

TEACHERS' ATTITUDES TOWARDS EVALUATION PROCESS

dr. sc. **Ante Kolak**, assistant professor

Abstract: Evaluation is an important part of education process at school. It plays a significant role in the development of school culture.

The paper deals with the terms connected with the evaluation process in teaching and explains the differences between monitoring, testing, assessing, measuring and grading. Special attention has been given to the purpose, function, types and forms of evaluation within the teaching process. The author also focuses on two different kinds of evaluation - student-directed and teacher-directed evaluation.

The research presented in this paper studied teachers' attitudes towards the evaluation process in primary schools. The aim was to examine their opinions and attitudes to evaluation. Furthermore, the author tried to identify variables which could be linked to positive or negative attitudes towards evaluation.

As a result, the number of years of working experience has turned out to be connected with teachers' attitudes towards evaluation in their classrooms, whereas other examined factors showed no significant differences in teachers' attitudes.

Keywords: evaluation, monitoring, grading, teachers' attitudes.

Introductory reflections

The term evaluation comes from the French word *evaluation* which means “determining values, grading, assessing” meaning that this process is the same as assessment. Evaluation is an extremely important part of the educational process whose aim is to fulfil certain goals and assess validity or evaluate the realisation. The evaluation process can be defined as “*systematic data collection in the process of learning and the achieved competence level: knowledge, skills, abilities, independence and work responsibility in accordance with predefined methods, procedures and elements whose components are monitoring, assessment and grading*” (Regulations of methods, procedures and elements of evaluation, NN 92/95). “*Evaluation characteristics indicate that it is a planned and systematic process determined by time periods and it includes analysis and evaluation; it is oriented towards exact criteria as starting points for analysis and evaluation which has a*

concrete benefit.” (Jurić, 2005, p. 288). Assessment and measuring can be added to the abovementioned components since they are also an integral part of evaluation. Regular monitoring, assessment and grading have an essential role in the entire process of education.

According to Regulations of methods, procedures and elements of evaluation, monitoring is *“a systematic observing and recording of remarks about the achieved competence level and assigned tasks defined by the national curricula, syllabi and programmes.”* Some authors pay special attention to measuring in the teaching process (Dockrell, 2001) and emphasize that teachers should not be laypeople in measuring. Measuring is defined as a procedure of objective validation of student’s achievements and it refers to expressing an individual’s knowledge and skills as specific numerical values (number of points in a test, for example) obtained by using a valid instrument and comparison of those values with other students’ results in *“the same situation.”* Measuring and assessment are not the same since assessment is not as precise as measuring – assessment determines a student’s position on a certain scale. Measuring and assessment provide a factual base serving as a foundation for grading, i.e. they are a pre-phase of grading. Grading represents assigning a numerical or descriptive value to monitoring and assessment results of students’ work according to grading components of each subject. In this process, as Grgin (1994) states, the value of student’s answer is assessed and this assessment is then expressed as a grade. Evaluating progress in student’s development implies determining the relation between assigned tasks and achieved results in acquiring learning content. Such evaluation, according to Pongrac (1980) is called assessment. Regulations of methods, procedures and elements of evaluation of students in primary and secondary school (NN 92/95) defines assessment as *“appraisal of the achieved level of competences in a subject or field or other forms of work at school during the school year.”* Encyclopaedic dictionary of pedagogy states that assessment is carried out in order to *“control quantitative and qualitative level of students’ educational achievements.”* According to these two definitions, authors, when defining assessment, emphasise the final aim of education and that is the level of acquisition of learning content. Some authors define assessment as *“collecting data (examination, evaluation, etc.) about how students approach defined aims”* (Andrilović & Čudina-Obradović, 1996, p. 120).

The second part discusses basic approaches to evaluation, types of evaluation and its fields. There are two basic approaches to each evaluation, the evaluation of a process and the evaluation of results i.e. consequences. Besides these, some authors identify two more categories: the evaluation of needs and efficacy (Posavac & Carey, 2003). Apart from approaches, there are differences between basic types of evaluation. Some authors emphasise complex theoretically-led programmes of evaluation in education and state that

“a programme theory and logic of a small or medium range are necessary for creating any type of programme or evaluation, whether it is a formative, summative, proactive, monitoring or interactive form of evaluation” (Halmi, 2010, 23). The author states that evaluation is often seen as being atheoretical or exclusively methodologically directed mission. The process of evaluation is one of the basic curricular tasks of schools and when fulfilling this task a diagnostic, formative, final or summative evaluation is used. A diagnostic evaluation records the current situation (when student-directed, it can, for example, be initial or final diagnostic; when teacher-directed, it records the current state of the teaching process). A formative evaluation is carried out during the process itself in order to eliminate possible errors (if it is directed at students, it, for example, eliminates difficulties with acquiring certain learning content, and if it is directed at teachers, it is carried out in order to improve the quality of teaching). Final or summative evaluation is carried out at the end of the school year (if it is student-directed), at the end of a trial period (teacher-directed) or at the end of a certain programme (projects, internship). When participants themselves evaluate, we call it internal evaluation and evaluation made by other individuals is external evaluation (Jurić, 1993; Kiper & Mischke, 2008). The evaluation of evaluation is quite rare in theory and teaching practice. A few general criteria are suggested for the assessment of evaluation quality: integrity, focus on improvement, sensitivity of evaluation, and being future-oriented (Kunkel & Tucker, 1977; according to Vizek Vidović et al., 2003). Integrity criterion is essential no matter whether evaluation is directed at students or teachers. Each evaluation should promote development, improvement and growth (of teacher's skills or student's achievements) and it should not be a mere criticism of a current situation. An evaluator should be capable of developing and growing to the same extent as participants. Evaluation is future-oriented if it contributes to improvement, and when its only aim is not only to record present conditions. Besides students, we can evaluate teaching and learning, school organization, school employees' qualities, school atmosphere... This chapter deals with two fields. The first is evaluation of students and their achievements and the second is evaluation of school employees, i.e. teachers.

Student-directed evaluation

A part of every teacher's work that has always been a challenge, is the evaluation of students' knowledge and skills. These activities, as well as grades, have various consequences for students, their parents, teachers and schools. There are numerous objections against evaluation procedures. Reflections on evaluation start from criticism emphasising that evaluation

procedures dehumanise teaching and create distrust between teachers and students. Furthermore, another negative aspect is students' anxiety. Students with poor results have a bad self-image which affects their self-esteem. Evaluation generates competitive atmosphere among students influencing class climate, students' satisfaction and hierarchy of social relations within a certain class. We often wonder whether a grade realistically reflects students' knowledge and whether an excellent grade is the indicator of success. Other perspectives regard evaluation as an integral part of the teaching process. Since evaluation and grading are burning and challenging issues, especially objections to grading, they have brought about certain changes in the school system. Numerical grades have been replaced by descriptive grades. Research results indicate that students graded descriptively have somewhat worse test results, but they are less anxious and more intrinsically motivated and have a better attitude to school. Additionally, there are no differences in academic locus of control for students graded numerically. Researchers conclude that a type of evaluation has different consequences with regard to observed variables, and there are different effects after a four-year education with positive influence of descriptive grading, but if they are studied after additional four years of education with numeric grades, there are no positive effects. Students graded descriptively have more extrinsic motivation and they assessed their success as being more dependent on factors out of their control (Arambašić et al., 1991). The results indicate that one should be careful when changing the evaluation system, which can produce some unexpected outcomes. One of the most important issues that we face when evaluating students is the purpose. Regarding the most common purpose of evaluation, Kyriacou (2001) names the following: getting feedback about student's progress, providing pedagogic feedback to students, motivating students, recording progress, stating present achievements, evaluating student's readiness for future learning, confirming teachers and school's efficacy. Feedback can serve as an indicator of teachers' successful or unproductive teaching and it can also indicate difficulties and misunderstandings which can be avoided in further teaching. Students can compare their own achievements with the expected standard in order to improve their work and become familiar with the requirements of certain activities. Evaluation can be motivational for students and encourage them to organise their work well and learn everything needed for a certain activity. Success often generates motivation and teachers should bear in mind that students need to be motivated since poor results and failure can be a consequence of students' lack of motivation. Studies show that every negative grade can be motivating for some students (Matijević, 2005). On the other hand, some students after getting one or two negative grades become completely helpless. Regular evaluation provides teachers with written recording of students' progress during a longer period of time and can serve as

a basis for assessment of students' present and future educational needs. Such records can be used in communication with others, including parents, as well as for future planning of teaching. A special evaluation activity can determine a standard of achieved success at a certain moment which can be a basis for obtaining a diploma or a certificate. Evaluation can determine whether students are ready for a certain type of learning, whether they have any learning difficulties, whether they have acquired the previous content in order to be ready to successfully continue learning. Evaluation of students' work shows their achievements and can serve as a proof of their progress and teacher's good work. Therefore it is also a good indicator of students and school's effectiveness. The abovementioned purposes point to the main functions of evaluation: diagnostic, prognostic, motivational and therapeutic function. Some authors state that educational standard should be used as the basis for the evaluation of student's achievement which is one of its main functions (Palekčić, 2007). There are three basic models of student evaluation: synthetic, analytic and combined. The choice of a model depends on a school level and specific requirements of a certain subject. Evaluation models should be different for formal and informal education, as well as for beginner and final grades of education. Analytic model is used in the history of language teaching, it evaluates students' knowledge according to seven variables (reading printed texts, reading written texts, calligraphy, grammar, punctuation, written assignments, pronunciation). Since the grading model after the first term of the first grade in primary school has a range of grades from one to five (there is a possibility to repeat a grade), some alternative branches suggest teaching processes which monitor and evaluate students without numerical grades and without an option of grade retention. The range of grades in different schools points to various models. Especially since PISA studies there has been an increased interest for Finland and its model of evaluation. After publishing PISA results (Finnish students were the best in all studied variables) "*Finland has become world-known and the interest has not diminished, while the Finns are still self-critical and discuss what else they should improve in their education system*" (Bašić, 2007, p. 136). When considering solutions for improving Croatian docimology model, we can rely on some Finnish docimology concepts taking into consideration the risks of mimetic systematic model which imitates other models and uses other people's experience. Their model of evaluation has a seven-level scale (from four to ten). The premises of the Finnish model have numerous characteristics different from the Croatian model. A distinct difference is the one regarding the highest grade since in Finland only a very small number of students deserve this grade. Only the ones who particularly stand out in fulfilling the requirements of the national education standard belong to this category, whereas in Croatian docimology model the best grade is for all those students

who satisfy education standards (Matijević, 2007). Since school system is a type of social system involving human factor, it automatically points to imperfections of a system such as a subsystem of grading. Some authors discuss fairness in evaluating students and stress the necessity of being fair at all times (Gipps & Murphy, 1994). Injustice in evaluation is a source of numerous students and parents' negative reactions. Errors in evaluation are usually classified as personal equations of an evaluator, halo effect, logical errors, mean, differentiation and contrast errors and adapting criteria to a group. Docimological anomalies can be observed in all the above mentioned categories and their conceptions are a part of teacher - evaluator's beliefs – only God deserves an excellent grade, or that every student who is present in the lesson must possess sufficient knowledge.

Teacher-directed evaluation

School inspectors, head teachers, counsellors and psychologists are usually the ones who evaluate teachers. Self-evaluation, a process in which teachers themselves record and evaluate their own work, has also become more common. There are two conceptual models of teacher evaluation: developmental and differential. A developmental model sees teachers as adults and emphasises two dimensions of teacher's personality: level of motivation and dedication to work and level of abstraction (Glickmann & Gordon, 2012). The level of abstraction (level of cognitive development and abstract thinking flexibility) classifies teachers as the ones with low, medium and high abstraction. This conceptual model consists of four categories of teachers.

High level of abstraction and motivation denotes a teacher who is a professional, whereas low level of abstraction and motivation means that a teacher is not suitable for this job. Low level of motivation with high flexibility means that a teacher is an analytical observer, while low flexibility and high motivation denote a "futile" teacher (Vizek Vidović et al., 2003; Glickman, 2012).

There are three orientations of teacher evaluation: non-directive, cooperative and directive. Non-directive orientation is acceptable if teachers themselves are able to improve their work after an analysis carried out by means of listening, explaining, encouraging and presenting by the observer of the teaching process. In cooperative orientation a supervisor or a teacher can take the initiative when elaborating on an observed problem. There is a possibility of including a mediator and it is conducted by means of discussion, problem solving and demonstration. Directive orientation has a high level of control over teachers since an evaluator directs, evaluates and strengthens a teacher (Rijavec, M. & Miljević Riđički, R. 2000). Differential evaluation

means that different circumstances require different approaches and takes into consideration the fact that teachers reluctantly agree to be evaluated. There are four types of teacher evaluation: clinical, cooperative, administrative and self-evaluation (Kagan & Warren, 2012). Clinical evaluation provides teachers with objective information about their teaching skills with the help of a supervisor. In cooperative evaluation there is cooperation between two expert teachers who observe their lessons, analyse and have a role of a critical friend. In self-evaluation teachers independently work on their professional development and growth, without being evaluated externally. Administrative evaluation is the least favourite and it is also known as control or inspection.

Methodology of empirical research

The aim of the research was to examine the attitudes of lower primary school teachers to different aspects of evaluation and identify variables which can influence the direction of examined attitudes.

There are two hypotheses:

H1 – Participants' attitudes are mostly positive,

H2 – Participants' attitudes are significantly different depending on the participants' certain characteristics and specific working conditions.

The dependent variable is participants' attitudes to evaluation and independent variables are their age, work experience and school size. The sample consisted of 1308 teachers, 1210 female and 98 male teachers. There are five subsamples regarding their age, six for work experience and three for school size. The research was conducted using questionnaires and assessment and the instrument is a combination of a questionnaire and an assessment scale. The instrument contains 17 items relating to different segments of evaluation. The participants answered to items positively or negatively on a Likert five-point scale.

Results and discussion

Items	M	SD	N
I create informal ways of assessment.	3.82	0.87	1307
I ask students to constructively comment on their and other students' work.	4.17	0.73	1307
I ask students to help others evaluate their work by making comments.	4.18	0.69	1307
I entice students to make and elaborate on their decisions.	4.36	0.68	1307
I entice students to listen critically.	4.4	0.67	1307
I involve students in active creation of grading criteria.	4.14	0.77	1307

In my work I use lists for self-evaluation.	3.44	1.05	1307
I reflect on my own docimological practice and make plans for improvement when necessary.	4.04	0.86	1307
I evaluate my work results and systematically ask for feedback from students, parents, counsellors, colleagues, head teachers...	3.96	0.86	1307
I look forward to new ideas in student evaluation and experiment with them when planning my teaching.	4.16	0.8	1307
I conduct action research and / or cooperate with researchers who study pedagogical practice.	3.24	1.17	1307
I am up to date with recent pedagogical research regarding evaluation and I implement these results in my work.	3.91	0.87	1307
I use expert and scientific works, seminars and other opportunities to broaden my knowledge of evaluation.	4.42	0.7	1307
I am always ready to introduce changes to student evaluation.	4.44	0.7	1307
I have a portfolio of professional development with my ideas and experience from practice.	3.12	1.18	1307
I openly exchange support, advice, feedback and criticism with my colleagues.	4.38	0.74	1307
My knowledge and skills are being completely used in my work place.	4.06	0.79	1307

Table 1. Descriptive parameters of items for measuring evaluation

The highest arithmetic mean is observed in items “I am always ready to introduce changes to student evaluation” (M=4.44), “I use expert and scientific works, seminars and other opportunities to broaden my knowledge of evaluation” (M=4.42), and “I entice students to listen critically” (M=4.40) with the lowest standard deviation (SD=0.67). The lowest arithmetic mean is for the item “I have a portfolio of professional development with my ideas and experiences from practice” (M=3.12) with the highest standard deviation (SD=1.18). Participants' answers reflect readiness to change current evaluation procedures of the teaching process and work with students. Apart from that, participants express a desire to broaden their knowledge in this field. It is also obvious that the process of self-evaluation is in its beginnings and we have to pay special attention to the mentioned techniques, procedures and self-evaluation criteria (Ljubetić, 2007).

Factor analysis of common factors was carried out on 17 items, which according to their content, mostly measure evaluation of the teaching process. Kaiser-Meyer-Olkin coefficient, which is 0.93, confirmed the adequacy of item correlation matrix of assessment scale for evaluation of the teaching process thus indicating that the observed items belong together psychometrically. Bartlett's test (χ^2 136, $p < 0.01$] =9122.43) indicates that there is no linear

dependence between the items thus confirming the matrix adequacy for analysis.

Kaiser-Meyer-Olkin measure		0.93
Bartlett's test	estimated χ^2	9122.43
	Df	136
	P	<0.01

Table 2: Indicators of item correlation matrix adequacy for factor analysis

Items	factor 1
I am up to date with recent pedagogical research regarding evaluation and I implement these results in my work.	0.69
I evaluate my work results and systematically ask for feedback from students, parents, counsellors, colleagues, head teachers...	0.66
I ask students to help others evaluate their work by making comments.	0.66
I entice students to make and elaborate on their decisions.	0,65
I involve students in active creation of grading criteria.	0.65
I look forward to new ideas in student evaluation and experiment with them when planning my teaching.	0.65
I ask students to constructively comment on their and other students' work.	0.63
I entice students to listen critically.	0.63
I am always ready to introduce changes to student evaluation.	0.62
I use expert and scientific works, seminars and other opportunities to broaden my knowledge of evaluation.	0.62
I reflect on my own docimological practice and make plans for improvement when necessary.	0.59
I create informal ways of assessment.	0.58
In my work I use lists for self-evaluation.	0.57
I conduct action research and / or cooperate with researchers who study pedagogical practice.	0.55
I have a portfolio of professional development with my ideas and experience from practice.	0.53
I openly exchange support, advice, feedback and criticism with my colleagues.	0.47
My knowledge and skills are being completely used in my work place.	0.40

Table 3: Matrix of factor structure

Table of factor structure indicates that a factor named Factor of evaluation has a high saturation on items "I am up to date with recent pedagogical research regarding evaluation and I implement these results in my

work” (saturation=0.69), “I evaluate my work results and systematically ask for feedback from students, parents, counsellors, colleagues, head teachers...” (saturation=0.66), “I ask students to help others evaluate their work by making comments” (saturation=0.66). All items have a moderate saturation with the first factor and according to their content they describe evaluation of the teaching process by teachers whether it is evaluation of their own work or evaluation of teacher’s work by students.

Age group	M	SD	N
up to 30 years	3.97	0.53	193
from 31 to 40 years	3.97	0.51	395
from 41 to 50 years	4.03	0.54	424
from 51 to 60 years	4.04	0.53	212
61 and older	4.13	0.43	83
total	4.01	0.52	1307

Table 4: Descriptive statistics for Assessment scale according to participants' age

Table 4 indicates that the teachers who are more than 61 years old express the highest level of agreement with the items of the instrument. It is estimated that they mainly agree with the ways of evaluation suggested by the items and they most commonly carry it out. The teachers from the first two subsamples (up to 40 years) express the lowest level of agreement, although their attitudes also have a positive direction. A subsample with the highest level of agreement with the mentioned statements also has the lowest standard deviation. In order to determine whether there is a significant difference, we carried out a one-way analysis of variance.

	Sum sq	df	Mean sum sq	F	p
Between groups	2.4	4	0.6	2.22	p>0.05
Within groups	352.24	1302	0.27		
Total	21415.61	1307			

Table 5: Analysis of variance for age difference for Assessment scale

The analysis indicates that there are no statistically significant differences between the teachers regarding their age when expressing their opinion about evaluation (F=2.22; df1=4; df2=1302; p>0.05).

Length of work experience	M	SD	N
up to 2 years	3.97	0.51	113
from 2 to 5 years	3.91	0.51	111
from 5 to 10 years	3.96	0.52	152
from 10 to 15 years	3.97	0.57	203
from 15 to 20 years	4.02	0.55	197
more than 20 years	4.07	0.49	531
total	4.01	0.52	1307

Table 6: Descriptive statistics for Assessment scale according to work experience

Table 6 indicates that the teachers generally agree with the items and the level of agreement is mainly homogenous regardless of the length of their work experience. The lowest level of agreement is expressed by the teachers from the second category (from 2 to 5 years of work experience) even though they mainly agree with the items offered and the differences in arithmetic means are insignificant. The teachers from the last two categories, with more than 15 years of work experience, express the highest level of agreement with the items. The data points to analogy with attitudes of teachers from older age groups. This leads to the conclusion that life and work experience increase teachers' experience regarding evaluation and a long-term professional development can change some attitudes towards evaluation. In order to determine whether there is a significant difference, we carried out a one-way analysis of variance.

	Sum sq	df	Mean sum sq	F	p
Between groups	4.13	5	0.83	3.06	p<0.01
Within groups	350.51	1301	0.27		
Total	21415.61	1307			

Table 7: Analysis of variance for work experience for Assessment scale

The analysis of variance indicates that there is a significant difference between the age groups considering their opinion about evaluation ($F=3.06$; $df_1=5$; $df_2=1301$; $p<0.01$). In order to determine between which subsamples there is a significant difference in attitudes to evaluation, we carried out a post-hoc test (Scheffe test).

Years of work experience	Years of work experience	Arithmetic mean difference	p
up to 2 years	from 2 to 5 years	0.06	p>0.05
	from 5 to 10 years	0	p>0.05
	from 10 to 15 years	0	p>0.05
	from 15 to 20 years	-0.06	p>0.05
	more than 20 years	-0.11	p>0.05
from 2 to 5 years	from 5 to 10 years	-0.05	p>0.05
	from 10 to 15 years	-0.06	p>0.05
	from 15 to 20 years	-0.11	p>0.05
	more than 20 years	-0.16	p<0.01
from 5 to 10 years	from 10 to 15 years	-0.01	p>0.05
	from 15 to 20 years	-0.06	p>0.05
	more than 20 years	-0.11	p>0.05
from 10 to 15 years	from 15 to 20 years	-0.05	p>0.05
	more than 20 years	-0.1	p>0.05
from 15 to 20 years	more than 20 years	-0.05	p>0.05

Table 8: Post-hoc test (Scheffe test) results for evaluation according to participants' work experience

A post-hoc test calculates the significance of difference between the two groups after the analysis of variance taking into consideration multiple comparisons and the level of significance. It indicates that there is a significant difference between the participants from group with 2 to 5 years of work experience and those with more than 20 years of work experience (difference=-0.16; p<0.01) since more experienced teachers express a higher level of agreement with the statements about evaluation as a characteristic of the teaching process. It is evident that more experienced teachers more often create informal ways of evaluation, they entice students to participate in evaluation of their own and other students' work, they make plans for improvement, they systematically record results, they are better informed about recent pedagogical research and professional papers and works, they exchange their opinions more openly and they more commonly have a portfolio of their professional development. This conclusion is reasonable since teachers with little work experience do not feel confident enough about evaluation. It seems that the work experience category plays a vital role in developing teacher's competence for evaluation.

School size	M	SD	N
Small – up to 300 students	4.04	0.52	373
Medium – between 300 and 700 students	3.98	0.53	511
Big – more than 700 students	4.03	0.52	422
Total	4.01	0.52	1306

Table 9: Descriptive statistics for Assessment scale according to school size

Table 9 indicates that the teachers from certain subsamples mainly agree with the statements. Arithmetic means show that the teachers from small (up to 300 students) and big schools (more than 700 students) express the highest level of agreement. The teachers working in medium-sized schools express a somewhat lower level of agreement ($M=3.98$). Standard deviations in all subsamples are almost identical. In order to determine whether there is a significant difference, we carried out a one-way analysis.

	Sum sq	df	Mean sum sq	F	p
Between groups	1.12	2	0.56	2.07	$p>0.05$
Within groups	353.25	1303	0.27		
Total	21395.1	1306			

Table 10: Analysis of variance for school size for Assessment scale

The analysis of variance indicates that there is no significant difference between the teachers regarding the school size and their attitudes to evaluation ($F=2.07$; $df_1=2$; $df_2=1303$; $p>0.05$).

Concluding remarks

The research has confirmed the first hypothesis and established that participants' attitudes have a chiefly positive direction. The participants indirectly express the need for changing the evaluation model of their own work as well as of students' work. They also express a positive attitude and willingness to develop as evaluators to the same extent as the ones being evaluated. The research was directed at two fields of evaluation – evaluation of teachers and students. In teacher evaluation, i.e. evaluation of their own work, the participants tend to prefer differential evaluation, especially cooperative. These findings confirm the results of the studies mentioned in the theoretical part of this chapter. Therefore it is necessary to motivate teachers to jointly record and analyse teaching and to strengthen the role of a critical friend in this part of evaluation. Taking into consideration participants' motivation, this type

of evaluation could achieve the best results. It is significant that there is more readiness for elements of self-evaluation. It is necessary to study more thoroughly why this type of evaluation is least being carried out as well as the obstacles and difficulties in self-evaluation. One can assume that teachers should be more carefully introduced to the abovementioned techniques and procedures of self-evaluation and provided with necessary leadership and support.

The second hypothesis is only partially accepted. Considering all participants' characteristics only the length of work experience turns out to be more significant since it has a more positive influence in the sense of teachers being ready for changes regarding evaluation. Other variables describing participants' characteristics and working conditions do not indicate any statistically significant relationship with participants' attitudes. In order to determine the relationship, it is necessary to include some other characteristics and more specific working conditions. It would be interesting to study teachers' attitudes to evaluation depending on participants' motivation, work dedication and success indicators of teaching, which could all be topics of future research.

References:

1. Andrilović, V., & Čudina-Obradović, M. (1996). *Psihologija učenja i nastave*. Zagreb: Školska knjiga.
2. Arambašić, L., Vizek-Vidović, V., & Lugomer-Armano, G. (1991). Some personality characteristics of primary school students under two different evaluation systems of academic achievement. In I. S. Ribeira & L. S. Almeida (Eds.), *School psychology and human development* (pp. 324-336). Portuguese: University of Minho.
3. Bašić, S. (2007). *Obrazovni standardi – didaktički pristup metodologiji izrade kurikuluma*. In V. Previšić (Ed.), *Kurikulum – teorije, metodologija, sadržaj i struktura* (pp. 117-156). Zagreb: Školska knjiga.
4. Dockrell, B. (2001). *Pristupi mjerenju u obrazovnom procesu*. In Ch. Desforges (Ed.), *Uspješno učenje i poučavanje* (pp. 283-298). Zagreb: Educa.
5. Dockrell, B. (2001). *Mjerenje, poučavanje i učenje*. In Ch. Desforges (Ed.), *Uspješno učenje i poučavanje* (pp. 299-316). Zagreb: Educa.
6. Gipps, C., & Murphy, P. (1994). *A fair test? Assessment, achievement and equity*. England: Open University Press.
7. Glickman, C. D. *Developmental supervision: alternative practices for helping teachers improve instruction*. (ED208487). Available at <http://www.eric.ed.gov/ERICWebPortal/>
8. Glickman, C. D., & Gordon, S. P. *Clarifying Developmental supervision. Supervision in context*. (EJ353888). Available at <http://www.eric.ed.gov/ERICWebPortal/>

9. Grgin, T. (1994). *Školska dokimologija*. Jastrebarsko: Slap.
10. Halmi, A. (2010). Kompleksni teorijski vođeni programi evaluacije u obrazovanju. *Pedagogijska istraživanja*, 7(1), 23-43.
11. Jurić, V. (1993). Praćenje i vrednovanje odgojno-obrazovnog procesa. In B: Drandić (Ed.), *Priručnik za ravnatelje odgojno-obrazovnih ustanova* (pp. 321-337). Zagreb: Znamen.
12. Jurić, V. (2007). Kurikulum suvremene škole. In V. Previšić (Ed.), *Kurikulum – teorije, metodologija, sadržaj i struktura* (pp. 253- 308). Zagreb: Školska knjiga.
13. Kagan, D. M., & Warren, E. *Differential supervision for early, middle, and late field experiences in teacher education* (EJ442242). Available at <http://www.eric.ed.gov/ERICWebPortal/>
14. Kiper, H., & Mischke, W. (2008). *Uvod u opću didaktiku*. Zagreb: Educa.
15. Kyriacou, C. (2001). *Temeljna nastavna umijeća*. Zagreb: Educa.
16. Ljubetić, M. (2007). (Samo)vrednovanjem do kvalitetnije odgojno-obrazovne prakse. In V. Previšić, N. N. Šoljan, N. Hrvatić (Eds.), *Pedagogija prema cjeloživotnom obrazovanju i društvu znanja*, Zagreb: HPD.
17. Matijević, M. (2005). Evaluacija u odgoju i obrazovanju. *Pedagogijska istraživanja*, 2(1), 23-43.
18. Matijević, M. (2007). Evaluacija u nastavnom kurikulumu škole. In V. Previšić (Ed.), *Kurikulum – teorije, metodologija, sadržaj i struktura* (pp. 309- 350). Zagreb: Školska knjiga.
19. Palekčić, M. (2007). Od kurikuluma do obrazovnih standarda. In V. Previšić (Ed.), *Kurikulum – teorije, metodologija, sadržaj i struktura* (pp. 39-105). Zagreb: Školska knjiga.
20. Pongrac, S. (1980). *Ispitivanje i ocjenjivanje u obrazovanju: primjena u osnovnom obrazovanju odraslih*. Zagreb: Školska knjiga.
21. Posavac, E. J. & Carey, R. G. (2003). *Program evaluations. Methods and case studies*. New Jersey: Prentice Hall.
22. *Pravilniku o načinima, postupcima i elementima vrednovanja*. Narodne novine, 92/95.
23. Vizek Vidović, V. et al. (2003). *Psihologija obrazovanja*. Zagreb: IEP – Vern.

Stavovi učitelja o procesu evaluacije

Sažetak: Uvodni dio rada usmjeren je na terminološke odrednice vezane uz proces vrjednovanja u nastavi razlikujući sastavnice vrjednovanja kroz praćenje, provjeravanje, procjenjivanje, mjerenje i ocjenjivanje. U razradi pojedinih sastavnica usmjerena je pozornost na svrhu, funkciju, vrste i oblike vrjednovanja unutar nastavnoga procesa. Razlikujući različite nastavne aktivnosti vezane za proces vrjednovanja istaknute su prednosti i nedostaci pojedinih, a i moguće pogriješke.

Predstavljeno je istraživanje u kojemu je istraživačko pitanje bilo usmjereno na ispitivanje mišljenja učitelja o procesu vrjednovanja u osnovnim školama. Cilj je bio doznati njihova mišljenja i stajališta vezana uz vrjednovanje kao iznimno važan dio nastavnoga procesa, a i identificirati varijable koje mogu utjecati na smjer mišljenja i stajališta učitelja.

Istraživanje je provedeno na uzorku učitelja razredne nastave diljem Hrvatske (N=1308). U istraživanju se koristilo metodom anketiranja i procjenjivanja, a instrument je kombinacija ankete i ljestvice procjene. Za utvrđivanje deskriptivnih pokazatelja upotrijebljeni su deskriptivni parametri, a faktorska analiza zajedničkih faktora upotrijebljena je kako bi se utvrdile latentne dimenzije koje se nalaze u podlozi interkorelacije čestica. Za usporedbu ispitanika po zadanim obilježja upotrijebljena je jednosmjerna analiza varijance.

Na 17 čestica koje prema sadržaju mjere evaluaciju nastavnoga procesa provedena je faktorska analiza zajedničkih faktora kojom su ekstrahirana tri faktora po Kaiser-Guttmanovu kriteriju. Pogodnost matrice korelacija čestica skale evaluacije nastavnoga procesa za faktorsku analizu potvrdio je Kaiser-Meyer-Olkinov koeficijent koji iznosi 0,93. Pouzdanost tipa unutarnje konzistencije, Cronbachova alfa, potvrdila je da ljestvica evaluacije ima visoku pouzdanost. U cilju potvrđivanja hipoteze koja se odnosi na mišljenje učitelja vezano uz pojedina obilježja ispitanika provedena analiza varijance upućuje na postojanje statistički značajne razlike između ispitanika vezanih uz radno iskustvo i dob, dok na ostalim obilježjima nije utvrđena značajnost.

S obzirom da je obradom podatka utvrđeno da čestice u ljestvici evaluacije visoko korespondiraju tomu da izmjere ispitivani koncept, mogu se upotrijebiti u daljnjim analizama i istraživanjima kao pouzdane mjere, a utvrđene razlike vezane uz pojedina obilježja ispitanika mogu odrediti smjer u daljnjem stručnom usavršavanju učitelja vezanom uz proces vrjednovanja.

Ključne riječi: vrjednovanje, praćenje, ocjenjivanje, mišljenja učitelja.

Einstellungen der Lehrer zum Evaluationsprozess im Unterricht

Zusammenfassung: Der Einführungsteil der Studie konzentriert sich auf die terminologischen Determinanten, die sich auf den Prozess der Evaluation im Unterricht beziehen, wobei die Elemente der Evaluation durch Beobachtung, Prüfung, Einschätzung, Messung und Benotung unterschieden werden. Bei der Ausarbeitung der einzelnen Komponenten wurde die Aufmerksamkeit auf den Zweck, Funktion, Arten und Formen der Evaluation im Unterrichtsprozess gerichtet. Bei der Differenzierung zwischen verschiedenen Lernaktivitäten, die mit dem Prozess der Evaluation im Zusammenhang stehen, wurden die Vor- und Nachteile der Einzelnen und die möglichen Fehler hervorgehoben.

Es wird eine Studie vorgelegt, in der die Forschungsfrage auf die Meinungsfrage von Lehrern über den Evaluationsprozess in den Grundschulen gerichtet war. Das Ziel des Beitrages war es, ihre Meinungen und Ansichten in Bezug auf die Evaluation als extrem wichtigen Teil des Lernprozesses herauszufinden, aber auch die Variablen zu identifizieren, die die Meinungen und Einstellungen der Lehrer beeinflussen könnten.

Die Studie wurde an einer Stichprobe von Grundschullehrern in ganz Kroatien durchgeführt (N = 1308). In der Untersuchung wurden die Umfrage- und Bewertungsverfahren benutzt, und das Instrument war eine Kombination aus

Befragung und Bewertungsskala. Zur Feststellung von deskriptiven Indikatoren wurden deskriptive Parameter benutzt, und die Faktorenanalyse von gemeinsamen Faktoren wurde verwendet, um die latenten Dimensionen zu identifizieren, die sich in der Basis der Interkorrelation befinden. Für den Vergleich der Befragten nach den gegebenen Eigenschaften wurde die einfaktorielle Varianzanalyse verwendet.

Es wurde eine Faktoranalyse von gemeinsamen Faktoren an 17 Fragen durchgeführt, die inhaltlich die Evaluation des Unterrichtsprozesses messen. Dabei wurden drei Faktoren nach dem Kaiser-Guttman-Kriterium extrahiert. Das Kaiser-Meyer-Olkin-Kriterium, das 0,93 beträgt, bestätigte die Eignung der Korrelationsmatrix von der Evaluationsskala des Unterrichtsprozesses für die Faktoranalyse. Die Verlässlichkeit vom internen Konsistenztyp, Cronbachs Alpha, bestätigte, dass die Evaluationsskala eine hohe Zuverlässigkeit aufweist. Um die Hypothese zu bestätigen, die sich auf die Meinungen der Lehrer in Bezug auf die Einzelmerkmale der Befragten bezieht, weist die durchgeführte Varianzanalyse darauf hin, dass statistisch signifikante Unterschiede zwischen den Befragten bestehen, die mit Berufserfahrung und Alter zusammenhängen, während die anderen Merkmale keine Signifikanz aufweisen.

Da mit Hilfe der Datenverarbeitung festgestellt wurde, dass die Fragen in der Evaluationsskala stark mit der Tatsache korrespondieren, das befragte Konzept zu messen, können sie bei weiteren Analysen und Forschungen als zuverlässige Maßnahmen verwendet werden, und die festgestellten Unterschiede im Bezug auf die Einzelmerkmale der Befragten können die Richtung der weiteren beruflichen Fortbildung von Lehrkräften im Zusammenhang mit dem Evaluationsprozess bestimmen.

Schlüsselbegriffe: Evaluation, Beobachtung, Benotung, Meinungen der Lehrer.