Agreement between the Assessments of Parents and Preschool Teachers of Different Developmental Areas of Preschoolers

Velki, Tena; Romstein, Ksenija; Ružić, Valentina

Source / Izvornik: **Croatian Journal of Education : Hrvatski časopis za odgoj i obrazovanje, 2021, 23., 1139 - 1169**

Journal article, Published version
Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

https://doi.org/10.15516/cje.v23i4.4241

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:141:844225

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-04-24**

Repository / Repozitorij:

FOOZOS Repository - Repository of the Faculty of Education

Croatian Journal of Education
Vol.23; No.4/2021, pages: 1139-1169
Original research paper
Paper submitted: 15th December 2020
Paper accepted: 28th January 2021
https://doi.org/10.15516/cje.v23i4.4241

Agreement between the Assessments of Parents and Preschool Teachers of Different Developmental Areas of Preschoolers

Tena Velki[1], Ksenija Romstein[1] and Valentina Ružić
[1]*Faculty of Education, Josip Juraj Strossmayer University of Osijek*

Abstract

People who know the child well and spend a significant amount of time with him/her (usually parents/caregivers and preschool teachers) are crucial for identifying developmental delays in preschool children, since the decision about the need for intervention or treatment is often based largely on their assessments. The aim of this study[1] was to examine concordance, i.e. agreement between the parents and preschool teachers' assessments of different aspects of preschool children's development. In total, 60 pairs of raters (parents-preschool teachers) participated in the study and gave assessment for children aged from 30 months to 54 months (42.6% were boys and 57.4% were girls). Both parents and preschool teachers filled out the Ages & Stages-3 Questionnaire (ASQ-3; Squires et al., 2009) translated into Croatian language for the purpose of the research project. Interrater agreement was tested for 5 developmental areas: communication, gross motor, fine motor, problem solving, personal-social area and general concerns. Intraclass correlation coefficients have shown poor to moderate correlation for interrater agreements (ICC=0.36-0.84), while interrater reliability measured by Cohen›s Kappa has shown fair to moderate agreement (κ=0.31-0.61), depending on developmental area. Ages & Stages-3 has not proved to be a completely reliable measurement instrument, where it is necessary to use

[1] Data used for this paper were obtained in the research project *Paradigm shift in teaching and learning in early childhood: constructivism and developmental approach vs. behaviorism* (research project of J. J. Strossmayer University of Osijek, INGI-2015-29).

different raters as well as additional instruments for a more accurate assessment of developmental delays.

Key words: *Ages & Stages-3; development of preschoolers; developmental delays; interrater agreement; interrater reliability.*

Introduction

According to the World Health Organization (WHO, 2012), a child at neurorisk is the one exposed to risk factors during pregnancy, childbirth or infancy. The WHO and the World Bank estimated that more than one billion people, or approximately 15% of the world's population, live with some type of disability. Currently, however, there are no comprehensive data on the exact overall number of children with developmental disabilities due to various definitions and diagnostic methods used in different countries (WHO, 2012).

It is estimated that 10-15% of live-born children can be considered as children at neurorisk (Boban, 2016). The recognition of the importance of the first years of life, in which the child's motor and cognitive abilities develop intensively, followed by (or in parallel with) the child's emotional and social development, has been encouraged in societies around the globe. Modern science provides strong evidence that early experiences shape the foundations for lifelong learning, behavior, and mental health, and that development of abilities is a gradual process in which simpler skills help shape more complex ones (Ljubešić, 2016). Identifying and monitoring children is crucial for early detection of neurodevelopmental delays and inclusion within the early intervention system. Namely, learning and development in the first years of life can be influenced by interventions in various developmental areas (Guralnick, 2005), with long-term positive effects.

The importance of early intervention is increasingly recognized by various experts (doctors, psychologists, educational rehabilitators, social workers, etc.), but also by government officials and stakeholders, both in the world and in Croatia. In the Republic of Croatia, early intervention was included in the legal system for the first time in 2011 (Milić Babić et al., 2013) as one of the forms of social support for families. The Social Welfare Act (Official Gazette 157, 2013, Art. 84, para. 1) states that early intervention is "a social service that includes professional stimulating assistance for children, and professional and counseling assistance for their parents, including other family members and foster parents of children, in the case of any identified developmental risk for developmental disability". The legislative framework in Croatia provides the possibility of supporting the child and the family through early intervention up to the school starting age (Social Welfare Act, 2013, Art. 84, para. 3). Establishing legal frameworks is a fundamental precondition for the systematic implementation of early intervention, with the main goal of providing support to the families in the process of raising children at risk.

In the last twenty years, in the context of early intervention, screening for developmental delays emerged as one of the preconditions for the adequate support of child's development. The aim of screening is early identification of children with developmental delays and their enrollment in the early intervention system. To emphasize the importance of screening for developmental delays in early childhood, the American Academy of Pediatrics (AAP) emphasizes that "early identification of developmental disorders is key to the well-being of children and their families" (Council on Children With Disabilities, 2006). Ljubešić (2016) states that children with neurodevelopmental risks or disabilities who do not receive an appropriate intervention support in early age are actually doubly deprived. On the one hand, under the influence of their disability, they spontaneously "withdraw" from the environment and are respectively less exposed to adequate stimulation. On the other hand, their parents are stressed due to the obligations concerning child's functioning, which leads them to have less responsive interactions with the child than they would have in other circumstances.

Early identification is actually a continuous process of gathering information about a child from a variety of sources, including direct observations from parents/caregivers and professionals (Squires et al., 1996). When we talk about the primary health care system, developmental disabilities are usually noticed either during regular check-ups by a doctor/pediatrician or propelled by parental concerns related to some aspect of the child's development, yet sometimes this concern is expressed by a third party (e.g. preschool teacher) (Choo et al., 2019). As developmental risks increase with the child's chronological age, it is recommended to examine children for developmental delays at least once a year, in order to monitor their development. Likewise, periodic developmental screening is recommended for all children.

The need to involve different professionals in the assessment of development increases as a child grows up. That enables each professional to assess developmental aspects and possible delays in the child's development from their own perspective. Screening can be performed as part of the work of pediatric services or other professionals (Council on Children with Disabilities et al., 2006), in order to ensure screening for disabilities and involvement in early intervention. However, when identifying potential disabilities, it is important to collect data not only from medical doctors and specialists, but also from the person closest to the child (parents/caregivers, families, preschool teachers) because they are able to observe the child in various environments and provide information from different perspectives.

It should be emphasized that screening for developmental delays does not lead to diagnoses but to identification of developmental areas where the child's level of functioning is lower than expected when compared to chronological age, considering the need to include the family in the early intervention support. The potential goal of early screening is twofold: to include as many children as possible, with as little "permeability" as possible in terms of not recognizing children with developmental disabilities. The second goal is to reduce the number of so-called "false positive" cases, i.e. reduction in the number of children who are identified as at risk and who are

referred to further diagnostics, without having actual developmental disability. It is thereby necessary to compare two measures - the sensitivity and the specificity of the screening measure. The sensitivity measure shows how good the screening method is in recognizing people with a particular disease, i.e. disability, while the specificity measure shows how good it is in recognizing those without it, thus eliminating the need for further assessment of those individuals (Parikh, 2008). These two measures are reciprocal, so the decision on which one to choose should be based on the estimated benefits. By increasing the sensitivity of the screening measure, we increase the number of children who are identified as potentially at risk and who are referred for a more detailed assessment, i.e., the possibility of "missing" such a child is reduced. This approach in return increases the likelihood of recognizing children with potential disabilities, but also increases the possibility of identifying children who do not actually have a disability. Also, increasing specificity ensures that we do not identify any child as a child with potential disabilities if they do not actually have one, thus reducing the potential time and material costs of the assessment. Sensitivity is also called assessment of *true positive* cases while specificity is called assessment of *true negative* cases (Nunnally & Bernstein, 1994). Completely accurate estimation would have a 100% sensitivity and a 100% specificity, but no measure has such a characteristic because a certain degree of error is always associated with the assessment.

The use of standardized assessment methods helps to achieve the goal of identifying children at risk as accurately as possible, but even with such measures there are certain problems. One of them is the question of the reliability of assessments in general, and the other is the agreement of assessments between different assessors. Namely, it is often the case that one assessor shows a certain level of concern regarding the child's development, while at the same time other adults who are in contact with the child do not express those concerns. The cause of such disagreements could be in the diversity of situations in which potential observers observe the child, but also in the child's experiences - the child feels safe, protected, free and relaxed in different situations so his behavior differs accordingly. In fact, this does not necessarily mean that one observation is wrong, but instead points out that situational factors should be considered, and information from different sources should be combined. Therefore, it is important to systematically monitor, examine and compare child assessments collected from different sources in order for the screening for developmental disorders to be as effective as possible. Previous research shows that parental assessments are quite reliable (Bodnarchuk & Eaton, 2004; Ratner & Silverman, 2000; Guralnick, 2005), but also that they can vary considerably depending on which ability is being assessed, i.e., in which environment this ability is manifested. On the other hand, there are certain differences in the assessments of individual parents - mother and father (Achenbach et al., 1987) which are explained by the fact that each of them perceives the child in different circumstances and can be focused on different aspects of development. Research conducted on twins (van der Valk et al., 2000) has shown that the discrepancy between parental assessments reflects the uniqueness of data provided by only one parent, rather than the unreliability or bias

of one of the assessors. Meta-analyses of assessments obtained in numerous studies show that the average correlation between assessments of two parents was 0.59, while the correlation between pairs of preschool teachers was 0.64 and between parents and preschool teachers 0.27 (Achenbach et al., 1987).

Taking into consideration that a large number of measures used to examine children's development also includes certain critical values (results that indicate that a particular child is suspected for developmental delay), it is possible for some children to be classified as suspected based on the assessment of one assessor and not based on the assessments of the other one. In other words, it is necessary to keep in mind that the screening result may vary depending on who is assessing the child. Therefore, it is important to analyze the correlation of classifications of different assessors. One approach to check classification accuracy is to calculate the relative risk odds ratio (Fleiss, 1981), which is used in epidemiological research. This ratio indicates the probability that in a particular group of people being assessed, those identified as at risk for a particular disorder will indeed have that disorder, relative to the likelihood of the presence of that condition in a group of individuals not identified as at risk. The comparison between the odds ratio for those who have and those who do not have that risk factor (certain critical score) is expressed as the ratio of the probability of the presence of an outcome if that risk factor is present and the probability of the presence of that outcome when that risk factor is absent.

The aim of this research was to examine the agreement between the assessments of parents and preschool teachers on different developmental aspects of preschool children. Cohen's kappa coefficient was used to measure inter-rater reliability or agreement between two raters (parents and preschool teachers). It represents the extent to which the data collected in the study are correct representations of the variables measured. For this purpose, kappa coefficient represents reliability of decision that a child is at risk, depending on who is doing the assessment. Agreement for 5 developmental areas will be compared: communication, gross motor, fine motor, problem solving and the personal-social area, as well as the agreement on general assessments, i.e. concerns about the child's development. An additional goal of the research was to compare classification of children using set up criteria (mild and severe developmental delay) between parents and preschool teachers as assessors. Since those assessments are used for further decisions regarding interventions and treatments of children, it is important to see how different assessors agree in deciding whether the child is at risk for developmental disability.

Method

Participants

The study involved 60 parents (95% of them were mothers) of children aged 30 months to 54 months, and 60 preschool teachers of the same children. Analysis of interrater agreement was done for 60 pairs of assessments (parent-preschool teacher), of which 41.7% were assessments of male children and 58.3% of female children. All

assessed children were included in regular preschool groups and their parents gave consent for the assessments. The average age of parents was M=29.51 (SD=6.31), while preschool teachers (100% female) had an average length of service M=15.44 (SD=10.68). Details about the sample are shown in Table 1.

Table 1
The selected sample of children for the evaluation of interrater agreement (parent-preschool teacher assessments)

Age of children	*f*	%
36 months	7	11.7
42 months	16	26.7
48 months	17	28.3
54 months	20	33.3
Total	60	100.0

Instrument

Ages and Stages-3 Questionnaire (ASQ-3; Squires et al., 2009) was used in the study. This questionnaire was translated into Croatian language for the purposes of the research project, and the consent of the authors was obtained. The ASQ-3 consists of a series of items that measure different aspects of development. In the first part of the questionnaire, parents/caregivers/preschool teachers provided basic demographic data (child's date of birts, gender and relationship to that person). The second part of the questionnaire contained 30 questions divided into 5 subscales, i.e. developmental areas: communication (k=6), gross motor (k=6), fine motor (k=6), problem solving (k=6), and personal-social area (k=6). The participants answered by choosing one of the three possibilities as the most appropriate for their child: "yes", "sometimes" and "not yet". For each scale, a total score was calculated (possible range 0 to 60). Based on the total score (depending on the age and developmental area), two main criteria were defined: a mild delay, which suggested the presence of fewer developmental discrepancies in accordance with age, and severe delay, which suggested the significantly lower score in accordance with age. The third part of the questionnaire consisted of 10 questions about parental/caregivers'/preschool teachers' concerns in different areas. A linear combination of these 10 questions presents a general concern where the sum of the first 5 questions (each positive answer brings 1 point) is subtracted from the sum of the second 5 answers (possible range of answers from 5 to 5, with a higher value indicating a lower level of concern). The questionnaire is divided into age spans from 2 months to 60 months of age. For the purposes of this research, the following versions were used: 36 months (from 34 months and 16 days to 38 months and 30 days), 42 months (from 39 months and 0 days to 44 months and 30 days), 48 months (from 45 months and 0 days to 50 months and 30 days) and 54 months (from 51 months and 0 days to 56 months and 30 days).

Table 2 shows internal Cronbach's α consistency coefficient for all developmental areas, together for parents and preschool teachers (total) and separately for each group.

Table 2
Cronbach's α internal consistency coefficients

Developmental areas	α total	α parents	α preschool teachers
Communication	0.60	0.58	0.75
Gross Motor	0.59	0.66	0.54
Fine Motor	0.65	0.74	0.61
Problem Solving	0.70	0.72	0.69
Personal-Social	0.47	0.31	0.49

The results show low internal consistency for the Personal-Social area and relatively low consistency for the Communication and Gross Motor area. Satisfactory internal consistency was obtained for the Problem Solving and Fine Motor areas, for both groups of raters, parents and preschool teachers ($\alpha > 0.60$), although these consistencies were not high either.

Procedure

The research was conducted in 5 Croatian counties: Osijek-Baranja County, Brod-Posavina County, Varazdin County, Virovitica-Podravina County and Vukovar-Srijem County. Parents were given a questionnaire in kindergartens, which they could fill out at home, and later return to the preschool teachers. Preschool teachers were given the same questionnaire to fill out. About one third of the parents returned the completed questionnaire. The research was voluntary for both parents and preschool teachers. Each assessment took about 20 minutes. All parents were offered feedback on the child's overall profile based on the ASQ-3 score, and they were offered counseling on appropriate pedagogical activities for children suspected of having developmental delays (which was voluntarily accepted by both parents and preschool teachers).

Results and Discussion

Only those questionnaires whose assessments of the child were obtained from both parents and preschool teachers were taken into analysis. The results were formed for all scales (subscales), and descriptive statistics for these scales can be seen in Table 3. The values obtained on the parents' assessments (marked with the letter P) and the values of the preschool teachers' assessments (marked with the letter T) are shown separately.

As shown in Table 3, the asymmetry indices (skewness and kurtosis) do not exceed +/- 4 in any group, so parametric statistics was applied in further analyses. The theoretical range of responses on the subscales ranged from 0 to 60, with a higher number indicating a higher stage of development for a particular developmental area. Table 3 shows that the minimum number of points (0 points) was not achieved on any subscale, while the maximum number of points was achieved on all subscales. Such results speak in favor of a slightly lower sensitivity of the measured instrument.

Also, the indices of deviation from the normal distribution, especially for parental assessments, show higher values, close to the marginal ones (especially for the subscales Communication, Gross Motor, and Personal-Social), which also speaks in favor of their poorer applicability in practice.

Table 3
Descriptive statistics for measured variables

Developmental areas	*N*	*Min*	*Max*	*M*	*SD*	*Sk*	*Ku*
Communication P	60	25.00	60.00	55.35	6.67	-2.25	4.00
Gross Motor P	60	20.00	60.00	55.00	8.18	-2.23	3.89
Fine Motor P	60	5.00	60.00	48.08	13.22	-1.19	0.90
Problem Solving P	60	5.00	60.00	51.92	11.16	-2.24	4.00
Personal-Social P	60	35.00	60.00	53.83	6.73	-1.16	0.75
Overall Concerns P	60	0.00	4.00	3.08	1.01	-1.08	0.61
Communication T	60	30.00	60.00	54.08	7.99	-1.44	1.16
Gross Motor T	60	20.00	60.00	51.58	9.59	-1.38	1.72
Fine Motor T	60	15.00	60.00	48.58	11.68	-0.96	0.21
Problem Solving T	60	5.00	60.00	50.50	12.06	-1.56	2.65
Personal-Social T	60	25.00	60.00	52.85	9.08	-1.56	1.93
General Concerns T	60	1.00	5.00	4.37	1.05	-1.71	2.42

P - Parents; T - Preschool Teachers; *M* - Mean; *SD* - Standard Deviation; *Sk* - Skewness; *Ku* - Kurtosis

Parental and preschool teachers' agreement on child developmental assessments

In accordance with the aim of the research, the agreement of the parent-preschool teacher assessors for 5 developmental areas was examined: communication, gross motor, fine motor, problem solving, and personal-social area, as well as general concerns for the child's development. Table 4 shows the intraclass correlation coefficients as a measure of agreement of assessments from two sources.

Table 4
Interrater agreement (parents-preschool teachers) for different developmental areas and general concerns

Average measures	Intraclass Correlation (*ICC*)	Confident Interval Lower Bound	Confident Interval Upper Bound	$F_{(60,60)}$
Communication	0.783	0.638	0.870	4.60**
Gross Motor	0.565	0.275	0.739	2.30**
Fine Motor	0.840	0.733	0.904	6.24**
Problem Solving	0.827	0.712	0.896	5.70**
Personal-Social	0.361	-0.066	0.616	1.56*
General Concerns	0.489	0.124	0.702	1.96**

* $p < 0.05$; ** $p < 0.01$

All obtained subscales show a statistically significant agreement of assessors, whereas the personal-social area and general concern for the child's development show poor

agreement (ICC < 0.5). Also, gross motor area shows moderate agreement (ICC = 0.50-0.75), while communication, fine motor and problem solving areas show good agreement (ICC = 0.75-0.90). Considering the obtained results, it can be concluded that parents and preschool teachers mostly agree in assessments when it comes to the three areas of child development, i.e. communication, fine motor and problem solving, assessing the level of child development equally for those areas. The lowest agreement was obtained for the personal-social area, which is expected, given the environment in which the child was when the assessment was being conducted. Parents assessed the child within the family environment, where the social component was actually reduced to a few close family members with whom the child often interacts (e.g. joint activities and play). However, preschool teachers assessed the same component in the child's environment with peers, where other aspects of social interactions were more prominent (e.g. coping in a group of peers, starting and maintaining communication with peers, sharing toys, etc.). The second potential reason for differences in assessment is in the formal education of preschool teachers, which attaches great importance to child's social skills, especially in the dimension of independence and relationships with others, in various courses (e.g. developmental psychology, pedagogy etc.). Parents who do not have formal education in the field of early childhood education and those who do not have more children in the family find it more difficult to assess appropriate behaviors in relation to the chronological age of the child.

Parental and preschool teachers' agreement in child classification

Considering that the applied questionnaire also contains critical values for child delays at two levels (mild and severe delays) for five developmental areas, the reliability of the assessors' agreement was checked using the Kappa coefficient (Table 5).

Table 5

Interrater reliability (Cohen's Kappa coefficient) for categorization of children in three groups (typical development, mild delay and severe delay)

Developmental areas	K
Communication	0.306**
Gross Motor	0.159
Fine Motor	0.508***
Problem Solving	0.609***
Personal-Social	0.102

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

The results show reliable agreement between parents and preschool teachers only for 3 developmental areas (communication, fine motor and problem solving), with fair agreement obtained for the communication area, moderate for fine motor and substantial for the problem solving area. When a deviation is taken into account, which would mean that parents and/or preschool teachers should introduce developmentally appropriate incentive activities (for children categorized as suspected in a certain

developmental area) or refer for further assessment due to potentially more serious developmental deviation (for children categorized as critical in a certain development area), parents and preschool teachers agree much less than in the general assessment of the development areas (intraclass correlation coefficients for all development areas were statistically significant, Table 3).

In addition to the different environments in which the child is being assessed (kindergartens and families), different assessors may have different levels of competence to perform the child assessment. Preschool teachers, due to their formal education, have a certain level of knowledge about the expected development of preschool children, as well as about possible delays, which can lead to differences in the assessments of parents and preschool teachers, especially for parents without similar education. However, just knowing a child, and his or her development and progress, can affect assessment, especially for preschool teachers who have not been long with the educational group in which they assess a specific child, for whom the parent, regardless of their education, may be able to give a more reliable assessment.

Given the dynamic of children's development at this age, it would be useful to compare Kappa coefficients across the age span. Unfortunately, given the small number of pairs of assessors in each of the four age categories, such a comparison on the data collected in this study was not possible. In fact, in some categories, due to the small number of assessors, the assessments of all children of a certain age are in the same category (e.g. the preschool teacher assesses all children of that age as typical), so the kappa coefficient according to age is not a good enough indicator. However, certain trends in age changes have been observed which show that kappa coefficients differ with age in individual subscales, probably because the behaviors included in a particular assessment change with age and appear differently in specific situations.

The original US questionnaire ASQ-3 contains two cut-off criteria according to which children's developmental delays are divided into three groups: children with typical developmental pattern (no developmental delays); children with mild delay (children who show mild delays from the population average in one of the 5 developmental areas, and for whom stimulating activities are recommended); and children with severe delay (children who show significant delays compared to the average population in one of the 5 developmental areas, for whom professional help is recommended, i.e. referral of the child for further assessment). Given that the standardization of ASQ-3 is underway in the Republic of Croatia, and preliminary results are in the process of publication (Velki & Romstein, 2020), Tables 6 and 7 show the children delays based on American cut-off criteria.

A comparison of the odds ratio (OI) was made in the manner suggested by Fleiss (1981), for those who have and those who do not have a risk factor (a certain critical score) according to the assessment of one of the assessors. The calculated odds ratios show the likelihood that children classified as delayed (mild or severe) based on parental assessments will be equally classified based on preschool teachers assessments. Also,

ratios show that children classified as delayed (mild or severe) based on preschool teachers' assessments were equally classified based on parental assessments. The results are shown in Table 6.

Table 6

Proportion of children that the first rater assessed as children with developmental delays (criterion) and the second as either children with developmental delays or children with typical development (odds ratio - OR)

	OR for parental criterion of children with delay vs. preschool teachers' assessment agreement		OR for the preschool teacher criterion of children with delay vs. parental assessment agreement	
Developmental areas	Delay	Typical	Delay	Typical
Communication *	2/2 = 1.00	–	2/7 = 0.29	5/7 = 0.71
Gross Motor	3/5 = 0.60	2/5 = 0.40	3/12 = 0.25	9/12 = 0.75
Fine Motor	5/10 = 0.50	5/10 = 0.50	5/6 = 0.83	1/6 = 0.17
Problem Solving *	9/9 = 1.00	–	9/15 = 0.60	6/15 = 0.40
Personal-Social	1/4 = 0.25	3/4 = 0.75	1/8 = 0.13	7/8 = 0.88

*$p < 0.01$

Differences in the OR proportions of delayed children for each assessor were examined with the t-test for proportions. The results show that the difference was significant for communication ($t = 4.14$; $p < 0.01$) and problem solving ($t = 3.16$; $p < 0.01$), while there was no significant difference in the remaining three aspects of development. It can be seen that the risk ratios are different for certain aspects of development, with only the aspects of communication and problem solving assessed as suspected for all children by both parents and preschool teachers. The lowest OR was obtained for the parental assessment of personal-social area (only one in 8 children identified by preschool teachers is also identified by the parent).

Furthermore, it is evident that the greatest difference is obtained between the assessment of children with delays in the personal-social area and in the fine motor area, which confirmed that these developmental areas are "the most difficult" to assess. The latter indicates the existence of the impact of formal education of preschool teachers on the assessment of these areas. The percentages of children classified into two specific categories (mild and severe delays) based on the assessments of both assessors are shown in Table 7.

When it comes to children with severe delays, parents and preschool teachers substantially agree with the assessment of delay (agreement is complete for ages 36 and 48 months). The results show that the assessors' agreement is higher when it comes to classification of children with severe delays, probably because the criteria are more distinctive and the assessment is less influenced by situational factors. More disagreement was found in the assessors' classification of children with mild delays probably because it is sometimes difficult to assess whether a child's behavior is within a normal developmental range or whether it shows a milder delay compared to the population average.

Table 7

Percentage of children that parents and preschool teachers estimated to deviate from typical development for different developmental areas

Age	Developmental areas	Communication		Gross motor		Fine motor		Problem Solving		Personal-Social	
		P	T	P	T	P	T	P	T	P	T
36 months	Severe delay	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %
7 pairs	Mild delay	0.00 %	0.00 %	0.00 %	42.86%	14.29 %	0.00 %	0.00 %	14.29 %	14.29 %	0.00 %
42 months	Severe delay	0.00 %	0.00 %	0.00 %	6.25 %	6.25 %	0.00 %	6.25 %	0.00 %	0.00 %	0.00 %
16 pairs	Mild delay	6.25 %	6.25 %	0.00 %	0.00 %	6.25 %	6.25 %	0.00 %	12.50 %	18.75 %	12.50 %
48 months	Severe delay	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %
17 pairs	Mild delay	0.00 %	5.88%	5.88 %	5.88 %	23.53 %	17.65 %	0.00 %	0.00 %	0.00 %	0.00 %
54 months	Severe delay	5.00 %	5.00 %	10.00 %	15.00%	0.00 %	5.00 %	10.00 %	20.00 %	0.00 %	15.00 %
20 pairs	Mild delay	0.00 %	20.00 %	10.00 %	20.00%	15.00 %	5.00 %	30.00 %	40.00 %	0.00 %	15.00 %

P – Parent, T – Preschool Teacher

For data interpretation, it is necessary to take into account that there are actually a small number of children in the population who show significant deviations, i.e. meet the criteria for diagnosis (acc. to Boban (2016), there is an estimation of about 10-15% of neurorisk children in the population). In practice, we can find a greater number of children who show smaller deviations (mild delays according to the described categorization) or milder temporary delays in the development (Guralnick, 2005). Such data are understandable if we take into account that the child's development does not take place in a linear path, equally for all children, but is rather dynamic, so milder discrepancies are expected. In addition to explaining the reasons for discrepancies in the assessments, it should be taken into account that parents and preschool teachers have different previous experiences and knowledge in assessments and they compare the child's progress with different reference groups. Results show that teachers estimated overall more children as deviate than parents, which stresses the importance of including preschool teachers' observations in everyday practice of assessing child development. Preschool teachers have experience in working with a larger number of children and the opportunity for comparing behavior of one child to a group of peers (unlike parents, who usually have the opportunity to observe a much smaller number of children, usually in the family surroundings). That is why the preschool teachers can be seen as an important source of information about the child's development. The results of this kind of research can point out the importance of preschool teachers' assessments to the parents and other experts.

Similar results have been obtained in studies in other countries (e.g., Kerstjens et al., 2009; Ortiz-Leon et al., 2018). Given the obtained results, which indicate a rather weak agreement in the assessments of developmental delays by parents and preschool teachers, it is recommended to use both assessments when identifying potential

developmental delay of the child. By using both assessments in making decisions about inclusion of children in the treatment, we decrease a chance of not-identifying, i.e. "missing" some of the information that can indicate a need for intervention. Having in mind the importance of early identification of possible risks, these results stress the need for combining different assessors, rather than discarding either one source of information. In addition, the observed differences between assessors can make all persons involved in a child's life more sensitive to the fact that they observe the child in selected occasions and so they can miss certain behaviors that are shown in different situations.

In addition, it is interesting to note that a lower agreement of the assessors was obtained precisely on the subscales that showed lower reliability. For subscales with lower internal consistency, agreement measures between assessors are also low or not statistically significant at all. Therefore, it is not recommended to use the subscales Communication, Gross Motor, and especially Personal-Social separately as a measure of screening for developmental deviations, but to combine these subscales with other assessments. As already mentioned, the reasons for lower reliability and weaker agreement of raters may be the inadequate training and different experience of assessors, which is a common problem in research with different assessors (e.g. Hsue et al., 2014). However, it is also possible that lower reliability and higher measurement error reflect the characteristics of the instrument itself. Furthermore, the influence of cultural factors on the assessments of preschool teachers and parents is possible, which should be taken into account when using foreign, non-adapted screening tests. For example, different results and sometimes norms were obtained for different countries such as Brazil (Correia et al., 2019) or India (Ali et al., 2011). Since it is not clear whether those results are a consequence of psychometric characteristics of the scale or of different circumstances of assessments (as explained earlier), those scales should be used with extreme caution.

In light of these findings, with the aim to make the assessment of children more reliable, which can lead to earlier inclusion in treatments, it is important to further explore new ways of assessing child's behavior as well as to train different people in making those assessments and make them aware of strengths and limitations of each of them.

Methodological limitations

The conducted research has certain methodological shortcomings that should be considered when interpreting the obtained results. The translated American instrument (Ages & Stages-3) was used in the research, for which the validation and standardization on the population of children in the Republic of Croatia is still in progress (Velki & Romstein, 2020). Given that the cut-off criteria obtained on the American population of children were used to calculate the measures of the agreement between assessors (ICC, Kappa, OR), the question is whether identical data would have been obtained

if standardization and certain cut-off criteria for the population of children in the Republic of Croatia had been made. Furthermore, although the survey was conducted in 5 counties, the obtained sample is not representative for the whole Croatia. Not all existing versions of the questionnaire were used for all age ranges (2 to 60 months) and a relatively small number of assessor pairs were included in the survey (especially for 36-month age estimates). Future research should certainly include a more representative sample that covers the entire age range with a larger number of assessor pairs, and it would be desirable to revise subscales with low reliability as well as to standardize the criteria for the population of Croatian children.

In addition, this study did not take into account the demographic characteristics of the assessors, such as the age, formal education level and socioeconomic status of the parents, which may also have an effect on the early recognition of developmental delays. Furthermore, assessments from only one of the parents were collected, and previous research emphasizes that parents also differ when it comes to observing child's development. Neale and Cardon (1992) emphasize the importance of such different perceptions and identification of standards for different types of assessments. The biases of individual assessors, i.e. differences in their assessments, are most often identified as measurement errors and further attributed to the low reliability of the assessment techniques and used screening measures. However, it is difficult to create one measure that would be completely reliable in all situations. For example, parents are often insensitive to children's emotional difficulties (Angold et al., 1995), and children may also show different communication patterns in different situations (Achenbach et al., 1987). Thus, the parent who most often brings a child to kindergarten may be better acquainted with his or her behavior outside the home than a parent who sees the child less often in such situations. It is therefore possible that the low assessors' agreement (lower interrater reliability) does not necessarily reflect the low reliability of the screening measure itself, but shows how much the estimated characteristics depend on the situational variations that should certainly be taken into account.

Practical implications

Some research shows that without regular screening, only 30% to 50% of children with developmental disabilities are diagnosed before entering kindergarten (Glascoe, 2000), and that regular screening increases this likelihood and referral for further evaluation (Vitrikas et al., 2017). There is an increase in the frequency of early diagnosis and the use of standardized screening measures in countries around the world (e.g. Radecki et al., 2009; Boyle et al., 2011), and Croatia certainly follows this trend. The results of this research show that not all disagreement necessarily points to poor characteristics of measures or inaccuracies in assessments, but may in fact reflect different situations and factors that different assessors take into account when making the assessment. Further research of available screening measures, but also specific situations and differences in the perception of individual child behaviors by different assessors, would make it

possible to identify details that lead to the neglect of individual symptoms or to the neglect of individual differences in the development of an individual child.

Developmental assessment and identification of children in whom a particular trait is not at an age-appropriate level signals the need for further estimations and monitoring of that child (Ortiz-Leon et al., 2018). Research shows that some areas of early development, such as fine and gross motor skills, are significantly associated with cognitive and communication development (Houwen et al., 2016) and with speech development, respectively. Such research emphasizes the importance of early intervention and identification of areas that need to be further encouraged and developed in children.

In this study, odds ratios, interclass correlation coefficients and Kappa coefficients were used to test the assessment agreement for different assessors and the agreement of classification of children into three groups (typical development, mild and severe delays). The obtained results suggest that more information and training of some assessors is necessary so that they can more accurately assess the developmental changes and behaviors that affect them. The most attention in the training of assessors should be given to those parts of the assessment that have proved to be the least accurate and where there are most disagreements between assessors (communication and personal-social area).

Awareness of differences in children's behavior in different environments, as well as behaviors that are consistent with the child's age and those that indicate possible risks, will increase the reliability and accuracy of screening, regardless of the screening method used. Of course, the reliability of assessments is increased by the use of validated screening measures (on which more work is needed) and by constant validation of sensitivity and specificity of available measures for screening.

Conclusion

Early screening of children, which can reliably and accurately identify discrepancies in specific developmental areas, is extremely important for adequate early intervention support. Differences in the assessments of different assessors and non-specific assessments may be the result of poor reliable screening measures, but also a reflection of specific differences in the information available to individual assessors. The agreement between the assessments of parents and preschool teachers for different development areas obtained in this research is moderate, and for some subscales even very low. This is partly a reflection of the psychometric characteristics of the measurement instrument, but also of different environments and reference groups that parents or preschool teachers use for comparing the child's developmental delays. For example, some parents do not have other children to make comparisons between their development or they do not possess sufficient knowledge about children's development, while others could have more children and maybe greater knowledge about typical children's development than preschool teachers (i.e. parents who are also pediatricians or psychologists). The

results suggest the importance of combining assessments from different sources and the importance of assessing different aspects of development (given that one separate developmental area can be an unreliable source of information about potential delay). It is important to continue testing assessor agreement and assessment reliability for developmental delays on larger samples. Furthermore, it is necessary to consider the demographic characteristics of assessors, and to combine assessments from family members and different professionals (preschool teachers, pediatricians, psychologists, rehabilitators, speech therapists, etc.), in order to obtain sensitive and reliable instruments for screening and monitoring the development of early and preschool age children.

Reference

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232. https://doi.org/10.1037/0033-2909.101.2.213

Ali, S. S., Balaji, P. A., Dhaded, S. M., & Goudar, S. S. (2011). Assessment of growth and global developmental delay: A study among young children in a rural community of India. *International Multidisciplinary Research Journal: Neurophysiology, 1*(7), 31-34.

Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research, 5*(4), 237-249.

Boban, Dž. (2016). Epidemiološki podaci o potencijalnim čimbenicima rizika za razvojne teškoće za područje Grada Dubrovnika [Epidemiological data on the potential risk factors for developmental difficulties in the City of Dubrovnik's area]. In *Neurorazvojni poremećaji i rana intervencija*, *zbornik pozvanih predavanja sa simpozija* (pp. 14). Udruga Dubrovnik zdravi grad i Zavod za javno zdravstvo Dubrovačko-neretvanske županije.

Bodnarchuk, J. L., & Eaton, W. O. (2004). Can parent reports be trusted? Validity of daily checklists of gross motor milestone attainment. *Journal of Applied Developmental Psychology, 25*(4), 481-490. https://doi.org/10.1016/j.appdev.2004.06.005

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., Visser, S., & Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics*, *127*(6), 1034-1042. https://doi.org/10.1542/peds.2010-2989

Choo, Y. Y., Agarwal, P., How, C. H., & Yeleswarapu, S. P. (2019). Developmental delay: Identification and management at primary care level. *Singapore Medical Journal, 60*(3), 119-123. https://doi.org/10.11622/smedj.2019025

Correia, L. L., Rocha, H. A. L., Sudfeld, C. R., Rocha, S. G. M. O., Leite, Á. J. M., Campos, J. S., & Silva, A. C. (2019). Prevalence and socioeconomic determinants of development delay among children in Ceará, Brazil: A population-based study. *PLoS ONE*, *14*(11), e0215343. https://doi.org/10.1371/journal.pone.0215343

Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee, & Medical Home Initiatives for Children With Special Needs Project Advisory Committee (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics, 118*(1), 405-420. https://doi.org/10.1542/peds.2006-1231

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). Wiley.

Glascoe, F. P. (2000). Early detection of developmental and behavioral problems. *Pediatrics in Review, 21*(8), 272-280. https://doi.org/10.1542/pir.21-8-272

Guralnick, M. J. (2005). Early intervention for children with intellectual disabilities: Current knowledge and future prospects. *Journal of Applied Research in Intellectual Disabilities, 18*(4), 313-324. https://doi.org/10.1111/j.1468-3148.2005.00270.x

Houwen, S., Visser, L., van der Putten, A., & Vlaskamp, C. (2016). The interrelationships between motor, cognitive, and language development in children with and without intellectual and developmental disabilities. *Research in Developmental Disabilities, 53*, 19-31. https://doi.org/10.1016/j.ridd.2016.01.012

Hsue, B. J., Chen, Y. J., & Wang, Y. E. (2014). The intra- and inter-rater reliability of component analysis of rise from supine in the children with typical development and developmental delay. *Research in Developmental Disabilities, 35*(1), 162-170. https://doi.org/10.1016/j.ridd.2013.10.011

Kerstjens, J. M., Bos, A. F., ten Vergert, E. M., de Meer, G., Butcher, P. R., & Reijneveld, S. A. (2009). Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. *Early Human Development, 85*(7), 443-447. https://doi.org/10.1016/j.earlhumdev.2009.03.001

Ljubešić, M. (2016). Razvojni sistemski model u ranoj intervenciji [Developmental systematic model in early intervention]. In *Neurorazvojni poremećaji i rana intervencija*, *zbornik pozvanih predavanja sa simpozija* (pp. 10). Udruga Dubrovnik zdravi grad i Zavod za javno zdravstvo Dubrovačko-neretvanske županije.

Milić Babić, M., Franc, I., & Leutar, Z. (2013). Iskustva s ranom intervencijom roditelja djece s teškoćama u u razvoju [Experiences with early intervention of parents of children with developmental disabilities]. *Ljetopis Socijalnog Rada, 20(3),* 453-480.

Neale, M., & Cardon, L. (1992). *Methodology for genetic studies of twins and families*. Kluwer Academic Publishers. https://doi.org/10.1007/978-94-015-8018-2

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory.* McGraw-Hill.

Ortiz-León, O., Granados-Rojas, A., Cavazos-Olivo, J., De Benito-Avendaño, L., Hugo Obregón-García, V., Duran-Avendaño, X. M., Ramírez-Pérez, R. V., Cárdenas-Medina, J. A., Vargas-Terréz, B. E., Luna-Guevara, G., Rodríguez-Machain, A. C., & Jaimes-Medrano, A. L. (2018). Internal and inter-rater reliability of the ASQ-3 in Mexican preschoolers. *Salud Mental, 41*(2), 65-72. https://doi.org/10.17711/sm.0185-3325.2018.011

Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmology, 56*(1), 45-50. https://doi.org/10.4103/0301-4738.37595

Radecki, L., Sand-Loud, N., O'Connor, K. G., Sharp, S., & Olson, L. M. (2011). Trends in the use of standardized tools for developmental screening in early childhood: 2002-2009. *Pediatrics, 128*(1), 14-19. https://doi.org/10.1542/peds.2010-2180

Ratner, N. B., & Silverman, S. (2000). Parental perceptions of children's communicative development at stuttering onset. *Journal of Speech, Language, and Hearing Research: JSLHR, 43*(5), 1252-1263. https://doi.org/10.1044/jslhr.4305.1252

Squires, J., Nickel, R. E., & Eisert, D. (1996). Early detection of developmental problems: Strategies for monitoring young children in the practice setting. *Journal of Developmental and Behavioral Pediatrics: JDBP, 17*(6), 420-427. https://doi.org/10.1097/00004703-199612000-00008

Squires, J., Twombly, E., Bricker, D., & Potter, L. (2009). *Ages and Stages Questionnaires: Third Edition*. Paul H. Brookes Publishing.

Van der Valk, J. C., van den Oord, E. J. C. G., Verhulst, F. C., & Boomsma, D. I. (2001). Using parental ratings to study the etiology of 3-year-old twins' problem behaviors: Different views or rater bias? *Journal of Child Psychology and Psychiatry, 42*(7), 921-931. https://doi.org/10.1111/1469-7610.00788

Velki, T., & Romstein, K. (in press). Validation and standardization of Croatian version of ASQ-3 for children aged 3 to 5 years. *Paediatira Croatica*.

Vitrikas, K., Savard, D., & Bucaj, M. (2017). Developmental delay: When and how to screen. *American Family Physician, 96*(1), 36–43.

World Health Organization (2012). *Early childhood development and disability: A discussion paper*. World Health Organization.

Zakon o socijalnoj skrbi. *Narodne novine* 157/2013, 152/2014, 99/2015, 52/2016, 16/2017, 130/2017.

Tena Velki
Faculty of Education
Josip Juraj Strossmayer University of Osijek
Cara Hadrijana 10, 31000 Osijek, Croatia
tena.velki@gmail.com

Ksenija Romstein
Faculty of Education
Josip Juraj Strossmayer University of Osijek
Cara Hadrijana 10, 31000 Osijek, Croatia
kromstein@foozos.hr

Valentina Ružić
Heinzelova 84, 10000 Zagreb, Croatia
vruzic12@gmail.com

Slaganje procjena roditelja i odgojitelja za različita područja razvoja predškolaca

Sažetak

Kada promatramo djecu predškolske dobi te potencijalna razvojna odstupanja, od ključne važnosti su nam osobe koje dobro poznaju dijete (obično roditelji/skrbnici i odgojitelji) i provode značajan dio vremena s njim, jer se odluka o potrebi za intervencijom ili tretmanom najčešće donosi na temelju njihovih procjena djeteta. Cilj je istraživanja[1] *ispitati slaganje procjena roditelja i odgojitelja za različite aspekte razvoja djece predškolske dobi. U istraživanju je sudjelovalo 60 parova procjenjivača (roditelj-odgojitelj) koji su procjenjivali djecu u dobi od 30 mjeseci do 54 mjeseca (42,6 % procjene muške djece te 57,4 % ženske djece). I roditelji i odgojitelji popunili su upitnik „Dob i razvojne razine-3" (Ages & Stages-3; Squires i sur., 2009) koji je za potrebe istraživačkoga projekta preveden na hrvatski jezik. Slaganja procjenjivača provjerena su za 5 područja dječjega razvoja: komunikacija, gruba motorika, fina motorika, rješavanje problema, osobno-društveno te opću zabrinutost. Intraklasni koeficijenti korelacije pokazali su slabe do umjerene povezanosti slaganja procjenjivača (ICC = 0,36–0,84) dok je pouzdanost slaganja procjenjivača mjerena Kappa koeficijentom pokazala blago do značajno slaganje (κ = 0,31–0,61) ovisno o razvojnom području. „Dob i razvojne razine-3" nije se pokazao u potpunosti pouzdan mjerni instrument pri čemu je nužno za mjerenje razvojnih odstupanja koristiti više procjenjivača kao i dodatne mjerne instrumente.*

Ključne riječi: *Ages & Stages-3; pouzdanost slaganja; razvoj predškolaca; razvojna odstupanja; slaganje procjenjivača.*

Uvod

Prema Svjetskoj zdravstvenoj organizaciji (World Health Organisation, 2012) neurorizično dijete je ono dijete koje je tijekom trudnoće, poroda ili u novorođenačkoj dobi bilo izloženo rizičnim čimbenicima. WHO i Svjetska banka navode da se procjenjuje kako više od jednoga bilijuna ljudi, tj. približno 15 % svjetske populacije živi s nekim

[1] Podatci prikupljeni u sklopu projekta: Promjena paradigme poučavanja u ranom djetinjstvu: konstruktivizam i razvojni pristup vs. biheviorističke strategije (interni znanstveno-istraživački projekt Sveučilišta J. J. Strossmayera u Osijeku, INGI-2015-29).

oblikom teškoće. Međutim, trenutačno ne postoje sveobuhvatni podatci o tome koliko djece ima neku teškoću jer se procjene učestalosti razlikuju ovisno o definiciji i načinu dijagnosticiranja u pojedinim zemljama (World Health Organisation, 2012).

Danas se smatra da 10-15 % živorođene djece pripada skupini neurorizične djece (Boban, 2016). Sve više se u društvu prepoznaje važnost prvih nekoliko godina života u kojima se razvijaju motoričke i kognitivne sposobnosti djeteta, nakon kojih (ili usporedo s njima) slijedi djetetov emocionalni i socijalni razvoj. Suvremena znanost pruža jake dokaze da rano iskustvo oblikuje temelje za cjeloživotno učenje, ponašanje i mentalno zdravlje i da je razvoj sposobnosti postupan proces u kojem jednostavnije vještine pomažu u oblikovanju složenijih (Ljubešić, 2016). Prepoznavanje i praćenje djece ključno je za rano otkrivanje neurorazvojnih odstupanja i uključivanje u sustav rane intervencije. Naime, na učenje i razvoj u prvim godinama života može se djelovati intervencijama u različitim razvojnim područjima (Guralnick, 2005), uz dugoročne pozitivne učinke.

Važnost rane intervencije sve više prepoznaju stručnjaci različitih profila (liječnici, psiholozi, edukacijski rehabilitatori, socijalni radnici), ali i predstavnici vlasti i donositelji zakona, kako u svijetu tako i u Hrvatskoj. U zakonski sustav Republike Hrvatske rana intervencija uvrštena je prvi put 2011. godine (Milić Babić i sur., 2013), kao jedan od oblika socijalne podrške obiteljima. Zakon o socijalnoj skrbi (Narodne novine 157, 2013, čl. 84, st. 1) navodi da je rana intervencija „socijalna usluga koja obuhvaća stručnu poticajnu pomoć djeci i stručnu i savjetodavnu pomoć njihovim roditeljima, uključujući i druge članove obitelji te udomitelja za djecu, kod nekog utvrđenog razvojnog rizika ili razvojne teškoće djeteta.“. Zakonodavni okvir u Hrvatskoj daje mogućnost podrške djetetu i obitelji putem rane intervencije do 3., odnosno do 7. godine djetetova života (Zakon o socijalnoj skrbi, 2013, čl. 84, st. 3). Utvrđivanje zakonskih okvira temeljna je pretpostavka sustavnom provođenju rane intervencije kojoj je glavni cilj pružanje podrške obiteljima u procesu podizanja djece rođene s neurorizicima.

Unazad dvadesetak godina, u kontekstu rane intervencije, pojavljuje se probir djece na razvojne teškoće kao jedna od pretpostavki pravovremenoga poticanja dječjega razvoja. Cilj probira je rana identifikacija djece s razvojnim odstupanjima i uključivanje u sustav rane intervencije. Kako bi istaknula važnost probira na razvojna odstupanja u ranom djetinjstvu, Američka pedijatrijska akademija (American Academy of Pediatrics, AAP) naglašava da je „rana identifikacija razvojnih odstupanja ključna za dobrobit djece i njihovih obitelji“ (Council on Children With Disabilities, 2006). Ljubešić (2016) navodi da djeca s neurorazvojnim rizicima ili teškoćama, koja u ranoj dobi ne dobiju primjeren oblik intervencije, zapravo budu dvostruko zakinuta. S jedne strane ona se, već pod djelovanjem svoje teškoće spontano „povlače“ iz okoline, zbog čega su izloženi manjoj količini stimulacije, a uz to i njihovi roditelji budu pod stresom zbog obaveza koje se nameću vezano uz djetetova ograničenja, što dovodi do toga da su manje responzivni u interakcijama s djetetom nego bi to bili u drugačijim okolnostima.

Rana identifikacija zapravo je kontinuirani proces prikupljanja informacija o djetetu iz različitih izvora, uključujući izravna opažanja roditelja/skrbnika i stručnjaka (Squires i sur., 1996). Kad govorimo o sustavu primarne zdravstvene zaštite, razvojna odstupanja obično se primjećuju ili tijekom redovnoga pregleda liječnika/pedijatra ili pak potaknuto roditeljskom brigom vezanom uz neki aspekt djetetova razvoja, a ponekad tu zabrinutost iskazuje neka treća osoba (npr. odgojitelj) (Choo i sur., 2019). S obzirom na to da se razvojni rizici povećavaju s djetetovom kronološkom dobi, preporučuje se barem jedanput godišnje raditi pregled djece koja su suspektna na odstupanje kako bi se pratio njihov razvoj. Također, preporučuje se periodički razvojni probir za svu djecu.

S djetetovim odrastanjem povećava se i potreba uključivanja različitih stručnjaka u procjenu razvoja, kako bi svaki od njih mogao iz svoje perspektive procijeniti aspekte razvoja i eventualna odstupanja u djetetovu razvoju. Probir se može vršiti i u sklopu rada pedijatrijskih službi i u sklopu rada ostalih stručnjaka (Council on Children With Disabilities i sur., 2006) kako bi se u što većoj mjeri osigurao probir na teškoće i uključivanje u ranu intervenciju. Međutim, u identificiranju potencijalnih odstupanja važno je prikupljanje podataka ne samo od liječnika i stručnjaka, već i od djetetu najbližih osoba (roditelja/skrbnika, obitelji, odgojitelja) jer su oni u mogućnosti opažati dijete u drugačijim okolnostima i tako pružiti informacije iz različitih kutova.

Potrebno je naglasiti da se probirom za razvojna odstupanja ne dolazi do dijagnoze već se identificiraju područja u kojima dijete manifestira sniženo funkcioniranje u odnosu na očekivano, tj. u odnosu na kronološku dob, na temelju čega se može razmotriti potreba za uključivanjem obitelji u sustav rane intervencije i savjetovanja. Potencijalni cilj ranoga probira je dvojak: obuhvat što većega broja djece, uz što manju „propusnost" probira u smislu neprepoznavanja djece koja imaju razvojna odstupanja. Drugi cilj je smanjiti broj tzv. lažno pozitivnih slučajeva, tj. smanjenje broja djece koja se prepoznaju kao rizična i koja se upućuju na daljnje obrade, a da kod njih zapravo razvojno odstupanje ne postoji. Pri tome je potrebno u odnos staviti dvije mjere – osjetljivost i specifičnost mjere probira. Osjetljivost govori o tome koliko je metoda probira „dobra" u prepoznavanju osoba s određenom bolesti/teškoćom, dok specifičnost govori o tome koliko je „dobra" u prepoznavanju onih koji nemaju bolest/teškoću, tj. eliminiranju potrebe za daljnjom procjenom kod tih pojedinaca (Parikh, 2008). Navedene dvije mjere u recipročnom su odnosu pa odluku o tome koje vrijednosti odabrati treba donijeti na temelju procijenjenih prednosti. Povećanjem osjetljivosti mjere probira povećava se broj djece koja se identificiraju kao potencijalno rizična i koja se upućuju na detaljniju procjenu, tj. smanjuje se mogućnost „propuštanja" takvoga djeteta. Time se povećava vjerojatnost da ćemo prepoznati djecu s potencijalnim teškoćama, ali se ujedno povećava i mogućnost identificiranja djece koja zapravo nemaju teškoću. S druge strane, povećanje specifičnosti osigurava da ne identificiramo ni jedno dijete kao dijete s potencijalnom teškoćom ako je ono zapravo nema, čime se smanjuju potencijalni vremenski i materijalni troškovi procjene. Osjetljivost se još naziva procjena stvarno

pozitivnih slučajeva dok se specifičnost naziva procjena stvarno negativnih slučajeva (Nunnally i Bernstein, 1994). Potpuno točno previđanje imalo bi 100 % osjetljivost i 100 % specifičnost, ali ni jedna mjera nema takve karakteristike jer se uz procjenu uvijek veže određeni stupanj pogreške.

Upotreba standardiziranih načina procjene pomaže u postizanju cilja što točnije identifikacije rizične djece, ali i kod takvih mjera postoje određeni problemi. Jedan od njih je pitanje pouzdanosti procjena općenito te problem slaganja procjena između različitih procjenjivača. Naime, često je slučaj da jedan procjenjivač pokazuje određenu razinu zabrinutosti vezano uz djetetov razvoj, dok u isto vrijeme ostali odrasli koji su u kontaktu s djetetom ne iskazuju zabrinutost. Uzrok takvih neslaganja može biti u različitosti situacija u kojima pojedine osobe u djetetovom životu (potencijalni opažači) vide dijete, ali i u doživljaju djeteta – dijete se osjeća različito sigurno, zaštićeno, slobodno i opušteno u različitim situacijama, pa se u skladu s time njegovo ponašanje i razlikuje. Pri tome zapravo nije nužno da jedno opažanje bude pogrešno, već treba uzeti situacijske faktore u obzir i kombinirati informacije iz različitih izvora. Zbog toga je važno sustavno pratiti, istraživati i uspoređivati procjene djeteta prikupljene iz različitih izvora, kako bi probir na razvojna odstupanja bio što učinkovitiji. Dosadašnja istraživanja pokazuju kako su procjene roditelja prilično pouzdane (Bodnarchuk i Eaton, 2004; Ratner i Silverman, 2000; Guralnick, 2005), ali i da se mogu prilično razlikovati ovisno o tome koja se sposobnost procjenjuje, tj. u kojem okruženju ona dolazi do izražaja. S druge strane, postoje određene razlike u procjenama pojedinoga roditelja – majke i oca (Achenbach i sur., 1987) koje se objašnjavaju time što svaki od njih opaža dijete u različitim okolnostima i može biti usmjeren na različite aspekte razvoja. Istraživanje provedeno na blizancima (van der Valk i sur., 2000) pokazalo je da neslaganje između procjena roditelja odražava jedinstvenost podataka koje daje samo jedan roditelj, a ne nepouzdanost ili pristranost jednog od ocjenjivača. U metaanalizama procjena dobivenih u brojnim istraživanjima utvrđeno je da je prosječna korelacija između procjena dvaju roditelja 0,59, dok je korelacija između parova odgajatelja 0,64, a između roditelja i odgajatelja 0,27 (Achenbach i sur., 1987).

S obzirom na to da velik broj mjera koje se koriste za ispitivanje razvoja djece predviđaju i određene kritične rezultate (rezultate koji upućuju na to da je određeno dijete suspektno na odstupanje u razvoju), moguće je da određeni broj djece bude klasificirano kao suspektno na temelju procjena jednoga opažača, a ne i na temelju procjena drugog (drugih) opažača. Drugim riječima, nužno je imati na umu da se rezultat probira može razlikovati ovisno o tome tko procjenjuje dijete. Stoga je važno analizirati povezanost klasifikacija ovisno o procjenjivaču. Jedan pristup analizi točnosti klasifikacije jest izračunavanje omjera izgleda relativnih rizika (Fleiss, 1981), koji se koristi u epidemiološkim istraživanjima. Navedeni omjer označava vjerojatnost da će u određenoj skupini ljudi koji su procjenjivani, oni identificirani kao rizični za određeni poremećaj zaista i imati taj poremećaj, u odnosu na vjerojatnost prisutnosti toga stanja u skupini osoba koje nisu identificirane kao rizični. Usporedba između omjera izgleda za one koji imaju i one koji nemaju taj čimbenik rizika (određeni kritični rezultat)

izražava se kao omjer vjerojatnosti za prisutnost nekog ishoda ako je taj čimbenik rizika prisutan i vjerojatnosti za prisutnost toga ishoda kad je taj čimbenik rizika odsutan.

Cilj je ovoga istraživanja ispitati slaganje procjena roditelja i odgojitelja za različita područja razvoja djece predškolske dobi. Cohenovim kappa koeficijentom mjerena je pouzdanost slaganja među procjenjivačima, odnosno slaganje između roditelja i odgojitelja. Kappa koeficijent predstavlja mjeru u kojoj podatci prikupljeni u istraživanju točno reprezentiraju mjerene varijable. U svrhu ovoga rada, kappa koeficijenti predstavljaju pouzdanost odluke od tome da je dijete rizično za razvojno kašnjenje, ovisno o tome tko ga procjenjuje. Usporedit će se slaganja na 5 područja dječjega razvoja: komunikacija, gruba motorika, fina motorika, rješavanje problema i osobno-društveno područje te slaganje u općoj zabrinutosti za djetetov razvoj. Dodatni cilj istraživanja je usporediti klasifikaciju djece prema zadanom kriteriju (suspektna i rizična djeca) između roditelja i odgojitelja kao procjenjivača. S obzirom na to da se procjene koriste za donošenje daljnjih odluka vezanih uz intervencije i tretmane djeteta važno je ispitati na koji se način različiti procjenjivači slažu u procjenama radi li se o potencijalno neurorizičnom djetetu.

Metoda

Sudionici

U istraživanju je sudjelovalo 60 roditelja (95 % majki) djece u dobi od 30 mjeseci do 54 mjeseca, te 60 odgojitelja. Analiza slaganja procjena napravljena za 60 parova procjena (roditelj-odgojitelj), od čega su 41,7 % procjena muške djece, a 58,3% ženske djece. Sva procjenjivana djeca pohađala su vrtić, a njihovi roditelji dali su pristanak na za procjenu. Roditelji su u prosjeku imali M = 29,51 (SD = 6,31) godina dok je prosječni radni staž odgojiteljica (100 % žene) bio M = 15,44 (SD = 10,68) godina. Detalji su prikazani u Tablici 1.

Tablica 1.

Instrument

U istraživanju je primijenjen upitnik *Dob i razvojne razine-3* (*Ages and Stages-3*, ASQ; Squires i sur., 2009), koji je za potrebe istraživačkoga projekta preveden na hrvatski jezik. Suglasnost autora instrumenta dobivena je za istraživačke svrhe. ASQ sastoji se od niza čestica koje mjere različite aspekte razvoja. U prvom dijelu upitnika roditelj/skrbnik/odgojitelj daje osnovne demografske podatke (datum rođenja i spol djeteta te odnos osobe s djetetom). Drugi dio upitnika sastoji se od 30 pitanja podijeljenih u 5 podskala koje ispituju različita razvojna područja: komunikacija (k = 6), gruba motorika (k = 6), fina motorika (k = 6), rješavanje problema (k = 6) i osobno-društveno područje (k = 6). Sudionici na ova pitanja odgovaraju jednom od 3 moguće kategorije „da“, „ponekad“ i „još ne“. Za svaku skalu izračunava se ukupan zbroj bodova (0 do 60). Treći dio upitnika sastoji se od 10 pitanja koja ispituju zabrinutost roditelja/skrbnika/odgojitelja po različitim područjima. Linearnom kombinacijom ovih 10 pitanja dobiva

se opća zabrinutost pri čemu se od zbroja prvih 5 odgovora (svaki pozitivan odgovor donosi 1 bod) oduzme zbroj drugih 5 pitanja (teoretski raspon odgovora se kreće od - 5 do 5 pri čemu veći broj označava manji stupanj zabrinutosti). Upitnik je podijeljen u kronološke dobi od 2 do 60 mjeseci. Za potrebe ovoga istraživanja koristile su se verzije: 36 mjeseci (od 34 mjeseca i 16 dana do 38 mjeseci i 30 dana), 42 mjeseca (od 39 mjeseci i 0 dana do 44 mjeseca i 30 dana), 48 mjeseci (od 45 mjeseci i 0 dana do 50 mjeseci i 30 dana) i 54 mjeseca (od 51 mjesec i 0 dana do 56 mjeseci i 30 dana).

U Tablici 2 prikazani su koeficijenti unutarnje pouzdanosti tipa Cronbachov α za sva područja razvoja i to zajedno za roditelje i odgojitelje te za svaku skupinu posebno.

Tablica 2.

Rezultati pokazuju nisku pouzdanost podskale Osobno-društveno područje te relativno nisku pouzdanost podskala Komunikacija i Gruba motorika. Zadovoljavajuća pouzdanost dobivena je za podskale Rješavanje problema i Fina motorika, i to za obje skupine procjenjivača, roditelje i odgojitelje ($\alpha > 0{,}60$), iako ni ta pouzdanost nije visoka.

Postupak

Istraživanje je provedeno u 5 županija (Osječko-baranjskoj, Brodsko-posavskoj, Varaždinskoj, Virovitičko-podravskoj i Vukovarsko-srijemskoj). Roditelji su u dječjem vrtiću dobili primjerak upitnika koji su ispunili kod kuće i zatim vratili odgojitelju. Odgojitelji su također dobili isti upitnik za popunjavanje. Oko jedne trećine roditelja vratilo je popunjen upitnik. Istraživanje je i za roditelje i odgojitelje bilo u potpunosti dobrovoljno. Vrijeme potrebno za popunjavanje procjene bilo je oko 20 minuta. Na kraju je svim roditeljima bila ponuđena opcija davanja povratne informacije o procjeni koju su načinili, a za djecu suspektnu na neki razvojni poremećaj ponuđena je opcija davanja povratne informacije vrtićima u svrhu primjerenoga pedagoškog rada s djetetom (na što su dobrovoljno pristali i roditelji i odgojitelji).

Rezultati i rasprava

U obradu su uzeti su samo oni upitnici čije su procjene djeteta dobivene i od stane roditelja i odgojitelja. Formirani su rezultati na svim ljestvicama (podskalama), a deskriptivna statistika za te ljestvice prikazana je u Tablici 3. Zasebno su prikazane vrijednosti dobivene na procjenama roditelja (označene slovom R) i vrijednosti procjena odgojitelja (označene slovom O).

Tablica 3.

Kao što je vidljivo, indeksi asimetrije (Skewness i Kurtosis) ni u jednoj skupini ne prelaze vrijednosti +/- 4 pa je u daljnjim analizama primijenjena parametrijska statistika. Teoretski raspon odgovara na podskalama kretao se od 0 do 60, pri čemu veći broj ukazuje na viši stupanj razvoja određenoga razvojnog područja. U Tablici 3 vidljivo je da ni na jednoj podskali nije postignut minimalan broj bodova (0 bodova) dok je na svim podskalama postignut maksimalan broj bodova. Takvi rezultati govore u prilog

o nešto slabijoj osjetljivosti mjerenoga instrumenta. Također i indeksi odstupanja od normalne distribucije, posebice za roditeljske procjene, pokazuju više vrijednosti, blizu graničnih (posebice za podskale Komunikacija, Gruba motorika i Osobno-društveno), što također govori u prilog o njihovoj slabijoj primjenjivosti u praksi.

Slaganje roditelja i odgojitelja o razvojnim procjenama djeteta

U skladu ciljem istraživanja ispitano je slaganje procjenjivača roditelj-odgojitelj na 5 područja dječjeg razvoja: komunikacija, gruba motorika, fina motorika, rješavanje problema i osobno-društveno područje te opća zabrinutost za djetetov razvoj. U Tablici 4 prikazani su intraklasni koeficijenti korelacije kao mjera slaganja procjena iz dvaju izvora.

Tablica 4.

Sve dobivene podskale pokazuju statistički značajno slaganje procjenjivača pri čemu osobno-društveno područje i opća zabrinutost za djetetov razvoj pokazuju slabo slaganje (ICC < 0,5), područje grube motorike umjereno (ICC = 0,50 - 0,75) dok je za područje komunikacije, fine motorike i rješavanje problema dobiveno dobro slaganje (ICC = 0,75 - 0,90). S obzirom na dobivene rezultate možemo zaključiti kako se roditelji i odgojitelji većinom slažu u procjenama kada su u pitanju tri područja djetetova razvoja (komunikacija, fina motorika i rješavanje problema) procjenjujući stupanj djetetova razvoja podjednako za ova područja. Najniže je slaganje dobiveno za osobno-društveno područje što je i očekivano s obzirom na okolinu u kojoj se dijete nalazi kada se provodi procjena. Roditelji procjenjuju dijete unutar obitelji gdje se zapravo društvena komponenta svodi na nekoliko užih članova obitelji s kojima je dijete često u interakciji (npr. zajedničke aktivnosti i igra) dok odgojitelji istu komponentu procjenjuju u okruženju djeteta s vršnjacima gdje drugi aspekti odnosa dolaze do izražaja (npr. snalaženje u grupi vršnjaka, započinjanje i održavanje interakcija s vršnjacima, dijeljenje igračke i sl.). Drugi potencijalni razlog razlikama u procjenama jest u formalnom obrazovanju odgojitelja u kojemu se socijalnim vještima, posebice u dimenziji samostalnosti i odnosa s drugima, daje velika važnost u sklopu različitih kolegija (npr. Razvojna psihologija) te je roditeljima koji nemaju formalno obrazovanje na području ranoga odgoja i onima koji nemaju više djece u obitelji teže procijeniti primjerena ponašanja u odnosu na djetetovu kronološku dob.

Slaganje roditelja i odgojitelja o klasifikaciji pojedinoga djeteta

S obzirom na to da primijenjeni upitnik sadrži i kritične vrijednosti za odstupanja u ponašanju i to za dvije razine (suspektno i kritično odstupanje) za pet razvojnih područja, provjerena je i pouzdanost slaganja procjenjivača kappa koeficijentom (Tablica 5).

Tablica 5.

Rezultati pokazuju pouzdano slaganje između roditelja i odgojitelja samo za 3 područja razvoja (komunikacija, fina motorika i rješavanje problema) pri čemu je slabo slaganje dobiveno za područje komunikacije, umjereno za područje fine motorike te dobro za područje rješavanja problema. Kada se uzme u obzir odstupanje koje bi značilo da roditelji i/ili odgojitelji trebaju uvesti razvojno primjerene poticajne aktivnosti (za djecu kategoriziranu kao suspektnu za određeno razvojno područje) ili uputiti na daljnju obradu zbog potencijalno ozbiljnijega razvojnoga odstupanja (za djecu kategoriziranu kao kritičnu za određeno razvojno područje), roditelji i odgojitelji puno se manje slaže nego kod opće procjene razvojnih područja (intraklasni koeficijenti korelacije za sva razvojna područja su bili statistički značajni, Tablica 3).

Osim različite okoline u kojoj se promatra dijete za koje se radi procjena (dječji vrtić i obitelj), različiti procjenjivači mogu imati i različit stupanj kompetentnosti za obavljanje procjene djeteta. Naime, svojim obrazovanjem odgojitelji posjeduju određenu razinu znanja o očekivanom razvoju djeteta predškolske dobi kao i o potencijalnim mogućim odstupanjima što može dovesti do razlika u procjenama roditelja i odgojitelja posebice za roditelje koji nemaju srodno obrazovanje. Međutim, i samo poznavanje djeteta i njegova napretka, može utjecati na procjenu, posebice za odgojitelje koji nisu dugo u odgojnoj skupini u kojoj procjenjuju specifično dijete, a za koje roditelj, bez obzira na obrazovanje, možda može dati bolju procjenu.

S obzirom na dinamičnost razvoja djece ove dobi, korisno bi bilo usporediti kappa koeficijente kroz dob. Nažalost, s obzirom na malen broj parova procjenjivača u pojedinoj od četiri dobne kategorije, takva usporedba na podatcima prikupljenim u ovom istraživanju nije moguća. Naime, u nekim se kategorijama, zbog maloga broja procjenjivača, procjene sve djece određene dobi nalaze u istoj kategoriji (npr. odgojitelj svu djecu te dobi procjenjuje kao urednu) pa kappa koeficijent prema dobi nije dovoljno dobar pokazatelj slaganja procjena. Ipak, uočeni su određeni trendovi u promjenama s dobi koji pokazuju da se kappa koeficijenti razlikuju ovisno o dobi u pojedinim podskalama, vjerojatno zato što se ponašanja koja su uključena u pojedinu procjenu mijenjaju s dobi i postaju različito uočljiva u pojedinim situacijama.

Originalni američki upitnik ASQ-3 sadrži granične kriterije prema kojima se razvojna odstupanja djece dijele u tri skupine: djeca urednoga razvoja (bez razvojnih odstupanja), suspektna djeca (djeca koja pokazuju blaža odstupanja od prosjeka populacije na nekom od 5 razvojnih područja, preporuča se nadzor i poticajne aktivnosti za ovu skupinu djece) i kritična djeca (djeca koja pokazuju značajnija odstupanja od prosjeka populacije na nekom od 5 razvojnih područja, preporuča se profesionalna pomoć, odnosno upućivanje djeteta na daljnju obradu). S obzirom na to da je u Republici Hrvatskoj u tijeku standardizacija ASQ-3 te su preliminarni rezultati u postupku objave (Velki i Romstein, 2020), u tablicama 6 i 7 prikazana su odstupanja djece na temelju američkih graničnih vrijednosti.

Napravljena je usporedba omjera izgleda (OI) na način koji predlaže Fleiss (1981), za one koji imaju i one koji nemaju čimbenik rizika (određeni kritični rezultat) prema

procjeni jednoga od procjenjivača. Izračunati omjeri izgleda pokazuju vjerojatnost da djeca koja su klasificirana kao odstupajuća (suspektna ili kritična) na temelju procjena roditelja budu jednako klasificirana na temelju procjena odgojitelja. Usporedno su izračunati omjeri izgleda da djeca koja su klasificirana kao odstupajuća (suspektna ili kritična) na temelju procjena odgojitelja budu jednako klasificirana na temelju procjena roditelja. Rezultati su prikazani u Tablici 6.

Tablica 6.

Razlike u proporcijama OI suspektne djece kod pojedinoga procjenjivača testirane su *t*-testom za proporcije. Rezultati su pokazali da je razlika značajna za aspekt komunikacije ($t = 4{,}14$; $p < 0{,}01$) i rješavanje problema ($t = 3{,}16$; $p < 0{,}01$) dok nema značajne razlike u preostala tri aspekta razvoja. Vidljivo je da su omjeri rizika različiti za pojedine aspekte razvoja pri čemu jedino za aspekte komunikacije i rješavanje problema svu djecu koju roditelji procjenjuju suspektnom istom takvom procjenjuju i odgojitelji. Najniži OI dobiven je za procjenu roditelja u osobno-društvenom području (samo jedno od 8 djece identificirane od strane odgojitelja suspektnim smatraju i roditelji).

Također, vidljivo je da postoji najveća razlika između procjene djece suspektnima na ljestvici osobno-društveno područje i ljestvici fine motorike, što ponovno pokazuje da su upravo ta područja razvoja „najteža" za procjenu. Potonje govori u prilog postojanju utjecaja formalnoga obrazovanja odgojitelja na procjenu ovoga područja. Postotci djece svrstane u dvije specifičnije kategorije (suspektni i rizični) na temelju procjena oba procjenjivača, prikazani su u Tablici 7.

Tablica 7.

Kada su u pitanju kritična djeca, roditelji i odgojitelji u većoj se mjeri slažu s procjenom odstupanja (slaganje je potpuno za dob od 36 te 48 mjeseci). Rezultati pokazuju da je slaganje procjenjivača veće kad se radi o svrstavanju u skupinu rizične djece, vjerojatno zato što u kriteriji jasniji i procjena je pod manjim utjecajem situacijskih čimbenika. U slaganju procjena kada se radi o suspektnom djetetu, nalazimo više neslaganja među procjenjivačima jer je ponekad teško procijeniti je li određeno ponašanje djeteta u granicama urednoga razvoja ili pak pokazuje blaže odstupanje od urednoga (prosjeka populacije).

U interpretaciji rezultata potrebno je uzeti u obzir da je u populaciji zapravo mali broj djece koja pokazuju značajna odstupanja, odnosno zadovoljavaju kriterije za postavljanje dijagnoze (prema Boban (2016) procjenjuje se 10 – 15 % neurorizične djece u populaciji). U praksi nailazimo na veći broj djece koja pokazuju manja odstupanja (suspektna djeca prema opisanoj kategorizaciji) odnosno na blaže privremene zastoje u razvoju (Guralnick, 2005). Takvi su podatci razumljivi ako uzmemo u obzir da se razvoj djeteta ne odvija linearnom putanjom, jednako za svu djecu, već je skokovit pri čemu su blaža odstupanja očekivana. Također, u objašnjenju razloga neslaganja u procjenama treba uzeti u obzir da roditelji i odgojitelji imaju različita prethodna iskustva

i znanja u procjenama te napredovanje djeteta uspoređuju s različitim referentnim skupinama. Rezultati pokazuju da je veći broj djece procijenjeno kao suspektno od strane odgojitelja nego od strane roditelja što naglašava važnost uključivanja procjena od strane odgojitelja u svakodnevnu praksu procjene razvoja djece. S obzirom na to da odgojitelji imaju iskustva s opažanjem većega broja djece tijekom svojega rada i imaju priliku uspoređivati ponašanje jednoga djeteta spram cijele grupe vršnjaka (dok roditelji obično imaju iskustva s opažanjem manjega broja djece i to u obiteljskom okruženju), odgojitelji mogu biti važan izvor informacija o djetetovu razvoju. Rezultati ovakvih istraživanja mogu roditeljima, ali i drugim stručnjacima pokazati da procjene odgojitelja ne bi trebale biti zanemarene.

Slični rezultati dobiveni su i u istraživanjima u drugim zemljama (npr. Kerstjens i sur., 2009; Ortiz-Leon i sur., 2018). S obzirom na dobivene rezultate koji upućuju na slabije slaganje u procjenama razvojnih odstupanja roditelja i odgojitelja preporuka je svakako koristiti procjene i jednih i drugih pri utvrđivanju potencijalnih odstupanja u razvoju djeteta. Korištenjem oba izvora procjene kod donošenja odluke o uključivanju djece u tretmane smanjujemo vjerojatnost neprepoznavanja, tj. „propuštanja" određenih informacija koje upućuju na potrebu za intervencijom. S obzirom na važnost rane identifikacije mogućih razvojnih rizika, ovi rezultati naglašavaju važnost kombiniranja procjenjivača, umjesto zanemarivanja jednoga od izvora procjene. Također, opažene razlike među procjenjivačima mogu povećati osjetljivost svih uključenih u život djeteta na to da oni imaju priliku opažati dijete u ograničenim okolnostima i zato mogu propustiti primijetiti ponašanja koja se očituju u različitim okolnostima.

Također, zanimljivo je primijetiti da je manje slaganje procjenjivača dobiveno upravo na podskalama kod kojih je dobivena i manja pouzdanost. Za podskale s nižom unutarnjom pouzdanosti mjere slaganja između procjenjivača također sucniske ili uopće nisu statistički značajne. Zato se za podskale Komunikacija, Gruba motorika, a posebice Osobno-društveno ne preporučuje njihovo zasebno korištenje kao mjere probira na razvojna odstupanja, već kombiniranje tih podskala s drugim procjenama. Kao što je već navedeno, razlozi manje pouzdanosti i manjega slaganja procjenjivača mogu biti nedostatan trening i različito iskustvo procjenjivača, kao što je čest problem u istraživanjima s različitim procjenjivačima (npr. Hsue i sur., 2014), ali i niža pouzdanost i veća pogreška pri mjerenju što je odraz samoga mjernoga instrumenta. Također, moguć je utjecaj kulturoloških čimbenika na procjene odgojitelja i roditelja, o čemu valja voditi računa kod preuzimanja stranih neadaptiranih testova probira. Na primjer, različiti rezultati i ponekad norme dobivene su za uzoroke djece iz različitih država kao što je Brazil (Correia i sur., 2019) i Indija (Ali i sur., 2011). S obzirom na to da nije jasno jesu li dobivene razlike u rezultatima posljedica psihometrijskih karakteristika korištenih skala ili različitih (ranije navedenih) okolnosti procjene takve skale treba koristiti s velikim oprezom.

U skladu s opisanim rezultatima, a u cilju povećanja pouzdanosti procjena djece koje mogu dovesti do ranijega uključenja u potrebne tretmane, nužna su daljnja

istraživanja načina procjene djece kao i edukacija različitih osoba u tim procjenama te prepoznavanje pojedinih prednosti i nedostataka.

Metodološka ograničenja

Provedeno istraživanje ima i određene metodološke nedostatke koje treba razmotriti prilikom interpretacije dobivenih rezultata. U istraživanju je primijenjen preveden američki mjerni instrument (Ages & Stages-3) za koji je postupak validacije i standardizacije za populaciju djece u Republici Hrvatskoj još u tijeku (Velki i Romstein, 2020). S obzirom na to da su za računanje mjera slaganje procjenjivača korišteni granični kriteriji dobiveni na američkoj populaciji djece, pitanje bi li identični podatci bili dobiveni da je prethodno napravljena standardizacija i određeni granični kriteriji za populaciju djece u Republici Hrvatskoj. Nadalje, iako je istraživanje provedeno u 5 županija, dobiveni uzorak nije reprezentativan za cijelu Hrvatsku. Nisu korištene sve postojeće verzije upitnika za sve dobne raspone (2 do 60 mjeseci) te je relativno mali broj parova procjenjivača uključen u istraživanje (posebice za procjene dobi od 36 mjeseci). Buduća istraživanja svakako bi trebala uključivati reprezentativni uzorak, na cijelom dobnom rasponu s većim brojem parova procjenjivača, a poželjno je napraviti i dorade pojedinih podskala s niskom pouzdanošću kao i standardizaciju s graničnim kriterijima za populaciju djece Republike Hrvatske.

Također, u ovom istraživanju nisu uzete u obzir demografske karakteristike procjenjivača, poput dobi i obrazovanja roditelja te socioekonomskoga statusa koje također mogu imati učinak na rano prepoznavanje razvojnih teškoća kod djece. Također, prikupljene su procjene samo jednoga roditelja, a dosadašnja istraživanja naglašavaju da se i roditelji nalaze u različitim situacijama i imaju različite mogućnosti uviđanja teškoća kod djeteta. Neale i Cardon (1992) naglašavaju važnost upravo takve različite percepcije i identificiranja standarda za različite vrste procjene. Pristranost pojedinoga procjenjivača, tj. razlike u njihovim procjenama najčešće se identificiraju kao pogreške mjerenja i pripisuju nepouzdanosti korištenih tehnika procjene i mjera probira, ali teško je osmisliti jednu mjeru koja bi bila potpuno pouzdana u svim situacijama. Na primjer, roditelji su često neosjetljivi na emocionalne teškoće djece (Angold i sur., 1995), a djeca mogu pokazivati i različite komunikacijske obrasce u različitim situacijama (Achenbach i sur., 1987). Tako roditelj koji najčešće dovodi dijete u vrtić može biti bolje upoznat s njegovim ponašanjem izvan kuće od roditelja koji dijete rjeđe vidi u takvim situacijama. Upravo je zato moguće da neslaganja procjenjivača ne odražavaju nužno nepouzdanost same mjere probira već pokazuju koliko su procjenjivane karakteristike ovisne o situacijskim varijacijama koje bi kod procjene odstupanja svakako trebalo uzeti u obzir.

Praktične implikacije

Neka istraživanja pokazuju da bez redovitoga probira samo 30 % do 50 % djece s razvojnim teškoćama bude dijagnosticirano prije polaska u vrtić (Glascoe, 2000) te da redoviti probir povećava tu vjerojatnost i upućivanje na daljnju evaluaciju

(Vitrikas i sur., 2017). Primjećuje se povećanje učestalosti rane dijagnostike i korištenja standardiziranih mjera probira u zemljama širem svijeta (npr. Radecki i sur., 2009; Boyle i sur., 2011), a Hrvatska bi svakako trebala slijediti taj trend. Iz rezultata istraživanja vidljivo je da ne moraju sva neslaganja nužno upućivati na loše karakteristike mjera ili pak netočnost procjena, već zapravo mogu odražavati različite situacije i čimbenike koji različiti procjenjivači uzimaju u obzir prilikom procjene. Daljnjim istraživanjima dostupnih mjera probira, ali i specifičnih situacija i razlika u percepciji pojedinih ponašanja djeteta kod različitih procjenjivača moguće je identificirati detalje koji dovode do zanemarivanja pojedinih simptoma ili do zanemarivanja individualnih razlika u razvoju pojedinoga djeteta.

Razvojna procjena i identifikacija djece kod koje određena osobina nije na razini primjerenoj za dob signalizira potrebu za daljnjom procjenom i praćenjem toga djeteta (Ortiz-Leon i sur., 2018). Istraživanja pokazuju da su neka područja ranoga razvoja, poput fine i grube motorike, značajno povezana s kognitivnim razvojem i razvojem komunikacije (Houwen i sur., 2016) pa s time i razvojem govora. Takva istraživanja dodatno naglašavaju važnost rane intervencije i identificiranja područja koje je potrebno dodatno poticati i razvijati kod djece.

U ovom istraživanju korišteni su omjeri izgleda, interklasni koeficijenti korelacije i kappa koeficijenti za procjenu slaganja procjena različitih procjenjivača i sukladnosti kategorizacije djece u tri skupine (uredni, suspektni i kritičan razvoj). Dobiveni rezultati sugeriraju da je potrebno više informacija i treninga nekih procjenjivača kako bi oni mogli točno procijeniti razvojne promjene i ponašanja koja na njih utječu. Najviše pažnje u obučavanju procjenjivača treba posvetiti upravo onim dijelovima procjene koji su se pokazali najmanje točni i kod kojih postoje najviše neslaganja između procjenjivača (komunikacija i osobno-društveno područje).

Osvještavanjem razlika u ponašanju djece u različitim okruženjima kao i ponašanja koja su u skladu s djetetovom dobi i ona koja upućuju na moguće rizike, povećat će se pouzdanost i točnost probira djece, neovisno o tome koja se metoda probira koristi. Naravno, pouzdanost procjena povećava se korištenjem validiranih mjera probira, pa i na tom aspektu treba raditi, stalnim provjerama i ispitivanjima osjetljivosti i specifičnosti dostupnih mjera.

Zaključak

Rani probir djece, koji pouzdano i precizno utvrđuje odstupanja na specifičnim razvojnim područjima od iznimne je važnosti za pravovremenu intervenciju i podršku obiteljima. Razlike u procjenama različitih procjenjivača i nedovoljno specifične procjene mogu biti posljedica nedovoljno pouzdanih mjera probira, ali i odraz specifičnih razlika u informacijama koje su pojedinim procjenjivačima dostupne. Slaganja između procjena roditelja i odgojitelja za različita razvojna područja dobivena u ovom istraživanju su umjerena, a za pojedine podskale čak i jako niska. To je djelomično odraz psihometrijskih karakteristika mjerenoga instrumenta, ali i različitih okolina i

referentnih grupa s kojima roditelji odnosno odgojitelji uspoređuju razvojna odstupanja djeteta. Na primjer, neki roditelji imaju samo jedno dijete i ne mogu uspoređivati njegov razvoj s braćom/sestrama ili pak nemaju dovoljno znanja o razvoju djece, dok drugi roditelji možda imaju više djece i više znanja o tipičnom djetetovu razvoju ponekad i više znanja od odgojitelja (npr. roditelji koji su ujedno pedijatri ili psiholozi). Rezultati upućuju na važnost kombiniranja procjena iz različitih izvora te važnost procjenjivanja različitih aspekata razvoja (s obzirom na to da jedno zasebno područje razvoja može biti nepouzdan izvor informacija o potencijalnom odstupanju). Važno je nastaviti ispitivanja ovakvih slaganja i pouzdanosti procjene razvojnih odstupanja na većim uzorcima, uzimajući u obzir demografske karakteristike procjenjivača te kombiniranjem procjena i članova obitelji i različitih stručnjaka (odgojitelji, pedijatri, psiholozi, rehabilitatori, logopedi i sl.) koji rade s ovom populacijom djece kako bi se dobili osjetljivi i pouzdani mjerni instrumenti za probir i praćenje razvoja djece rane i predškolske dobi.