. We want to note that this effect of age should be interpreted with caution, as it seems to be largely driven by one study that reports better performance in adults with dyslexia than in controls (Pothos & Kirk, 2004,  $g = -0.68$ ). Thus, future research should examine the possibility of an age effect in visual AGL in dyslexia in further detail by selecting adult participants with dyslexia from all educational levels and comparing them to children on the same visual AGL task.

Although the present meta-analysis suggests that visual artificial grammar learning might be poorer in dyslexia relative to non-dyslexic individuals overall, these results cannot address the issue of causality between implicit statistical learning and literacy skills in this population. Future longitudinal studies are needed to investigate the potential causal link between implicit statistical learning and literacy skills in individuals with and without dyslexia.

Additionally, several factors that could influence the effect sizes of individual studies were not included in the present meta-analysis due to the relatively small number of studies. One such factor is the complexity of the underlying grammar. The level of complexity potentially plays a role in whether participants are able to learn the underlying structure. In fact, a recent meta-analysis of AGL studies with typical populations showed that, indeed, there is a significant correlation between grammar complexity and learners' task performance (Schiff & Katan, 2014). Also related to the difficulty of the task at hand are factors such as the length of the sequences and the amount of exposure to these sequences. Whereas some studies use a fixed sequence length of 4 (Nigro et al., 2016), 5 (Ise et al., 2012), or 7 (Samara, 2013), other studies use sequences of varying lengths (between 2 and 6 (e.g. Pothos & Kirk, 2004; Pavlidou & Williams, 2014), 4 and 7 (Rüsseler et al., 2006) or 6 and 8 (Du, 2013) individual items). Similarly, whereas some studies include only 69 instances of a grammatical string (e.g. Laasonen et al., 2014; Pavlidou et al., 2009), others include as many as 108 instances (Nigro et al., 2016). Another factor worth investigating is the severity of dyslexia in individual participants, as this may be related to the severity of the deficit in implicit statistical learning. Finally, the modality (visual versus auditory) in which the stimuli are presented may affect the learnability of the grammar for individuals with dyslexia. Future research should investigate the potential effect of the abovementioned factors to gain further understanding of what types of methodological characteristics increase or decrease an AGL task's learnability for individuals with and without dyslexia.

## Acknowledgements

We thank Merel Hardeman for her help with the database searches and for her assessments of full-texts as a second reviewer. The Netherlands Organization for Scientific Research supported this study for (NWO Innovational Research Incentives Scheme Vidi).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ridd.2017.09.006>.

## References

- Abrams, M., & Reber, A. S. (1988). Implicit learning: Robustness in the face of psychiatric disorders. *Journal of Psycholinguistic Research*, 17(5), 425–439.
- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Text Revision (DSM-IV-TR).
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: DSM-V.
- Apfelbaum, K. S., Hazeltine, E., & McMurray, B. (2013). Statistical learning in reading: Variability in irrelevant letters helps children learn phonics skills. *Developmental Psychology*, 49(7), 1348–1365.
- Arciuli, J., & Cupples, L. (2006). The processing of lexical stress during visual word recognition: Typicality effects and orthographic correlates. *Quarterly Journal of Experimental Psychology*, 59(5), 920–948.
- Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, 14(3), 464–473.
- Arciuli, J., & Simpson, I. C. (2012). Statistical Learning Is Related to Reading Ability in Children and Adults. *Cognitive Science*, 36(2), 286–304.
- Bond, C. F., Jr., Wiitala, W. L., & Richard, F. D. (2003). Meta-analysis of raw mean differences. *Psychological Methods*, 8(4), 406.
- Borenstein, M. H., Higgins, L. V., & Rothstein, J. P. T. (2009). *Introduction to meta-analysis*. Chichester, England: Wiley.
- Copas, J., & Shi, J. Q. (2000). Meta-analysis: Funnel plots and sensitivity analysis. *Biostatistics*, 1(3), 247–262.
- Copas, J. (1999). What works?: Selectivity models and meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 162(1), 95–109.
- Del Re, A. C. (2014). *Package 'compute.es' for R software: Compute effect sizes*. <https://cran.r-project.org/web/s/compute.es/compute.es.pdf>.
- Deroost, N., Zeischka, P., Coomans, D., Bouazza, S., Depessemier, P., & Soetens, E. (2010). Intact first- and second-order implicit sequence learning in secondary-school-aged children with developmental dyslexia. *Journal of Clinical and Experimental Neuropsychology*, 32(6), 561–572.
- Du, W., & Kelly, S. W. (2013). Implicit sequence learning in dyslexia: A within-sequence comparison of first- and higher-order information. *Annals of Dyslexia*, 63(2), 154–170.
- \*Du, W. (2013). *Associative implicit learning in adult dyslexic readers*. University of Strathclyde Doctoral dissertation.
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463.
- Duval, S. (2005). *The trim and fill method*. *Publication Bias in Meta-analysis: Prevention, Assessment, and Adjustments* 127–144.

- 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.
- Facoetti a Paganoni, P., & Lorusso, M. L. (2000). The spatial distribution of visual attention in developmental dyslexia. *Experimental Brain Research*, *132*(4), 531–538.
- Fawcett, A. J., & Nicolson, R. I. (1995). Persistent deficits in motor skill of children with dyslexia. *Journal of Motor Behavior*, *27*(3), 235–240.
- Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science*, *24*(7), 1243–1252.
- Gathercole, S., & Alloway, T. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology*, *93*(3), 265–281.
- 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*, 336–360.
- Hedenius, M., Persson, J., Alm, P. a., Ullman, M. T., Howard, J. H., Howard, D. V., et al. (2013). Impaired implicit sequence learning in children with developmental dyslexia. *Research in Developmental Disabilities*, *34*(11), 3924–3935.
- Hedges, L., & Olkin, I. (1985). *Statistical models for meta-analysis*. New York: Academic Press.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*(2), 107–128.
- \*Ise, E., Arnoldi, C. J., Bartling, J., & Schulte-Körne, G. (2012). Implicit learning in children with spelling disability: Evidence from artificial grammar learning. *Journal of Neural Transmission*, *119*(9), 999–1010.
- Jiménez-Fernández, G., Vaquero, J. M. M., Jiménez, L., & Defior, S. (2011). Dyslexic children show deficits in implicit sequence learning, but not in explicit sequence learning or contextual cueing. *Annals of Dyslexia*, *61*(1), 85–110.
- Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology*, *77*(1), 30–60.
- Jost, E., Conway, C. M., Purdy, J. D., & Hendricks, M. A. (2011). Neurophysiological correlates of visual statistical learning in adults and children. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2526–2531). Austin, TX: Cognitive Science Society.
- \*Kahta, S., & Schiff, R. (2016). Implicit learning deficits among adults with developmental dyslexia. *Annals of Dyslexia*, *66*(2), 235–250.
- Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184.
- Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, *22*(17), 2693–2710.
- \*Laasonen, M., Väre, J., Oksanen-Hennah, H., Leppämäki, S., Tani, P., Harno, H., et al. (2014). Project DyAdd: Implicit learning in adult dyslexia and ADHD. *Annals of Dyslexia*, *64*(1), 1–33.
- Lum, J. G., Ullman, M. T., & Conti-Ramsden, G. (2013). Procedural learning is impaired in dyslexia: Evidence from a meta-analysis of serial reaction time studies. *Research in Developmental Disabilities*, *34*(10), 3460–3476.
- Mann, V. A., Shankweiler, D., & Smith, S. T. (1984). The association between comprehension of spoken sentences and early reading ability: The role of phonetic representation. *Journal of Child Language*, *11*(3), 627–643.
- Mavridis, D., & Salanti, G. (2014). Exploring and accounting for publication bias in mental health: A brief overview of methods. *Evidence-based Mental Health*, *17*(1), 11.
- Maybery, M., Taylor, M., & O'Brien-Malone, A. (1995). Implicit learning: Sensitive to age but not IQ. *Australian Journal of Psychology*, *47*(1), 8–17.
- Melby-Lervåg, M., Lyster, S. A., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352.
- Menghini, D., Finzi, A., Benassi, M., Bolzani, R., Facoetti, A., Giovagnoli, S., et al. (2010). Different underlying neurocognitive deficits in developmental dyslexia: A comparative study. *Neuropsychologia*, *48*(4), 863–872.
- Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, *1*(31), 1–9.
- Moher, D., Liberati, A., Tetzlaff, & Altman, D. G. (2009). Academia and clinic annals of internal medicine preferred reporting items for systematic reviews and meta-analyses. *Annals of Internal Medicine*, 264–269.
- Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: Reuniting the developmental disorders? *Trends in Neurosciences*, *30*(4), 135–141.
- \*Nigro, L., Jiménez-Fernández, G., Simpson, I. C., & Defior, S. (2016). Implicit learning of non-linguistic and linguistic regularities in children with dyslexia. *Annals of Dyslexia*, *66*(2), 1–17.
- Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit Learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, *130*(3), 401.
- Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development*, *76*(2), 324–339.
- \*Pavlidou, E. V., & Williams, J. M. (2014). Implicit learning and reading: Insights from typical children and children with developmental dyslexia using the artificial grammar learning (AGL) paradigm. *Research in Developmental Disabilities*, *35*(7), 1457–1472.
- \*Pavlidou, E. V., Williams, J. M., & Kelly, L. M. (2009). Artificial grammar learning in primary school children with and without developmental dyslexia. *Annals of Dyslexia*, *59*(1), 55–77.
- \*Pavlidou, E. V., Kelly, L. M., & Williams, J. M. (2010). Do children with developmental dyslexia have impairments in implicit learning? *Dyslexia*, *16*(2), 143–161.
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine*, *26*(25), 4544–4562.
- \*Pothos, E. M., & Kirk, J. (2004). Investigating learning deficits associated with dyslexia. *Dyslexia*, *10*(1), 61–76.
- R Development Core Team (2008). *R: A language and environment for > statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. [ISBN 3–900051-07–0 URL <http://www.R-project.org>.
- \*Rüsseler, J., Gerth, I., & Münte, T. F. (2006). Implicit learning is intact in adult developmental dyslexic readers: Evidence from the serial reaction time task and artificial grammar learning. *Journal of Clinical and Experimental Neuropsychology*, *28*(5), 808–827.
- Ramus, F., Pidgeon, E., & Frith, U. (2003). The relationship between motor control and phonology in dyslexic children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*(5), 712–722.
- Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, *13*(2), 212–218.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863.
- Rispens, J., & Katan, P. (2007). Subject-verb agreement and phonological processing in developmental dyslexia and specific language impairment (SLI): A closer look. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists*, *42*(3), 293–305.
- Rispens, J., Roeleven, & Koster, S. C. (2004). Sensitivity to subject-verb agreement in spoken language in children with developmental dyslexia. *Journal of Neurolinguistics*, *17*(5), 333–347.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.
- \*Samara, A. (2013). *Statistical learning of orthographic patterns in typically developing and dyslexic populations*. Doctoral dissertation, Bangor University.
- Schiff, R., & Katan, P. (2014). Does complexity matter? Meta-analysis of learner performance in artificial grammar tasks. *Frontiers in Psychology*, *5*, 1–10.
- Schmalz, X., Altoè, G., & Mulatti, C. (2016). Statistical learning and dyslexia: A systematic review. *Annals of Dyslexia*, 1–16.
- Schwarzer, G., Carpenter, J., & Rücker, G. (2010). Empirical evaluation suggests Copas selection model preferable to trim-and-fill method for selection bias in meta-analysis. *Journal of Clinical Epidemiology*, *63*(3), 282–288.
- Schwarzer, G. (2012). *Package 'meta' for R software: General package for meta-analysis*. <https://cran.r-project.org/web/packages/meta/meta.pdf>.
- Shankweiler, D., Crain, S., Katz, L., Fowler, A. E., Liberman, A. M., Brady, S. A., et al. (1995). Cognitive profiles of reading-disabled children: Comparison of language skills in phonology, morphology, and syntax. *Psychological Science*, *6*(3), 149–156.
- Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2014). Statistical learning is related to early literacy-related skills. *Reading and Writing*, *28*(4), 467–490.
- Stein, J. F., & Walsh, V. (1997). To see but not to read; The magnocellular theory of dyslexia. *Trends in Neurosciences*, *20*(4), 147–152.
- Stein, C. L., Cairns, H. S., & Zurif, E. B. (1984). Sentence comprehension limitations related to syntactic deficits in reading-disabled children. *Applied Psycholinguistics*, *5*(04), 305–322.

- Swanson, H. L., & Jerman, O. (2007). The influence of working memory on reading growth in subgroups of children with reading disabilities. *Journal of Experimental Child Psychology*, 96(4), 249–283.
- Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience*, 5(9), 721–728.
- Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, 41(3), 399–433.
- Vicari, S., Finzi, A., Menghini, D., Marotta, L., Baldi, S., & Petrosini, L. (2005). Do children with developmental dyslexia have an implicit learning deficit? *Journal of Neurology, Neurosurgery and Psychiatry*, 76(10), 1392–1397.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48.
- Waltzman, D., & Cairns, H. (2000). Grammatical knowledge of third grade good and poor readers. *Applied Psycholinguistics*, 21(2), 263–284.
- Wijnen, F. (2013). Acquisition of linguistic categories: Cross-domain convergence. In J. Bolhuis, & M. Everaert (Eds.). *Birdsong, speech, and language: Exploring the evolution of mind and brain* (pp. 157–177). Cambridge, MA: MIT Press.
- Yu, C. (2008). A statistical associative account of vocabulary growth in early word learning. *Language Learning and Development*, 4(1), 32–36.