The short- and long-term predictions of reading accuracy and speed from paired-associate learning

Poulsen, Mads; Elbro, Carsten

The short and long-term prediction of reading accuracy and speed from paired associate learning

Mads Poulsen and Carsten Elbro
University of Copenhagen

Mads Poulsen, PhD, Department of Nordic Studies and Linguistics, University of Copenhagen, Copenhagen, Denmark, m.poulsen@hum.ku.dk; Carsten Elbro, PhD, Department of Nordic Studies and Linguistics, University of Copenhagen, Copenhagen, Denmark, ce@hum.ku.dk.

Correspondence concerning this article should be addressed to Mads Poulsen, Department of Nordic Studies and Linguistics, Mads Poulsen, Emil Holms Kanal 2, Copenhagen S. DK-2300, Denmark. E-mail: m.poulsen@hum.ku.

The short and long-term prediction of reading accuracy and speed from paired associate learning

Abstract

Cross-sectional studies have established that performance on paired-associate learning tasks (PAL) is associated with reading performance. There are good reasons to expect such a relationship because learning to read involves learning the sounds of individual letters and possibly also sounds of strings of letters (e.g. spelling patterns). However, results from longitudinal studies have been mixed. A closer look at these studies suggests that PAL may be related to development of accuracy rather than speed in reading. This suggestion was investigated directly in the present longitudinal study. The study followed 137 students from Grade 0 to 5. In Grade 0 they completed measures of PAL, letter knowledge, phoneme awareness, and RAN. In Grades 1 and 5 decoding accuracy was measured with the addition of decoding speed in Grade 5. PAL in Grade 0 was found to be a unique predictor of decoding accuracy in Grades 1 and 5 after control for Grade 0 letter knowledge, phoneme awareness, and RAN. PAL in Grade 0 even contributed to Grade 5 decoding accuracy after additional control for Grade 1 decoding. Zero-order correlations between PAL and Grade 5 decoding speed were non-significant and close to zero. The results indicate that PAL measures a trait that may influence reading development over a substantial amount of time. Possible roles of

PAL in decoding development over time are discussed, e.g., how verbal learning may be a core component in the acquisition of associations between letter patterns ('spelling patterns') and their pronunciation.

Introduction

Learning to read written words involves learning to associate orthographic signs with sound, from the letter level to the word level. For this reason, the ability to store and retrieve word-like sounds has attracted attention as a possible marker of reading difficulties and a possible predictor of word reading development. The ability has been measured with paired-associate learning tasks (PAL), which have been shown to correlate with reading cross-sectionally (Georgiou, Liu, & Xu, 2017; Litt, de Jong, van Bergen, & Nation, 2013; Warmington & Hulme, 2012) and to differ between dyslexics and controls (Elbro & Jensen, 2005; Litt & Nation, 2014; Mayringer & Wimmer, 2000; Messbauer & de Jong, 2003, 2006; Wimmer, Mayringer, & Landerl, 1998).

However, the results with regard to longitudinal prediction of reading development have been mixed. Some studies have found longitudinal correlations (Horbach, Scharke, Cröll, Heim, & Günther, 2015; Nielsen & Juul, 2015; Poulsen, Juul, & Elbro, 2015), but Lervåg, Bråten, and Hulme (2009) found that PAL was not an independent predictor of reading after control for letter knowledge and phoneme awareness (and other precursors). The studies vary in whether they primarily measure reading accuracy or speed. The present study was concerned with the possibility that PAL is more strongly related to reading accuracy than speed, which could explain the pattern of published results.

Why may paired-associate learning be related to reading?

In the PAL tasks that are typically used in reading research, participants have to learn a set of associations between visual stimuli (e.g. drawings or non-familiar characters) and verbal responses

(e.g. nonwords). Usually, the participant has to learn a set of associations involving short words or nonwords. The typical PAL task requires the participant to store multiple associations between visual and spoken representations and to keep them in memory for long enough to be retrieved in any new order.

Such verbal learning may be useful at several stages of reading development. At the first stage, the sounds and names of letters have to be learned. Letter names are largely new and meaningless to the beginning reader. At later stages, verbal learning may also be important for the acquisition of sound patterns of sequences of letters. This is particularly the case for irregular (or "deep") orthographies. Such orthographies have been conservative and resisted changes in the spoken language which means that the relations between spelling and sound have become obscure – in more or less systematic ways. As a consequence, readers need to learn to associate strings of letters with novel strings of sounds in order to decode written words. Consider, for example, how adding more letters changes the most likely pronunciation of the letter *o* (Elbro, 2005):

-o- -> /ɒ/ (hot, mob)

-ou- -> /aʊ/ (doubt, about)

-ough -> /əʊ/ (though, dough)

-ought -> /ɔː/ (ought, thought)

Indeed, the pronunciation of *-ought* (as in *bought, brought, thought* etc.) is completely predictable even though the constituent letters have mostly non-standard pronunciations. Notice how such learning of letter patterns requires verbal learning, in the above example the sound patterns /aʊ/, /əʊ/ and /ɔt/ have to be learned. Hence, verbal learning may continue to be an underlying source of variation in decoding development beyond the first stage. The importance of verbal learning would be expected to continue over several years of decoding development especially in deep

orthographies, which contain multiple specific relations between letter sequences and sound sequences.

In addition, the development of whole-word recognition may also involve some kind of verbal learning, at least at the initial stage. When beginners attempt to read a word that they have not seen in print before they may, at first, form an assembled (or 'spelling') pronunciation of the word – before they reach the correct pronunciation and recognize the word (Elbro & de Jong, 2017; Elbro, de Jong, Houter, & Nielsen, 2012). During initial reading development, such spelling pronunciations may form stepping stones between strings of recoded letter-sounds and already known spoken words. However, the spelling pronunciations will have to be learned as variant pronunciations of already known words. In short, there are good reasons to expect that PAL taps aspects of verbal learning that are also required for learning to read words. A more general, theoretical account of reading development as verbal learning has recently been outlined by Elbro and de Jong (2017).

There is some agreement between researchers that storage and retrieval of phonological forms is central to the relationship between PAL and reading: PAL tasks with visual responses (e.g. drawn pseudo-letters) have been found to be only weakly associated with reading (Hulme, Goetz, Gooch, Adams, & Snowling, 2007; Litt et al., 2013; Litt & Nation, 2014). Some researchers have further emphasized the importance of the cross-modal aspect of the task, which clearly mirrors important aspects of reading development, from visual letter or word forms to their spoken counterparts. Evidence for a possible special role of cross-modal associations comes from experiments showing that a PAL task with visual stimuli and verbal responses explained variance in reading beyond a PAL task with both verbal stimuli and responses (Hulme et al., 2007). However, Litt and colleagues found that it was the verbal response part of the task that was correlated with reading (Litt et al., 2013; Litt & Nation, 2014). Furthermore, Litt and Nation found that the difference between

dyslexics and controls in a visual stimuli-verbal response PAL task could be explained by the dyslexics' difficulty with learning the verbal responses before the visual stimuli were introduced. These results indicate that the relationship between PAL and reading is driven by the verbal learning demands of the task rather than the cross-modal pairing demands.

PAL as a longitudinal predictor of decoding accuracy or speed?

As mentioned, relatively unequivocal results have emerged from cross-sectional studies of the relationship between PAL and reading, Yet, it remains a challenge that previous longitudinal studies have provided mixed results with regard to PAL's prediction of later decoding. It is possible, however, that the type of measure of reading outcomes could explain differences in the prediction value of PAL across studies. In particular, we examined the idea that PAL is related to decoding accuracy rather than speed.

When decoding *accuracy* has been the outcome measure, unique contributions from PAL to later decoding have been reported (Georgiou et al., 2017; Poulsen, Nielsen, Juul, & Elbro, 2017). In addition, Nielsen and Juul (2015) found that PAL was the strongest Grade 0 predictor of Grade 5 spelling accuracy.

We know of no studies that directly test the relationship between PAL and decoding *speed*, but several studies used decoding *fluency* as an outcome measure. Decoding *speed* is a matter of how fast words are decoded successfully. Reading times of *incorrectly* read words are not considered because they have an uncertain relationship with reading ability. They may reflect how students deal with not being able to decode accurately, some take their time, others just shoot and miss. *Fluency* is typically measured as number of correct answers pr. time unit. Fluency is thus a complex construct that depends on both *accuracy* and *speed*. Longitudinal studies of PAL with reading fluency as an outcome measure have found mixed results. In a three-year Norwegian study, Lervåg et al. (2009) did not find a robust unique prediction from PAL measured in Grade 1 of decoding

fluency measured in Grades 2 to 4, when letter knowledge and phoneme awareness were controlled. Norwegian uses a transparent orthography, so students can be expected to reach ceiling with respect to accuracy quickly, so measures of fluency probably reflected speed to a substantial degree (Lervåg, personal communication). In a one-year German longitudinal study, Horbach et al. (2015) found that a PAL task differentiated early readers from non-readers in kindergarten. PAL uniquely predicted decoding fluency one year later beyond phoneme awareness and rapid naming in the group of students who were non-readers in kindergarten. However, the unique longitudinal prediction was not significant in the group of students who were readers already in kindergarten. German is also a transparent orthography, and students can be expected to reach an accuracy ceiling quickly. Hence, it is possible that for students who were already readers in kindergarten, the fluency measure would mostly be a measure of decoding speed. For the kindergarten non-readers, on the other hand, a fluency measure might still contain substantial variance due to accuracy. If PAL is primarily related to accuracy, then this potential difference in how the fluency measure is influenced by accuracy and speed could explain why fluency was predicted by PAL in the group of kindergarten non-readers, but not in the group of kindergarten readers.

In accordance with the above reasoning, several cross-sectional studies have found PAL differences between dyslexics and controls in transparent orthographies (Mayringer & Wimmer, 2000; Messbauer & de Jong, 2003, 2006; Wimmer et al., 1998). Since some problems with accuracy continue to characterize dyslexic reading (Mayringer & Wimmer, 2000; Wimmer et al., 1998), these studies may reflect a link between PAL and reading accuracy.

In a cross-sectional study, Litt and colleagues (2013) found some indications that PAL was more closely related to decoding accuracy than speed. PAL was uniquely correlated with decoding accuracy ($\beta = .30$), but not fluency ($\beta = .14-.17$) after controlling for phoneme deletion and RAN in a sample of 7-11 year old British students. The result should be interpreted cautiously because the

difference in relationship was not significant. On the other hand, the decoding fluency measure was probably not a clean measure of speed given that the sample displayed plenty of accuracy variance on an untimed decoding measure. The present study measured accuracy and speed separately.

In summary, PAL is a potentially interesting predictor of word reading development because the ability to encode, store and recall phonological material could be useful at several stages of reading development. There is a fairly strong evidence for a correlation between PAL and decoding, but results from longitudinal studies have been mixed. We suggest that the mixed results may be a consequence of different outcome measures of reading – and that PAL may be more strongly related to decoding accuracy than speed.

The present study

The overarching goal of the present study was to add to our understanding of the cognitive prerequisites of reading development. More specifically, the aim was to investigate the possibility that the ability to encode, store, and/or recall phonological material predicts the development of reading ability, but possibly only the development of decoding accuracy, not speed. The study was conducted in a sample of Danish students who were followed from Grade 0 to Grade 5.

The research questions were as follows:

- Does PAL measured in late Grade 0 predict early decoding (Grade 1) beyond established predictors, that is, beyond letter knowledge, phoneme awareness, and rapid automatized naming (RAN)? This is effectively a question of concurrent prediction given the short time span.

- Does PAL predict further decoding development (in Grade 5) beyond early decoding (autoregressor) and standard predictors?

- Does the prediction from PAL depend on whether decoding skills are measured in terms of accuracy or speed?

An important feature of the study thus is that we measured decoding accuracy and speed independently, rather than as a composite fluency measure. This presupposes that decoding skill in Grade 5 can be meaningfully separated into accuracy and speed. Therefore, we also analyzed the relationship between our different decoding accuracy and speed measures, and whether decoding accuracy and speed contributes independently to reading comprehension. A Grade 5 fluency measure was included for the sake of comparison with previous studies, but the study focuses on the separate measures of accuracy and speed.

Method

Participants and design

One hundred and eighty-seven preschoolers from four schools in Copenhagen participated in the initial round of testing in March/April of their preschool year in 2009. They completed word list decoding tests in September of Grade 1. In September of G5, 137 of these students completed a final battery of decoding tests. These 137 students had a mean age of 6;10 (SD = 4 months) at the time of Grade 0 testing. 53% were boys, and 4% were bilingual. Some data from these students have previously been reported (Elbro et al., 2012; Juul, Poulsen, & Elbro, 2014; Nielsen & Juul, 2015; Poulsen et al., 2015; Poulsen et al., 2017).

Danish has a deep orthography (Elbro, 2005; Seymour, Aro, Erskine, & network, 2003). Danish children typically start in Grade 0 the year they turn six. The aim of Grade 0 is to prepare the children for school. Activities promoting letter knowledge and phoneme awareness are common in Grade 0, but students are not expected to learn to read before Grade 1. In the present study most of the students were unable to read at the end of Grade 0 (60% of the students were unable to read more than one word from a list of 16 simple 2-3 letter words). Formal reading instruction typically starts in the beginning of Grade 1. School begins early in August. Already by September, some progression in reading can be expected, which allows measuring actual reading without very strong

floor effects. Thus, September of Grade 1 represents a first possible measurement time for a reading autoregressor.

Materials

Grade 0 (May)

PAL. In this task, the students had to learn nonword names (*sput*, *laf*, and *ky*) of three non-familiar cartoon animals. Pilot testing revealed that this was a very difficult task in this age group. As a consequence, the following steps were taken to encourage the children to stay on task. The students were told that they were going to be introduced to some animals with weird names that they were going to learn. The hand-drawn animals were introduced in a booklet one at a time, each in the context of a small story that highlighted a special characteristic. After the story about an animal, the students were asked questions about the name and characteristic of the animal. The purpose of the story and the questions was to engage the students, make the animals memorable, and to give the experimenter and the students a reason for repeating the names of the animals several times in a natural way. The learning trials started after the first two animals had been introduced. On each trial, the two animals were presented on a page in alternating order, and the students were asked to name the animals. The test administrator guided the naming order by covering one animal at a time and asking "What is the name of the animal?" If the students made mistakes, they were corrected and asked to repeat the names to allow learning. When a student named the two animals correctly in three consecutive trials, the third animal was introduced in the same way as the first two. Naming trials with all three animals continued for a total of 15 trials (including trials with two animals). If all three animals were named correctly three times in a row, the task was terminated and the remaining naming opportunities were scored as correct. The score was the proportion of correctly named animals out of a maximum of 42. The reliability of this measure is uncertain (cf. Poulsen et al., 2015, for discussion).

RAN. The students completed two tests of rapid automatized naming adapted from Denckla and Rudel (1976), one with digits and one with objects. In the digits version of the task, students were asked to name five rows of 10 digits accurately and as fast as possible. The digits 1-5 were used and repeated two times in each row. In the objects version, the students had to name four rows of eight pictograms. There were four different pictograms (*sol* "sun", *saks* "scissors", *hjerte* "heart", *blomst* "flower"). Students were familiarized with the digits and pictograms before the tasks commenced. The score for the task was the number of items named correctly per second. The correlation between the two measures was $r = .63$ in the original sample.

Letter knowledge. The students completed two tests of letter knowledge. In an individually administered letter naming task (Elbro, Borstrøm, & Petersen, 1998), they were asked to name the 29 upper case letters in the Danish alphabet. In a group administered letter identification task (Borstrøm & Petersen, 2006), the students were asked to identify each of the 29 lower case letters based on the letter names. For each spoken letter name, they were asked to circle the correct letter among six alternatives. Cronbach's alpha was .92 and .82 for letter naming and identification, respectively, in the original sample.

Phonemic awareness. The students completed two tests of phonemic awareness. In an individually administered phoneme deletion task (adapted from Elbro et al., 1998), for each of 18 items, the students were asked to tell what word was left after deleting a specific phoneme from the beginning, middle, or end of the stimulus word. In a group-administered phoneme matching task, the students were asked for each item to select a picture from four alternatives that started with the same sound as a target picture. Cronbach's alpha was .91 and .80 for the deletion and matching tasks, respectively, in the original sample.

Grade 1 (September)

G1 list decoding accuracy. The students were asked to read aloud six lists of four real words. All the words were short (CVC, CV, or VC) orthographically regular words, i.e. all letters in these words are pronounced with their most frequent letter sound. The score was the percentage of the 24 words that were read correctly. The lists were a subset of a larger number of lists that included more difficult (e.g. longer and irregular) words, nonwords and sentences. The particular lists were chosen for the measurement of early decoding because the distribution of scores was not marked by strong floor effects as opposed to the more difficult lists. Cronbach's alpha was .92. Accuracy levels were too low to make measures fluency relevant (cf. Juul et al., 2014).

Grade 5

G5 Reading comprehension. Reading comprehension was measured with *Tekstlæseprøve 7* ('Test of text reading') (Møller, 2013). The test consisted of 38 maze items and 11 multiple-choice recall items. The students were given 20 minutes to complete the test. They were allowed to reexamine the text to answer the recall items. Cronbachs alpha was .91 for the cloze items and .77 for the recall items. The correlation between the cloze and the recall items was $r = .74$. The score was the percentage of correct answers across all items.

G5 Isolated decoding accuracy and speed. The test was administered individually on a laptop computer. Each trial started with three fixation asterisks appearing in the middle of the screen. After 500 ms, the asterisks were replaced with the stimulus, which stayed on the screen for 500 ms. Then the stimulus was replaced by a mask. The mask stayed on the screen until the experimenter pressed a button to indicate that the student had answered. The items consisted of 25 nonwords, which were three to five letters long with many consonant clusters to make items difficult (e.g. *gepsk* and *almsk*). The task was programmed in E-prime 2 (Schneider, Eschman, & Zuccolotto, 2012). Sound was recorded and accuracy and reaction time was scored manually using Praat (Boersma &

Weenink, 2014). Reaction time was measured as the latency from the stimulus was presented on the screen until the onset of the first correct answer. Decoding accuracy was computed as the proportion of correct answers. Cronbach's alpha was .90. Decoding speed was computed as the mean reaction time for correct items only. This was done to separate accuracy from decoding speed. The mean score was inverse transformed to provide a scale corresponding to items pr. second, so that larger numbers signify better performance. Split-half reliability was .92 (Spearman-Brown corrected).

G5 List decoding accuracy and fluency. In this individually administered test, the students were asked to read aloud 48 nonwords distributed across six lists. The nonwords were phonotactically legal and two to eight letters in length, many with consonant clusters (e.g. *skvemp*). Both standard and correct conditional pronunciations of individual letters were accepted as correct. The score was the proportion of correct answers. Cronbach's alpha was .91. List decoding fluency was also measured as the number of correct items pr. minute. Fluency should not be confused with decoding speed, that is the speed of successful decoding (see the Introduction). We report fluency for the sake of comparison with previous research.

Results

In the isolated decoding task, individual reaction times above or below two standard deviations from participant means were reduced to a two SD cut-off value. This was done to reduce the influence from extreme values; it was relevant for 6% of the latencies.

Descriptive statistics and zero-order correlations

Table 1 presents the descriptive statistics for the individual study variables.

Table 1. Descriptive statistics

	Mean	SD	Range	Skew	Kurtosis
Letter naming (proportion correct)	.85	.18	.17-1	-1.58	1.85
Letter identification (proportion correct)	.90	.10	.55-1	-1.52	2.09
Phoneme deletion (proportion correct)	.35	.28	0-.94	0.47	-0.97
Phoneme matching (proportion correct)	.85	.14	.45-1	-0.90	-0.17
PAL (proportion correct)	.77	.16	.24-1	-0.89	0.34
RAN - digits (items pr. sec.)	1.03	0.24	0.41-1.86	0.59	1.13
RAN - objects (items pr. sec.)	0.90	0.18	0.48-1.31	-.02	-0.13
G1 List decoding accuracy (proportion correct)	.45	.29	0-1	0.17	-1.19
G5 List decoding accuracy (proportion correct)	.83	.15	.40-1	-1.10	0.32
G5 List decoding fluency (correct pr. min.)	51.44	21.71	10.9-100.8	0.16	0.93
G5 Isolated decoding accuracy (proportion correct)	.68	.24	.08-1	-0.77	-0.49
G5 Isolated decoding speed (items pr. sec.)	1.25	0.40	0.34-2.08	-0.21	-0.66
G5 Reading comprehension (proportion correct)	.71	.21	.17-1	-0.44	-0.65

To simplify the analyses, we computed composite measures of phoneme awareness (phoneme deletion and phoneme matching), letter knowledge (letter naming and letter identification), RAN (RAN objects and digits), and Grade 5 decoding accuracy (isolated and list decoding accuracy) as averaged z-scores. The correlations within the letter knowledge, RAN and decoding measures were strong (r = .60, .65, and .82 respectively), confirming that the pairs tapped very similar abilities. The phoneme awareness measures were only moderately correlated (r = .35), probably due to floor effects on the deletion task and ceiling effects on the matching task. The composite measure did not have floor or ceiling problems and was the strongest correlate of G1 reading (cf. Table 2 below), which indicates that it was a valid measure. To normalize distributions, the following measures

were log or square root transformed as appropriate: PAL, Grade 1 decoding, the composite measure

of letter knowledge, and Grade 5 decoding accuracy.

The attrition rate was 26%. The students who only participated in Grade 0 and 1 did not differ

significantly on the Grade 0 and G1 study variables, with the exception of a slightly lower score on

letter knowledge, $t(185) = 2.63$, $p < 0.01$.

Table 2 presents correlations between the key variables. All of these key variables had skew and

kurtosis values below 1, except G1 decoding which had a kurtosis value of -1.19.

Table 2. Zero-order correlations

	1	2	3	4	5	6	7	8
1. Letter knowledge composite	-							
2. Phoneme awareness composite	.59***	-						
3. RAN composite	.37***	.27**	-					
4. PAL	.23**	.42***	.12	-				
5. G1 Decoding accuracy	.62***	.74***	.36***	.43***	-			
6. G5 Decoding accuracy	.31***	.32***	.34***	.41***	.45***	-		
7. G5 Decoding speed	.11	.06	.27**	.06	.10	.37***	-	
8. G5 Decoding fluency	.21*	.26**	.47***	.34***	.32***	.79***	.62***	-
9. G5 Reading comprehension	.28***	.24**	.31***	.21*	.29***	.50***	.50***	.67

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Letter knowledge and phoneme awareness in Grade 0 were strong short-term correlates of

Grade 1 decoding accuracy ($r = .62$ and .74, respectively). The long-term correlation with Grade 5

decoding accuracy dropped to $r = .31$ and .32. The Grade 0 RAN and PAL measures were

moderately correlated with short-term Grade 1 decoding (r's $= .36$ and .43, respectively). They

remained moderate correlates of Grade 5 decoding accuracy (r's $= .34$ and .41).

It is worth noticing that Grade 1 decoding accuracy correlated with Grade 5 decoding accuracy

($r = .45$, p $< .001$) and reading comprehension ($r = .29$, p $< .001$), but not with decoding speed ($r =$

.10, p > .10). In fact, the only significant early correlate of Grade 5 decoding speed was RAN ($r =$ .31, p < .001), whereas Grade 5 decoding accuracy was correlated with Grade 0 letter knowledge, phoneme awareness, RAN, and PAL.

Does PAL predict decoding accuracy in Grade 1?

The first research question was whether PAL was a unique short-term predictor from the May of Grade 0 to September of Grade 1. In a hierarchical regression analysis, Grade 1 decoding accuracy was used as the dependent variable. Letter knowledge, phoneme awareness, and RAN was entered at the first step, and PAL was entered at the second step. Table 3 presents the results. PAL did indeed explain significant unique variance in Grade 1 decoding, but only 2%. The other predictors explained substantial variance, 61%. The effect of RAN was not significant in the final model ($\beta =$ .11, $p =$ .053), thus letter knowledge and phoneme awareness accounted for most of the explained variance in the short-term prediction. This is not surprising given their high zero-order correlations with Grade 1 decoding (cf. Table 2 above). In summary, PAL only just predicted unique variance in the short term over the stronger predictors, letter knowledge and phoneme awareness.

Table 3. Prediction of Grade 1 decoding accuracy

Step	Predictor	R^2	ΔR^2	β
1		.61	.61***	
	Letter knowledge			.26***
	Phoneme awareness			.49***
	RAN			.11
2				
	PAL	.63	.02*	.14*

Note. * $p < .05$, *** $p < .001$. β is the standardized beta coefficient in the final model.

Are decoding accuracy and speed separable components in Grade 5?

Before analyzing the longitudinal prediction, the following analyses investigated whether the separation of decoding accuracy and speed was valid in Grade 5. This appeared to be the case: First, Grade 5 accuracy and speed were highly correlated with Grade 5 decoding fluency ($r = .79$ and $.62$ respectively), but only moderately with each other ($r = .37$, cf. Table 2).

Second, the two separate Grade 5 decoding accuracy measures correlated highly with each other ($r = .82$), but much less so with the decoding speed measure ($r = .37$ and $.31$ for isolated and serial decoding respectively). Both the composite accuracy and speed measures correlated with reading comprehension (both r's $= .50$, cf. Table 2). Furthermore, a multiple regression with reading comprehension as the dependent variable showed that both Grade 5 decoding accuracy ($\beta = .39$, p $<$ .001) and speed ($\beta = .32$, p $<$ .001) were significant unique predictors explaining a total of 34% variance in reading comprehension. These results indicate that the accuracy and speed measures were valid and separable.

Does PAL predict decoding development to Grade 5 beyond early decoding?

A new set of hierarchical regression analyses were run to test whether PAL predicted development of decoding beyond Grade 1. Table 4 presents separate analyses with Grade 5 decoding accuracy, speed, and fluency as dependent variables. Grade 1 decoding was entered at the first step, followed by letter knowledge, phoneme awareness, and RAN at the second step, and PAL at the third step.

In the analysis of Grade 5 decoding accuracy, PAL explained a significant 6% unique variance – when controlling for Grade 1 decoding and the other predictors. Letter knowledge and phoneme awareness were not significant predictors in the final model, but RAN was. Follow-up regression analyses showed that neither letter knowledge nor phoneme awareness predicted Grade 5 decoding accuracy after controlling for Grade 1 decoding as the only other predictor in the model. Thus, the

longitudinal correlations from letter knowledge and PA to Grade 5 decoding were entirely

explained by these precursors' effect on Grade 1 decoding.

The results were quite different when Grade 5 decoding speed was used as the dependent

variable. PAL did not explain any unique variance. This is not surprising given that the zero-order

correlation between PAL and Grade 5 decoding speed was very close to zero ($r = .06$). It is perhaps

more surprising that Grade 1 reading *accuracy* did not predict Grade 5 decoding *speed* either. Speed

was only measured on correct responses. Although early reading accuracy may be indicative of later

difficulties with accurate reading, early accuracy may simply not have much impact on speed of

correct decoding five years later. The only significant unique predictor in the final model was RAN

($\beta = .27$, p $< .01$).

Thus, PAL only predicted development of decoding accuracy, not speed. For good measure, we

confirmed that the difference in correlation coefficient between PAL and Grade 5 decoding

accuracy ($r = .41$), on the one hand, and PAL and Grade 5 decoding speed ($r = .06$), on the other,

was significant in a test of dependent correlations, $t_2 = 5.33$, $p < .001$ (Steiger, 1980).

Finally, in the analysis of fluency Table 4 shows that PAL was a unique predictor, and so was

RAN. This is not surprising given that fluency is a combination of accuracy and speed.

Table 4. Prediction of Grade 5 decoding accuracy and speed

		G5 Accuracy			G5 Speed			G5 Fluency		
Step	Predictor	R^2	ΔR^2	β	R^2	ΔR^2	β	R^2	ΔR^2	β
1		.20	.20***		.01	.01		.11	.11***	
	Grade 1 decoding			.30*			.01			.10
2		.24	.04		.07	.06*		.25	.14***	
	Phoneme awareness			-.10			-.04			.01
	Letter knowledge			.04			.02			-.07
	RAN			.21**			.27**			.42***
3		.30	.06***			.00		.31	.06**	
	PAL			.29***			.03			.26**

Note. * *p* <. 05, ** *p* < .01, *** *p* < .001. *β* is the standardized beta coefficient in the final model.

Discussion

The present study asked whether paired-associate learning (PAL) is a predictor of decoding in the short and the long term – by tracking reading development of students learning the deep Danish orthography from Grade 0 to Grade 5. The results showed that PAL in Grade 0 was a stable, moderate correlate of decoding both in Grade 1 and Grade 5. PAL predicted Grade 1 decoding beyond letter knowledge and phoneme awareness, but the unique effect of PAL was small. PAL also predicted Grade 5 fluency. But since fluency is a combination of accuracy and speed, analyses of separate measures of accuracy and speed were of greater interest. Indeed, PAL predicted Grade 5 decoding accuracy beyond both the Grade 0 controls and Grade 1 decoding. In contrast, there was no correlation between PAL and Grade 5 decoding *speed*. The results are in line with previous studies that have found correlations between PAL and decoding skills (Elbro & Jensen, 2005; Georgiou et al., 2017; Horbach et al., 2015; Litt et al., 2013; Litt & Nation, 2014; Mayringer & Wimmer, 2000; Messbauer & de Jong, 2003, 2006; Poulsen et al., 2015; Warmington & Hulme, 2012; Wimmer et al., 1998). The present longitudinal findings with control for early decoding skills further supports the notion that PAL measures a trait that influences decoding accuracy development over time. In sum, the results suggest that individual differences in the ability to store and recall phonological material influence both the short-term and the long-term development of accurate decoding skills. But that these individual differences have little or no long-term influence on how fast children decode, once they can decode words accurately.

The differential longitudinal effects of PAL on decoding accuracy and speed are in accord with a cross-sectional trend found by Litt et al. (2013). As mentioned in the introduction, the

selective effect of PAL on decoding accuracy is a potential explanation for the mixed results in the

literature: Most studies have found a correlation between PAL and decoding. The ones that have not

found a robust correlation have used fluency measures in transparent orthographies (Horbach et al.,

2015; Lervåg et al., 2009), where fluency can be expected to be determined mostly by decoding

speed rather than accuracy. But relationships between PAL and decoding has been found even in

transparent orthographies when the samples have overrepresented less skilled readers (Horbach et

al., 2015) entailing some variation in accuracy. If this explanation for the pattern of results is

correct, PAL measures a trait that is associated with decoding accuracy development across

orthographies, but which plays a smaller role in transparent orthographies because achieving

accuracy in shallow orthographies is less demanding.

The present study was not designed to answer exactly what trait in decoding PAL measures or

why PAL is a stable longitudinal correlate of decoding, thus we can only speculate. As mentioned

in the introduction, there are reasons why verbal learning as measured by PAL is likely to continue

to be important for decoding development in deep orthographies beyond the first year of reading

instruction. Letter-sound regularities may be stronger for *patterns* of letters rather than for *single*

letters. For the learner, this means that there is a high number of sound patterns to be learned, one

for each pattern of letters with a predictable pronunciation. These sound patterns may correspond to

whole words (as in the case of *-ought*), but in most cases they do not (as with *-ough*). So they are

essentially new phonological units which must be stored in memory – and activated when the

corresponding letter strings are encountered as part of an unfamiliar written word.

It is a further possibility that young readers memorise spelling pronunciations of whole words

(as mentioned in the Introduction section). A spelling pronunciation is the spoken form that is the

result of a simple letter-to-sound recoding, e.g. "waas" *for was*, or "egg-yipt" for *Egypt* (Elbro & de

Jong, 2017; Elbro et al., 2012). For the young reader, remembering the spoken form "waas" would

form a stepping stone between the written word and the standard spoken word form. The expense would be the need to learn an additional *spoken* form of an already known word (compare to learning a spoken dialect form), but the advantage for orthographic learning would be huge because the beginner would be able to decode words by means of a very limited number of grapheme-phoneme rules. In addition, such spelling pronunciations are often memorised and consulted by spellers to aid with the spelling of irregularly spelled words (Nielsen & Juul, 2015). In any case, learning spelling pronunciations requires verbal learning.

The above interpretation of the results is not entirely straight-forward, however. In the present study, decoding accuracy with nonwords were scored leniently. That is, responses were considered correct both if the pronunciation followed simple letter-sound correspondences and if they followed more advanced pronunciations of letter patterns. Hence, decoding accuracy was not a direct measure of knowledge of the sounds of letter sequences such as spelling patterns. Nonetheless, knowledge of the sounds of letter sequences may be a significant aid to the reading of non-words – simply by *reducing the number* of sound segments to be kept in working memory and blended in each nonword. It is, however, not clear why such an aid would support accuracy only and not speed. One might expect that decoding based on chunks of letters rather than single letters would speed up nonword decoding. One possibility is that students who have available representations of letter chunks may have a greater chance of hitting the target pronunciation (i.e. being accurate) under conditions that do not invite laborious letter-by-letter recoding. But the benefits of letter chunk representations for decoding speed may be too small to measure when the students hit the target. In the present study, both decoding measures were associated with a certain amount of time pressure, which could prohibit conscientious recoding. This may be representative of many text reading scenarios, including the present reading comprehension test, where some students may have been more concerned with completing the entire task than with getting every word right.

Limitations

PAL was measured somewhat differently in the present study as compared to other studies. The procedure was to a higher degree designed to support the students' learning by ensuring engagement, providing rich semantic content for the phonological forms to be learned, and by introducing the different forms gradually. The intention was to improve reliability and validity of the measure by limiting influences of general motivation and attention with this difficult task. It is thus possible that the stronger longitudinal effects of PAL in this study compared to previous studies were due to differences in the PAL format rather than differences in the decoding measures. It is difficult to pinpoint exactly what differences that would be responsible for the changes, but some things are worth considering. The complex format of the task with gradual introduction of items and narrative support appeared successful: The mean proportion of correct answers were .77. For comparison, the proportion was between .31-.55 in Lervåg et al. (2009). This could raise the concern that the present PAL task was sensitive to individual differences in semantic processing. This concern cannot be ruled out, but the correlation with Grade 5 nonword decoding accuracy was stronger ($r = .41$) than with Grade 5 reading comprehension ($r = .21$, significance of difference: $t_2 = 2.48$, $p < .05$). This suggests that the PAL measure is still a phonological rather than semantic measure on the assumption that a semantic measure would show a higher correlation with reading comprehension.

Grade 5 decoding speed was measured with an isolated non-word decoding task, and only latencies from accurate trials were used. Therefore, the speed score represents how fast the students successfully decoded nonwords. We do not know to what extent the results generalize to irregular word reading speed. The stimuli were masked after 500 ms, so the measure may have missed variance from slow, laborious decoding. But it should be noted that the longest nonwords were only five letters long. Unsuccessful decoding will in all likelihood slow down connected decoding, thus

PAL can be expected to influence text reading fluency if the materials are challenging, and indeed, PAL did predict Grade 5 decoding fluency. The speed measure was highly correlated with fluency and was a unique predictor of reading comprehension beyond accuracy, suggesting some level of validity.

There was a 26% attrition in the study. On most measures, the group of students who left the study did not differ from those who completed. But they did differ on letter knowledge. Although unlikely, we cannot rule out that this may have skewed the results.

It is interesting that PAL predicted variance in decoding accuracy over a five-year period even when controlling for Grade 1 reading accuracy (an autoregressor). However, it would have been interesting to investigate at which grade levels between G1 and G5 PAL contributes the most variance to decoding accuracy. The students did complete decoding tests at the end of Grade 1 and Grade 2, but unfortunately these measures had pronounced ceiling effects on accuracy, which invalidate them as meaningful control variables. Difficult items were added to the Grade 5 measure to avoid the ceiling effects of the previous measures.

Summary and conclusions

The present longitudinal study found that performance on a paired-associate learning (PAL) task before formal reading instruction predicted decoding performance in Grade 5, but only when decoding was measured in terms of accuracy, not speed. The prediction was significant after controlling for letter knowledge, phoneme awareness, RAN, and Grade 1 decoding skills. This indicates that PAL measures a trait that influences reading accuracy development over a substantial amount of time. Future studies of PAL's relationship with reading should be explicit with regard to which aspect of reading is under investigation, accuracy or speed. It is suggested that PAL is a measure of verbal learning of both single-letter and multi-letter phonological representations, which may be important for accurate decoding of both known and novel words. The load on phonological

working memory will be significantly eased if the reader can decode by chunks of letters rather than by single letters, and accuracy with irregularly spelled words may be greatly improved. Further studies are needed to test these possibilities directly.

References

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer (Version 5.3.77). Retrieved from http://www.praat.org/

Borstrøm, I., & Petersen, D. K. (2006). *Læseevaluering på begyndertrinnet [Evaluation of beginning reading]* (2nd edition ed.). Copenhagen, Denmark: Alinea.

Denckla, M. B., & Rudel, R. G. (1976). Rapid `automatized' naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*, 471-479.

Elbro, C. (2005). Literacy acquisition in Danish: A deep orthography in cross-linguistic light. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of Orthography and Literacy* (pp. 31-45). Mahwah, NJ: Lawrence Erlbaum.

Elbro, C., Borstrøm, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33*, 36-60. doi:10.1598/RRQ.33.1.3

Elbro, C., & de Jong, P. F. (2017). Orthographic learning is verbal learning. The role of spelling pronunciations. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of Reading Development* (pp. 169-189). Amsterdam: John Benjamins.

Elbro, C., de Jong, P. F., Houter, D., & Nielsen, A.-M. (2012). From Spelling Pronunciation to Lexical Access: A Second Step in Word Decoding? *Scientific Studies of Reading, 16*, 341-359. doi:10.1080/10888438.2011.568556

Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology, 46*, 375-384. doi:10.1111/j.1467-9450.2005.00468.x

Georgiou, G., Liu, C., & Xu, S. (2017). Examining the direct and indirect effects of visual–verbal paired associate learning on Chinese word reading. *Journal of Experimental Child Psychology, 160*, 81-91. doi:10.1016/j.jecp.2017.03.011

Horbach, J., Scharke, W., Cröll, J., Heim, S., & Günther, T. (2015). Kindergarteners' performance in a sound-symbol paradigm predicts early reading. *Journal of Experimental Child Psychology, 139*, 256-264. doi:10.1016/j.jecp.2015.06.007

Hulme, C., Goetz, K., Gooch, D., Adams, J., & Snowling, M. J. (2007). Paired-associate learning, phoneme awareness, and learning to read. *Journal of Experimental Child Psychology, 96*, 150-166. doi:10.1016/j.jecp.2006.09.002

Juul, H., Poulsen, M., & Elbro, C. (2014). Separating speed from accuracy in beginning reading development. *Journal of Educational Psychology*, 106, 1096-1106. doi:10.1037/a0037100

Lervåg, A., Bråten, I., & Hulme, C. (2009). The cognitive and linguistic foundations of early reading development: A Norwegian latent variable longitudinal study. *Developmental Psychology, 45*, 764-781. doi:10.1037/a0014132

Litt, R. A., de Jong, P. F., van Bergen, E. v., & Nation, K. (2013). Dissociating crossmodal and verbal demands in paired associate learning (PAL): What drives the PAL–reading relationship? *Journal of Experimental Child Psychology, 115*, 137 - 149. doi:10.1016/j.jecp.2012.11.012

Litt, R. A., & Nation, K. (2014). The nature and specificity of paired associate learning deficits in children with dyslexia. *Journal of Memory and Language, 71*, 71-88. doi:10.1016/j.jml.2013.10.005

Mayringer, H., & Wimmer, H. (2000). Pseudoname learning by German-speaking children with dyslexia: Evidence for a phonological learning deficit. *Journal of Experimental Child Psychology, 75*, 116-133. doi:10.1006/jecp.1999.2525

Messbauer, V. C. S., & de Jong, P. F. (2003). Word, nonword, and visual paired associate learning in Dutch dyslexic children. *Journal of Experimental Child Psychology, 84*, 77-96. doi:10.1016/S0022-0965(02)00179-0

Messbauer, V. C. S., & de Jong, P. F. (2006). Effects of visual and phonological distinctness on visual–verbal paired associate learning in Dutch dyslexic and normal readers. *Reading and Writing: An Interdisciplinary Journal, 19*, 393-426. doi:10.1007/s11145-005-5121-7

Møller, L. (2013). *Tekstlæseprøve 7 [Text reading test 7]*. Frederiksberg, Denmark: Hogrefe Psykologisk Forlag.

Nielsen, A.-M. V., & Juul, H. (2015). Predictors of early versus later spelling development in Danish. *Reading and Writing*. doi:10.1007/s11145-015-9591-y

Poulsen, M., Juul, H., & Elbro, C. (2015). Multiple mediation analysis of the relationship between rapid naming and reading. *Journal of Research in Reading, 38*, 124-140. doi:10.1111/j.1467-9817.2012.01547.x

Poulsen, M., Nielsen, A.-M. V., Juul, H., & Elbro, C. (2017). Early identification of reading difficulties: A screening strategy that adjusts the sensitivity to the level of prediction accuracy. *Dyslexia, 23*, 251-267. doi:10.1002/dys.1560

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide.* Pittsburg: Psychology Software Tools Inc.

Seymour, P. H. K., Aro, M., Erskine, J. M., & network, C. A. A. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143-174. doi:10.1348/000712603321661859

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 37*, 245-251.

Warmington, M., & Hulme, C. (2012). Phoneme Awareness, Visual-Verbal Paired-Associate Learning, and Rapid Automatized Naming as Predictors of Individual Differences in Reading Ability. *Scientific Studies of Reading, 16*, 45-62. doi:10.1080/10888438.2010.534832

Wimmer, H., Mayringer, H., & Landerl, K. (1998). Poor reading: A deficit in skill-automatization or a phonological deficit? *Scientific Studies of Reading, 2*, 321. doi:10.1207/s1532799xssr0204_2