Prediction Models of Literacy Skills and Deficits: a Review of the Research

Tereza Medřická, Anna Kucharská

Abstract: Literacy skills – decoding, reading comprehension and spelling – are an integral part of today's world and it is necessary for people to acquire them on an adequate level if they wish to participate actively in society. This does not always happen, however. In practice, we try to prevent the formation of literacy deficits or to at least mitigate their manifestation. Much attention in research is paid to the study of literacy as such, its development and its disorders, and, last but not least, to the potential ways of predicting these disorders. Early detection of individuals who are at risk paves the way for early interventions, thus minimizing the manifestation of problems. Each individual research study is highly specific because it depends greatly on how its authors handle the given topic. This paper provides an overview of variability in various areas of the study of the development of literacy skills, literacy deficits and prediction models, including examples of larger meta-analytic research studies which attempt to settle this variability.

Key words: Literacy skills, literacy deficits, prediction models, methodology, review of research

Introduction – Literacy Skills and Deficits

The term **literacy** certainly has an indisputable place in the modern world. On the lowest level, we talk about basic school skills such as **reading, writing and arithmetic.** These skills are a conduit for additional education and we use them as a tool to obtain information; they also predetermine our professional and social outlook, and they are a hobby as well as a mode of self-presentation and self-fulfilment (Česká školní inspekce, 2013; for more information on the term *literacy*, see e.g. Hart & Hartlová, 2010).

In terms of development, we can distinguish the precursors of literacy skills and conventional literacy skills (NIFL,[1] 2008; Helus, 2012). The precur-

---

[1]   *The National Institute for Literacy.*

sors or predictors of literacy skills are understood to be those abilities and skills which affect the development of our reading and writing later on but are not themselves directly applied in reading and writing itself. Language skills are an example of this. Conversely, conventional literacy skills already include the actual process of reading and writing – we distinguish decoding, reading comprehension and spelling. Precursors, as well as *early* conventional literacy skills, are developed at preschool age prior to the onset of formal school education. Children are surrounded by texts at every turn; they show a natural interest in them and they deal with them in a specific manner (Ferreiro & Teberosky, 1982). More advanced *conventional literacy skills* as such are then the result of school education. They are not the result of natural development but rather of targeted instruction.

The process of acquiring literacy skills is often disrupted for various reasons. Children with specific learning disorders (hereinafter referred to as „SLD" or „dyslexia") are a fairly large group that is clearly defined within the Czech school context. The „inability to learn corresponding skills on a level that would correspond to the intellectual ability of the given child and is not the result of a sensory or mental handicap, emotional or other psychological problems or the result of negligence in rearing or education" is typical for these children (Vágne-

rová & Klégrová, 2008, p. 369). **Disorders in the acquisition of literacy** in the broader sense of the word, however, are also found in pupils with mental retardation, in pupils with a specific language impairment or in slow learning, in children with a sensory disorder as a result of emotional or psychological problems, parental negligence, a socio-cultural handicap or of immaturity or an unsuitable educational approach (Berninger, 2001; Přinosilová, 2007; Vágnerová & Klégrová, 2008).

In practice, we try to prevent the development of literacy disorders or at least minimize their level of severity. It is for this reason that our attention is shifting away from the study of disorders as such to preschool age, where the prerequisites for the development of literacy skills are formed. If we are able to map out the development of the precursors of literacy and its deficits, we can follow up with targeted intervention efforts that will help develop the areas of insufficiency (for more on the topic of the *risk of dyslexia* see e.g. Kucharská, 2007).

This paper provides an overview of several selected studies dedicated to a deeper understanding of the development of literacy and its deficits and efforts to predict them. The authors attempt to capture the variability of access to various areas of the topic under study – from the theoretical understanding of the concept, through the definition of groups that are monitored, up to the

selection of the actual methodology and methods used for processing data. In conclusion, the authors list examples of meta-analytical studies which also work with this variability and focus (among other things) on the generalization of predictors (disorders) of literacy.

Literacy in Research

Quite a lot of attention is paid to the topic of the development of literacy in research today, as is apparent from the large numbers of records in the citation databases. Researchers monitor the development of literacy itself (e.g. Caravolas et al., 2012; Torppa et al., 2016) but also pay attention to its disorders, as well as the potential ways of predicting them (e.g. Snowling, Gallagher, & Frith, 2003; Kucharská, 2014; Bigozzi et al., 2016; Moll et al., 2016; Medřická, 2019). To put the research findings into practice, we need to have clear and unequivocal conclusions. The generalization of research findings is not simple, however, because individual research studies have their particularities and their conclusions are often not in direct agreement. **Meta-analytical studies and reviews**, which try to bridge the specifics of the individual research findings, attempt to find the convergence of a larger volume of research works and to unify the

conclusions of various authors and generalize these research findings.

An example of a large meta-analytical study is the *National Early Literacy Panel* (hereinafter referred as „NELP" – for more details, see NIFL, 2008). NELP is a large-scale meta-analysis that was carried out in the United States from 2002 to 2006. Its goal was to synthesize the results of about 500 available studies dedicated to the development of early literacy skills in preschool children from birth to five years old. NELP was a reaction to the *National Reading Panel* (hereinafter referred to as „NRP", for more details, see NICHD,[2] 2000), which used a meta-analysis of published studies to evaluate the efficiency of reading and writing teaching methods and reacted to the findings from the *National Assessment of Educational Progress* that more than one-third of fourth-grade pupils in the US do not reach a basic level of reading development. Several analytical studies pointed out the wide variability in (pre-)reading skills in children already at the time when they begin school. This turned the attention to preschool children. The goal of the related NELP was to obtain information on how to provide effective support for the early development of literacy (and language) and the influence of the family environment, which could be used in practice to improve children's

---

[2]   *The National Institute of Child Health and Human Development.*

readiness for future reading and writing instruction. The partial and initial goal of NELP, however, was to specify individual predictors, in other words abilities and skills, which predict higher reading and writing levels in preschool children later on.

Snowling and Melby-Lervåg (2016) carried out a later meta-analytical study in Europe, processing 95 publications dedicated to groups of children with a familial risk of dyslexia. They focused on the assessment of the prevalence of dyslexia in a given group and finding the risk and protective factors of dyslexia, the generalization of predictors and the assessment of intervention measures.

Below we provide an overview of possible approaches to studying literacy and its disorders, as well as the options for their prediction.

Methodological Approach

Longitudinal research is a tool for monitoring developmental changes over time (for more see e.g. Langmeier & Krejčířová, 2006) in which we monitor a specific group of individuals in various consecutive phases where we repeatedly assess the developmental level in the area of our interest – here this is literacy skills and their deficits and predictors. We see two approaches here – developmental and retrospective.

If we define a research group that is monitored according to symptoms present at the start of the research, we are working with a **developmental (prospective) model.** We monitor developmental changes at regular intervals and evaluate the current status at the time of measurement. If we are monitoring several partial groups, this approach allows us to compare not only ongoing changes in development within the group but also differences in the scope of developmental changes between the groups. We can see this approach e.g. in the work of Kucharská (2014), who monitors the developmental profiles of language, cognitive and pre-literacy skills in several groups of children. The groups of children who are monitored – children with typical development, children with familial risk of dyslexia and children with a specifically impaired speech and language development – are defined at the start of the research and are then recorded again on the basis of the data obtained during the first phase of the testing. The children's development is evaluated in a total of five testing phases for a period of five years, starting at preschool age (about five years of age) through to the third grade of elementary school, where the children's literacy skills and their deficits are assessed. Besides the dynamics of development in individual groups, Kucharská also compares the developmental profiles of individual groups against each other – the biggest risk group for future literacy skill deficits is the group of children with spe-

cifically impaired speech and language development, followed by the group of children with a familial risk of dyslexia, even though this group often comes close in terms of performance to the group of children with typical development.

If we define a research group that is monitored according to characteristics present at the end of the research, we are working with a **retrospective (anamnestic) approach** in which we go back from the present moment into the past during the assessment. We try to find a relationship between a defining characteristic in the present and specific expressions (or significant events) in the past – care must be taken here, however, when deducing causal relationships in order not to mistake them for simple correlation (Langmeier & Krejčířová, 2006). This approach allows us to look for predictors of literacy skills and difficulties associated with it in the earlier stages of development. Even in this approach, one can find differences in the predictors between the groups. For example, Medřická (2019) works with this approach and similarly to Kucharská (2014), she monitors identical groups of preschool-aged children up through to the third grade of elementary school for a period of five years. However, she uses latent profile analysis to divide the base group of children into two groups on the basis of their performance in the third grade – children with literacy deficits and those without. Unlike Kucharská,

Medřická looks at the continual development of children in retrospect, taking into account their literacy status in the third grade of elementary school. On the basis of the performance of children with literacy deficits, prediction models of literacy deficits are set up using *lasso*, in other words L-1 penalized regression, for each of the four previous testing phases (a year before the end of preschool, at the end of preschool, in the first grade of elementary school and in the second grade of elementary school). Individual models set up in this manner represent the best combination of testing methods used for the evaluation of development in the given testing period, which together best predict the later incidence of literacy deficits.

Both approaches can be combined and several authors end up doing this. First, they describe the chronological development in a group and after that they create prediction models. We can cite the work of Moll et al. (2016) as an example. They map the precursors of reading difficulties in a group of Czech and Slovak children at risk of dyslexia. In the first phase, they describe the differences in the developmental indicators between groups of children with a familial risk of dyslexia, those with speech/language difficulties and a control group. They assess their development in three phases over a period of three years – from preschool to the first grade of elementary school. They primarily monitor oral lan-

guage and code-related skills, and in the first grade, they evaluate literacy skills. In the second phase, they focus on the creation of a prediction model in which a two-group latent variable path model shows that early language skills predict code-learning skills and those in turn predict literacy skills. They therefore determine that the originator of dyslexia in the Slovak children is early language deficits and they also find impaired code-related skills measured at a time prior to the start of formal school education in children with reading difficulties in the first grade of elementary school.

Research Sample – Monitored Groups and Size

When studying the development of literacy and its disorders, we can monitor the general population, in which we usually compare individuals according to the level of development of their skills or individuals with deficits and without literacy impairments (e.g. Caravolas et al., 2012; Bigozzi et al., 2016; Torppa et al., 2016). We can also monitor previously defined **risk groups**, where a higher risk of difficulties is expected. The primary „advantage" of working with a risk group is summarized by Snowling and Melby-Lervåg (2016), who point to the fact that if we wish to monitor 50 persons with dyslexia within the standard population, for example, and if the prevalence

of dyslexia is 10 %, we need an initial sample of 500 persons, while the size of the basic group in the risk population decreases, depending on the incidence of persons with difficulties in the specific risk group.

Children with a familial risk of dyslexia are a typically monitored risk group (e.g. Snowling et al., 2003; Snowling et al., 2016), as are children with language or speech impairments, and some authors monitor both of these groups concurrently (e.g. Kucharská, 2014; Moll et al., 2016; Medřická, 2019). In their meta-analytical study, Snowling and Melby-Lervåg (2016) focus on a group of children with a familial risk of dyslexia while pointing to the interesting fact that several authors in the original studies do not monitor any potential comorbidity, such as e.g. language or speech disorders. The question to ask then is how often the group of children with a familial risk of dyslexia is created rather as a „mixed" risk group. They also point to the difference in the definition of the group with a familial risk as such – those individuals who have at least one parent or older sibling with dyslexia are generally considered to be children with a familial risk of dyslexia. Researchers have different opinions on how they back up this criterion – while some are satisfied with a parental self-report (e.g. Moll et al., 2016), others verify the parents' reading skills using (standardized) testing methods (e.g. Kucharská, 2014; Medřická, 2019).

We compare risk groups with each other or with a control group or in greater detail – we generate sub-groups of risk children with difficulties and those without, as well as control group children with difficulties and without based on evaluations of learned literacy skills. Authors select various combinations of the individual sub-groups for their analyses with consideration for their research questions, goals and hypotheses.

When processing original texts as part of their meta-analysis, Snowling and Melby-Lervåg (2016) find three types of different research designs: (a) studies which take reading skills as a continual variable and do not divide children into groups of those who are dyslexic and those who are not; in their analyses, they compare a risk group with a control group without additional sorting; (b) studies which compare children from a risk group with subsequent problems in literacy and a control group, and (c) studies which compare children from a risk group without problems in literacy and a control group.

Moll et al. (2016) is an example of a combination of various approaches within one work. While studying risk factors for the development of reading problems, (1) when searching for putative causes of dyslexia, they compared the preschool performance of children from a risk group with problems and a control group without problems and found substantial and statistically significant differences for code-related skills (letter-sound knowledge, phoneme awareness, and RAN) at the beginning of the last preschool year (T1) and again at the end of the last preschool year (T2) and for grammatical skills (only) in T2; (2) when mapping factors associated with the risk of dyslexia (i.e. putative *endophenotypes* of dyslexia), they compared the risk group without problems in reading and a control group without problems, as well as a risk group with reading problems and in phonological skill tests (word/pseudoword recall and non-word repetition), they found impaired performance both in risk children with reading problems as well as in risk children without reading problems in T1 and T2, just as in vocabulary in T1; (3) they compared a risk group with reading problems and a control group with reading problems to determine whether the risk factors differed depending on whether the children came from the control or the risk group and even though none of the T1 and T2 scales from the groups differed significantly, the authors generally found worse performance in children from the risk group, especially in the word/pseudoword recall and non-word repetition tests. Because of the small number of individuals, the authors grouped children with a familial risk of dyslexia together with children with speech/language difficulties into the risk group for the listed part of the analyses; once they focused on the differences between these two groups,

however, they found worse performance on tests of phoneme awareness and RAN (not in letter knowledge or reading) in risk children with reading deficits from the group with a familial risk of dyslexia. On the contrary, the children with speech/language difficulties performed worse in grammar.

Bigozzi et al. (2016), for example, worked with a sample of children from the **general population**. In a research study that monitored differences in predictors of literacy in Italian children with dyslexia and without, they worked with a group of 450 children who were attending standard preschools and elementary schools. The sample size dropped down to 427 individuals once the children entered elementary school because some children entered a different school from the one involved in the research project. Of the 427 children, a group of nine children was created, comprising children who were diagnosed with dyslexia in the third grade of elementary school and the complete data from the previous phase of testing was missing for three other children with dyslexia. The performances of these children in the earlier phases of testing were then compared to the performances of the children from the control group. The control group comprised 65 children without dyslexia and was created from within the original group – these were classmates of the children with dyslexia from the same class (to ensure the same teaching practices), with the same level of SES (assessed on a scale of 1 to 5) and of the same sex.

The maintenance of an adequate **size of the research sample** for the entire duration of the research can be a challenge for longitudinal research. There are many different reasons for this. In the Czech section of the ELDEL WP2[3] research project (Kucharská, 2014; Medřická, 2019), a repeated reduction of the research sample occurred as a result of several factors. Some of the families lost interest or motivation to participate in the project. Others moved or changed their contact information and the researchers were unable to renew cooperation with them. In others, it was impossible to set a time for some of the testing phases because of the parents' busy schedules or because the child was sick for an extended period of time. The most significant drop occurred as a result of revision of the children's affiliation into the groups that were monitored. The initial sign-up occurred on the basis of the interest of the parents and their subjective sorting into one of the recruitment groups on which the project focused. Following the initial testing phase, the affiliation of the children to the individual research

---

[3]   *Enhancing Literacy Development in European Languages, Work Package 2,* for more information please visit: www.eldel.cz.

groups was re-evaluated on the basis of the performance of the children or parents in selected tests or information the parents provided in the questionnaires. From the total of 149 children who were recruited, 48 were recruited into the group with impaired speech (and language) development and only 18 were left after the revision; there were 48 children recruited into the group with a familial risk of dyslexia, of whom 41 fulfilled the criteria (Kucharská, 2014). Over the course of the five years of monitoring the group of children (first as part of the ELDEL project and then as part of the GA UK[4] project), the total number was reduced from the original 149 children to a final 96 children who participated in the project through to the third grade of elementary school and of those, only 76 children passed the revision of the defining criteria for the main group that was monitored (Medřická, 2019).

Organizational difficulties may also arise in the long-term horizon. In the research of Caravolas et al. (2012), about 50 children (out of the initial 735 participants) moved away between the two phases of testing prior to the start of school. 23 children had to be removed from the research sample of Bigozzi et al. (2016) because after preschool, they started attending another elementary school than the one involved in the research project. From a total sample of 308 individuals, 64 children were excluded from processing in Moll et al. (2016) because in the final phase of testing at the end of the first grade of elementary school, these children had postponed the start of their elementary school education and therefore did not attend elementary school, meaning that they were obviously at a disadvantage in terms of formal literacy, which the authors verified directly as part of partial analyses of the data that was obtained.

Problems with the size of the research sample can also occur as a result of retrospective research *per se*. In a retrospective approach we reorganize the sample of children that is monitored on the basis of their performance during the last stage of monitoring. In studying literacy predictors, we divide the sample of children on the basis of the level of literacy reached. The size of the groups that we are able to get is therefore completely beyond our control, as opposed to the size of the recruitment groups, which we create at the start of the research, meaning that we actively regulate their size at the given moment. Depending on the level of representation of a disorder in the group of children that is monitored, we may end up with a small group, which is of no interest to us. Bigozzi et al. (2016) and Medřická (2019) were both

---

left with groups of children with literacy deficits that comprised a mere nine children. This fact subsequently influences the additional data processing in terms of the way we process the groups that are monitored, e.g. even in terms of selecting the statistical method or the data processing method or the level of usefulness of the results obtained.

Moll et al. (2016) attempted to deal with a small number of children in a group by grouping the risk groups into one. They primarily worked with three groups of children in their research. Two risk groups comprised children with a familial risk of dyslexia (the FR group), where a parent or a sibling was dyslexic, and children with clinical concerns about their speech and language development (the SLD group), whose problems were defined by a performance worse than a standard deviation of 1 (hereinafter referred to as „SD") below the corresponding age average of two out of three criterion tests (specifically, vocabulary and receptive and expressive grammar) or they fulfilled the criteria for a speech sound disorder, which is defined by the correct pronunciation of less than 80 % of the consonants in a pronunciation test. The control group comprised children with typical development (the TD group) without speech and language difficulties and without a family predisposition. The literacy skills (decoding and spelling) of the children were evaluated after a year of formal education at the end of the first grade. For the purposes of prediction, they created a composite score of decoding as a sum of the z-scores from individually given tests. They then defined reading difficulties by a performance that was 1,5 SD below the average of the control group. During the subsequent reorganization of the basic groups according to literacy/reading status, they ended up with a small size of the partial groups created, which complicated the subsequent statistical processing. That is why they decided to combine both risk groups into one „at-risk" (AR) group for some of the analyses. They therefore obtained four groups of children according to the reader status achieved – AR-normal reader, AR-poor reader, TD-normal reader and TD-poor reader. The fact that the TD-poor reader group is included in the partial analyses is quite unusual in this work. Here, the authors compare risk factors for the onset of literacy difficulties between the TD-poor reader and AR-poor reader groups and thereby determine whether these risk factors differ if the children are from the risk group. The analyses indicate that the partial skills monitored in the T1 and T2 preschool age testing phases are more disturbed in children from the risk group, especially on the word/pseudoword recall and non-word repetition scale; however, the key finding is the fact that on no scale in T1 and T2 are there significant differences between these two groups of children.

Medřická (2019) used this result as her basis. As she created retrospective prediction models, she also had to deal with a small number of children in risk groups at the end of the monitoring in the third grade of elementary school. She therefore used one basic group as the basis for the corresponding analyses. This group comprised all three of the groups of children who were monitored – children with a familial risk of dyslexia, children with disturbed language and speech development and children with typical development. Using latent profile analysis, which allows hidden groups to be found in a set of data (for more on the method see e.g. Oberski, 2016), and without having previously defined criteria for these groups, she separated 76 children from the group with a sub-group of nine children with literacy deficits and for them, she created prediction models for four monitored phases of their preschool and early school development.

Period Monitored

The periods for which research groups were monitored by the authors are different. The upper time limit is defined by the moment **when the acquisition of literacy skills, or their deficits, is assessed.** Bigozzi et al. (2016) assess the acquisition of reading skills in the third grade (at the beginning), which is when dyslexia is typically diagnosed in Italy.

Similarly, Medřická (2019) assesses the literacy skills of Czech pupils in the third grade, where she expects that problems in acquiring literacy have a deeper basis and are not merely the result of partial impairments with which children can struggle at the start of their school education. Moll et al. (2016) assess the skills of decoding and spelling in Czech and Slovak children already at the end of the first grade, i.e. after one year of formal education. Torppa et al. (2016) connect both approaches by using former findings that some partial literacy skills already reach a ceiling during the first year of education (decoding accuracy specifically), while others (such as decoding fluency and reading comprehension) are connected to them and are developed over a longer period of time. They also only monitor the level of word-reading fluency up to the second grade, while reading comprehension is still monitored in the third grade.

A cultural viewpoint can bring us other differences as well. Besides the orthographic depth of a specific language, these cultural differences are also reflected in the nature of the educational system, which defines the period when children enter elementary school and how formal education is approached, whether preschool education is obligatory and what children's attendance in preschool comprises, i.e. whether preschool-aged children are already exposed to some sort of formal instruction of literacy skills

or if their predictors are at least stimulated and to what extent.

Caravolas et al. (2012) monitored the predictors of early literacy skills in four groups of children whose mother tongues were different in terms of orthographic depth – specifically, English, Spanish, Czech and Slovak. The nature of orthography proved to be a significant factor affecting the development of literacy and the nature of difficulties in the development of literacy (for more see e.g. Caravolas, 2005). Caravolas et al. (2012), however, point to another possible factor, which is the different period for the start of school education or the formal education of literacy skills and also the different content of preschool education. The first phase of testing took place in the last year prior to the start of the first grade of elementary school, in which all children went to kindergarten or the reception year in England. The average age of the English children was younger than in the remaining three groups of children – by six months compared to the Spanish children and by almost 12 months compared to the Czech and Slovak children, which is the result of the different age at which children in each given country enter the first grade of elementary school. The content of the preschool education is also different. Testing in the first phase occurred at half-term, when the English children had already been exposed to formal literacy instruction for five to six months. The other groups had instruction in certain letter-sound knowledge and phonological awareness skills as part of kindergarten, and had learned to recognize their names and some signs that they saw in the classroom on a daily basis. These differences are then most probably reflected in the differences found between the groups. The English children had a better knowledge of letters, which is a point of focus in reception class instruction, and the English and Spanish children had better spelling skills, while the Czech and Slovak children had a slightly better picture-word matching reading measure. Generally, however, the groups reached relatively similar results in the first phase of testing in the measurements of early literacy that were monitored.

In one of their partial analyses, Torppa et al. (2016) monitored whether preschool attendance *per se* has an impact – of a total sample of 1815 children who took part in testing in the first, second and third grades, only 1546 took part in testing in kindergarten. The authors compared the development of reading in children who were tested in kindergarten and those who were not. The performance of both groups of children was similar, with the exception of reading fluency in the second grade. All the while, the authors point to the example of Finland, where preschool education focuses mainly on personal and social development as opposed to instruction in academic skills. Children,

however, are still often read to and play with letters, words and numbers and about 50 % of children learn to decode at least a few words prior to starting elementary school.

The „age" factor is also monitored by authors in the meta-analytical study that is part of the NELP (NIFL, 2008). As part of the secondary analyses, they monitor whether the age when predictors (preschool vs. kindergarten) are evaluated has an influence, as well as whether the age at which the final level of literacy is assessed (kindergarten vs. first or second grade) does. In terms of the first question, they found only minor differences in the influence of the predictor in relation to the time when the (early) predictors were monitored, where the strength of the prediction in such a case was higher if the predictors were monitored (usually) at preschool age and in variables in the moderate relationship range. In terms of the second question, they focused on comparing the strength of the prediction in the studies where the „final" literacy skills were assessed in kindergarten or at the start of school education (in the first or second grade). The found significant differences in the strength of prediction in relation to the time when literacy results were assessed in about 50 % of cases. Mainly, they found a stronger relationship if the literacy skills were already assessed in kindergarten – the authors understand this to be due to the time proximity

between the predictor measurement and the subsequent outputs, as well as the result of the fact that with the start of school, children are exposed (perhaps) to a greater level of variability in the way instruction in conventional literacy skills is approached, which also causes greater heterogeneity in their performance.

Similarly, in their meta-analytical work, Snowling and Melby-Lervåg (2016) take note of the developmental viewpoint where the authors of the original studies collect data on early literacy development at various developmental stages. As part of studying endophenotypes of dyslexia, where they compare the development of children with a familial risk of dyslexia and children in a control group, and by finding the risk factors associated with familial dyslexia, they sort the processed data into four groups: (a) infants and toddlers (0–3 years), (b) preschool (below 5.5 years and before formal reading instruction starts), (c) early primary school (up to fourth grade), and (d) late primary school/secondary school (from fifth grade).

Grasping the Concept of Literacy Skills and Their Deficits, Definition Criteria

If we wish to assess **literacy deficits,** we must first define them. In practice, literacy deficits are often understood to be **specific learning disabilities.** An SLD

diagnosis includes detailed pedagogical and psychological tests (see e.g. Vágnerová & Klégrová, 2008), of which the results must be considered in the broader context of the child's situation, where we evaluate the complete case history and the course of the child's school education up to this point (defined according to DSM-5; American Psychiatric Association, 2013). We can see this type of approach to literacy difficulties in the research of Bigozzi et al. (2016). They isolated 35 children out of a research group of 427, who were singled out by their teachers as individuals with difficulties in reading and were sent to centres specializing in the diagnostics and treatment of learning disabilities – 12 children were diagnosed with dyslexia.

Several authors, however, „limit" their work to the simple assessment of the level of development of literacy skills. For example, Torppa et al. (2016), who work with the *simple view of reading* concept, evaluate literacy skills on a scale of listening comprehension, reading fluency and reading comprehension. Literacy skills, however, are much more often assessed on the reading (decoding) and spelling scale (e.g. Caravolas et al., 2012; Moll et. al., 2016) or in greater detail, such as decoding, reading comprehension or spelling (e.g. Snowling et al., 2003; Kucharská, 2014; Medřická, 2019). Even here, however, we find considerable variability in the selection of the testing methods which the authors choose for

their assessment of the level of literacy.

Neither do we find a unified approach in the method used for establishing **deficit performance limits.** If authors work with a standardized test battery, the deficit limit can be established by the relevant (population) norm. „Normal" performance, however, is often defined by the average performance of children from the control group. Snowling et al. (2003) compare two approaches in the evaluation of the level of literacy skills. They assess literacy skills in eight-year-old children on a scale of basic reading, reading comprehension and spelling, where they calculate one composite score from the standard scores. In the first case, literacy deficits are defined with a limit of minus 1 SD below the average of this composite score in the control group. This criterion revealed literacy problems in 66,1 % of the children from the risk group and 13,8 % of the children in the control group. In the second case, they work with a traditional discrepancy principle, where dyslexia is defined as a significant difference between the expected reading level that is defined via the relevant level of intellect (IQ) and the actual reading level – literacy deficits fell more than 1,5 *SEs* of measurement below the expected value. This criterion revealed dyslexia in 32,1 % of the children form the risk group and 10,3 % of the children from the control group. The authors used the same approach to evaluate the level of literacy skills in six-year-old

children in order to compare the stability of the criteria used over time. 40 % of the children who were classified as dyslexic at six years old according to the discrepancy criteria did not have reading difficulties at the age of eight. Of those who did not have difficulties at the age of six, 20 % were classified as dyslexic at the age of eight. When the first criterion (i.e. 1 SD below the control group average) was used, 91 % of the children who were assessed as dyslexic at the age of six were also dyslexic at the age of eight. They therefore selected the first criterion for further processing (SD).

If the authors use the performance of the control group as a basis, deficient performance is that which falls (usually) 1 or 2 SD below the average. The farther the measured performance is from the control group average, the more severe the deficits are. It is up to the author's decision which limit they choose for their work. If the deficit limit is 1 SD below the average, more children will exhibit the deficit since children with a milder form of the disorder will be included. On the contrary, if the limit of the deficit is set at 1,5 or 2 SD below the average, the percentage of children with a milder form of the disorder is smaller since only children with a more severe form of the disorder will be included. The selection of the limit affects the representation of children with the deficit in the group that is monitored – see e.g. Kucharská (2014), who monitors differences in the inciden-

ce of deficits according to the selected level of severity of the impairment in her group of third-grade students. Snowling and Melby-Lervåg (2016) **point out the relationship between the selected deficit criterion value and the level of prevalence of literacy difficulties** in their meta-analysis. They find differences between individual processed studies from the European language environment in terms of the prevalence of dyslexia in children with a familial risk of dyslexia ranging from 29 to 66 %, with these differences being attributed partially to the selection of the criteria that establish the deficit.

We can evaluate literacy deficits as a **whole or as partial deficits** in the area (usually) of decoding, reading, comprehension or spelling. The performance in each of these areas can be evaluated using one test or using a set of several tests whose results we then synthesize. This can be carried out using composite scores, where we add or average out the standardized scores from individual tests (e.g. Caravolas et al., 2012; Moll et al., 2016) or we can define a deficit criterion, which defines in how many tests from the battery the individual must „fail" in order for their performance to be considered as deficient in the given partial skill (see below). In either case, it can occur that an individual scores „extremely" on one of the tests but the overall performance is then evaluated as much less severe because it is evened out by

their performance in other tests from the set. This can be eliminated by using the deficit performance in only one test from the battery to define a deficit.

Something similar may occur when partial deficits are generalized into one total deficit. In practice, individual deficits are combined freely, while some children have a deficit in only one area, other children combine two different deficits and another child fails in all areas. Kucharská (2014) compares two possible approaches to establishing an overall literacy skill deficit. She evaluates literacy skills using partial skills such as decoding, reading comprehension and spelling – she sets a composite score for each of these areas, which is the average of the scores of the individual tests assessing the given area. A literacy disorder is then defined as (a) deficits in at least two of the partial areas or (b) deficits in at least one partial area. In the end, model (b) includes a larger number of children and thereby a higher incidence of literacy disorders in the group of children that is monitored.

Prediction of Literacy Skills and Deficits

As already described in the previous chapters, the research studies dedicated to the study of the development of literacy and the options for predicting disorders show high variability in several areas. Last but not least, this is also true in terms of the selection of the statistical data processing method and the manner in which prediction models are created. **Meta-analytical studies and reviews** strive to find overlaps between a larger number of research studies.

An example of such a work in finding (early) **predictors of the development of literacy** is the meta-analytical work carried out as part of the *National Early Literacy Panel* mentioned in the introduction above (for more on this study and the entire project, see NILF, 2008). More than 7000 thematically focused outputs were mapped out as part of the NELP and subsequently, 234 publications were processed in detail. All of them focus on the relationship of preschool abilities and skills with subsequent levels of literacy development. Snowling and Melby-Lervåg (2016) worked with a slightly smaller volume of data, as they went through about 300 outputs and then processed a revision and meta-analysis of 95 publications based on 21 independent studies. As opposed to the NELP, Snowling and Melby-Lervåg focused on the risk group of children with a family history of reading difficulties and focused thematically on mapping out risk and protective factors in the development of literacy difficulties prevalent in preschool age.

The NELP authors divide the original research studies into two types on the basis of their statistical processing. (1) Studies which show observed rela-

tionships between early literacy variables and conventional literacy outcomes, and which primarily work with simple correlations. The authors carry out a meta-analysis of these studies, where the results are the average correlations across all the studies for individual variables – i.e. early literacy skills and predictors. Out of the number of variables contained in the original studies, those that appeared to be a predictor in at least three studies were selected for these types of analyses. (2) Multivariate studies, which take into account the fact that variables may share predictive variance with each other. These studies use multiple regression or similar analytic techniques and since the original studies work with a high variability of combinations of control variables, the NELP authors then limit their work to analysing the overview of the conclusions of individual original studies.

Snowling and Melby-Lervåg (2016) find two basic types of research design among the original studies. (a) Cross-sectional studies are based on group comparisons between children with a familial risk of dyslexia and a control group at a given time – stage of development – and find risk factors, which are weighed against the familial risk of dyslexia (mapping what are called endophenotypes of dyslexia). The limitation

remains, however, that this approach allows one to find certain associations but does not specify any potential causality. (b) Longitudinal prediction studies work with multiple regression and related statistical techniques, which help us understand the relationship between early cognitive skills and later reading. When processing their meta-analytic studies, the authors base their findings on the internationally compiled *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA[5]) statement. The meta-analysis confined itself primarily to the calculation of the mean effect size, if a minimum of two and more studies were available for the given hypothesis. In the opposite case, they would be limited to a systematic review and report effect sizes for individual studies, which primarily had to do with longitudinal studies – similarly to the NELP – where these studies monitor a set of more variables at once (multiple studies).

The NELP authors monitor the level of conventional literacy skills on the level of coding, reading comprehension and spelling. The research synthesis they presented identified groups of ten variables which have strong to moderate relationships with at least one conventional literacy outcome and therefore are a strong to moderate predictor of later

---

Table 1. Summary of Meta-Analytic and Multivariate Results for Literacy-Related Predictor Variables with Moderate to Strong Relationships with Conventional Literacy Outcomes (NIFL, 2008, p. 67, Table 2.4)

Predictor variable	Description	Decoding	Reading comprehension	Spelling	Multivariate Significance
Alphabet knowledge	Knowledge of the names and sounds associated with printed letters	++	+	++	Yes
Phonological awareness	The ability to detect, manipulate or analyse the auditory aspects of spoken language	+	+	+	Yes
RAN of letters or digits	The ability to rapidly name a sequence of repeating random letters or digits	+	+	NA	Yes
RAN of objects or colours	The ability to rapidly name a sequence of repeating random sets of pictures of objects or colours	+	+	+	Yes
Writing or writing name	The ability to write letters in isolation on request or to write one's own name	+	+	+	Yes
Phonological short-term memory	The ability to remember spoken information for a short period of time	--	+	+	Yes
Oral language	The ability to produce or comprehend spoken language, including vocabulary and grammar	+	+	+	Sometimes
Concepts about print	Knowledge of print conventions (e.g. left-right, front-back) and concepts (book cover, author, text)	+	++	+	Sometimes
Visual perception	The ability to match or discriminate visually presented symbols	--	--	+	No
Print awareness	Tasks combining elements of AK, concepts about print, and protodecoding	--	+	NA	NA

++   strong relationship (0,5 or more) based on zero-order correlations
+   moderate (0,3-0,49) relationship based on zero-order correlations
--   weak relationship (0-0,29) based on zero-order correlations
NA   no relevant data available for analysis

conventional literacy skills. Moreover, six of them retain their predictive strength even if other contextual factors are taken into consideration (e.g. IQ, age, socioeconomic status or some of the other variables that were monitored), as multivariate studies confirm. An overview of these variables is provided in Table 1 (below). As part of the secondary analyses, they then monitor whether the predictive strength changes according to the concretization of the given variable where several authors monitor different aspects of a specific scale. While they worked with a composite score for individual variables as part of their primary analyses, here they focused on various aspects of oral language and phonological awareness. They found that more complex aspects of oral language, such as grammar, definitional vocabulary and listening comprehension, have stronger predictive relationships with later conventional literacy skills (specifically with decoding and reading comprehension, and no data was available for spelling) than simple vocabulary knowledge does. Phonological awareness, which reflects the level of linguistic complexity achieved (phoneme, syllable, rhyme) or the type of cognitive operation required (identification, synthesis, analysis) can similarly also be measured in different ways.

In their meta-analytical study, Snowling and Melby-Lervåg (2016) confirmed an increased risk of the onset of dyslexia in children from the familial risk group across all language environments, which was contained in the original studies. Children at risk of dyslexia show slower development of speech and language skills, they acquire literacy and other skills crucial for the development of literacy at a slower rate and signs of dyslexia are already apparent when they are of preschool age. Similarly to the NELP, Snowling and Melby-Lervåg confirm that phonological awareness, knowledge of symbols (letters) and rapid naming (RAN) are strong predictors of literacy skills both in children from the familial dyslexia risk group and in children from the control group. The authors nevertheless find certain specifics as well. Letter knowledge is a long-term predictor in children with a familial risk because, as compared to children from the control group, these children reach what is called the „ceiling" in a certain skill later and that is why RAN appears to be the more significant unique predictor in this risk group.

Conclusion

The authors dealing with the study of acquiring literacy, the development of its disorders and the possibilities of their prediction show vast variability in their research works. This variability affects several different areas. The authors already differ in the way that they approach the topic in methodologically terms – whether they monitor specific develo-

pmental advances at a given moment or whether they go back into the past using a retrospective approach as part of longitudinal research. We have research available that monitors the development of literacy and incidence of disorders in the general population. Several authors, however, focus on specific risk groups (children with a familial risk of dyslexia, children with impaired language and speech development) in which the incidence of individuals with a literacy disorder is higher. When studying literacy, we monitor the development of literacy skills *per se* but also focus on the development of other skills and abilities which are fundamental for the development of literacy – the researchers' interest is therefore shifting to the preschool age, where they monitor the development of predictors and early literacy skills prior to the start of formal education. They continue with the monitoring in the first years of elementary school, before it is possible to assess the level of literacy already achieved – the authors differ in when they assess these skills here as well, however. This often depends on how specifically the authors assess the literacy skills (decoding, reading comprehension or spelling). We can encounter differences in the prevalence of literacy disorders among the studies, which is partially due to the actual definition of a disorder (e.g. specific learning disorders, performance below the average of the control group) or a limit set for a defining criterion. This is why we may encounter different conclusions in various studies. Reviews and meta-analytical studies strive to unify them, pointing to phonological awareness, letter knowledge and RAN as the emerging strongest predictors of the development of literacy.

References

American Psychiatric Association (APA, 2013). *DSM-5: Diagnostic and statistical manual of mental disorders*. Fifth edition. Arlington, VA: American Psychiatric Association.

Berninger, V. W. (2001). Understanding the 'lexia' in dyslexia: A multidisciplinary team approach to learning disabilities. *Annals of Dyslexia*, *51*, 23–48.

Bigozzi, L., Tarchi, C., Pezzica, S., & Pinto, G. (2016). Evaluating the predictive impact of an emergent literacy model on dyslexia in Italian children: a four-year prospective cohort study. *Journal of Learning Disabilities*, *49*(1), 51–64.

Caravolas, M. (2005). The nature and causes of dyslexia in different languages. In M. J. Snowling, & C. Hulme, *The Science of Reading: A Handbook* (336–355). Oxford: Blackwell Publishing.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., Salas,

N., Schöffelová, M., Defior, S., Mikulajová, M., Seidlová-Málková, G., & Hulme, Ch. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, *23*(6), 678–686.

Česká školní inspekce (2013). *Koncepční rámec čtenářské gramotnosti PISA*. Dostupné z http://www.csicr.cz/Prave-menu/Mezinarodni-setreni/PISA/Metodika-setreni/Koncepcni-ramec-ctenarske-gramotnosti-v-setreni-PI

Ferreiro, E., & Teberosky, A. (1982). *Literacy before schooling*. Ports mouth, NH: Heinemann Educational Books.

Hartl, P., & Hartlová, H. (2010). *Velký psychologiký slovník*. Praha: Portál.

Helus, Z. (2012). Reflexe nad problémy gramotnosti: Komentář k monotematickému číslu. *Pedagogika*, *62*(1–2), 205–210.

Kucharská, A. (2007). Riziko dyslexie. *Zpravodaj Pedagogicko-psychologické poradenství*, *48*, 28–43.

Kucharská, A. (2014). *Riziko dyslexie.* Praha: Univerzita Karlova v Praze, Pedagogická fakulta.

Langmeier, J., & Krejčířová, D. (2006). *Vývojová psychologie*. 2. Praha: Grada.

Medřická, T. (2019). *Variabilita vývoje počáteční gramotnosti u dětí s rizikem dyslexie: Predikční modely gramotnostních deficitů* (Disertační práce). Praha: Univerzita Karlova, Pedagogická fakulta.

Moll, K., Thompson, A. P., Mikulajova, M., Jagercikova, Z., Kucharska, A., Franke, H., Hulme, Ch., & Snowling, J. M. (2016). Precursors of reading difficulties in Czech and Slovak children at risk of dyslexia. *Dyslexia*, *22*, 120–136.

National Institute of Child Health and Human Development. (NICHD, 2000). *Report of the National Reading Panel: Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: U.S. Government Printing Office.

National Institute for Literacy. (NILF, 2008). *Developing early literacy. Report of the National Early Literacy Panel. A scientific synthesis of early literacy development and implications for intervention*. Washington, DC.

Oberski, D. (2016). *Mixture models: Latent profile and latent class analysis. Modern statistical methods for HCI*. Springer Nature, 275–287.

Přinosilová, D. (2007). *Diagnostika ve speciální pedagogice. Texty k distančnímu vzdělávání*. Brno: Paido.

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skills. *Child Development*, *74*(2), 358–373.

Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin*, *142*(5), 498–545.

Torppa, M., Georgiou, G. K., Lerkkanen, M., Niemi, P., & Poikkeus, A. (2016). Examining the simple view of reading in a transparent orthography: A longitudinal study from kindergarten to grade 3. *Merrill-Palmer Quarterly: Journal of Developmental Psychology*, *62*(2), 179–206.

Vágnerová, M., & Klégrová, J. (2008). *Poradenská psychologická diagnostika dětí a dospívajících*. Praha: Karolinum.

Mgr. et Mgr. Tereza Medřická, Ph.D.
doc. PhDr. PaedDr. Anna Kucharská, Ph.D.
Faculty of Education, Department of Psychology
Charles University
medricka.t@gmail.com
anna.kucharska@pedf.cuni.cz