Art. #2100, 11 pages, https://doi.org/10.15700/saje.v43n3a2100

Effects and challenges to implement differentiated mathematics teaching among fourth graders in Montenegro

Veselin Mićanović and Dijana Vučković
Teacher Training Department, Faculty of Philosophy, University of Montenegro, Podgorica, Montenegro
veselinm@ucg.ac.me
Biljana Maslovarić
Preschool Teacher Training Department, Faculty of Philosophy, University of Montenegro, Podgorica, Montenegro
Nada Šakotić and Tatjana Novović
Department of Pedagogy, Faculty of Philosophy, University of Montenegro, Podgorica, Montenegro

In this quasi-experimental pretest-posttest study we examined the effects of differentiated instruction (DI) in within-class ability groupings of 246 Montenegrin fourth-graders and their ability to solve algebraic equations. We assessed 2 parallel student groups at equal achievement levels to compare DI, in which teaching and work modes were adapted to students' grouping according to previous achievement and pretest scores, and traditional whole-class instruction. Pretest-posttest evaluations were administered to both groups, and observation indicators were evaluated to assess the level and type of student activities, engagement, and individualisation. Students in the homogeneous DI experimental groups with tailored instructions were significantly more successful at solving algebraic tasks than their peers in the traditional whole-class instruction control group. DI improved students' results, but teachers required specific training and significantly more preparation time.

Keywords: algebra; differentiated instruction; mathematical skills; quasi-experimental pretest-posttest; within-class ability grouping

Introduction

In 2017, the Montenegrin elementary school curricula were redesigned to envelop specific learning outcomes and knowledge standards as a basis to adjust instruction to student capabilities (Pavićević, Vučeljić, Lalić, Pavićević & Kostić, 2017). The curricula reform emphasised improving mathematical literacy, a global trend to equip children with life-long skills in the 21st century (Mejer, Turchetti & Gere, 2011). Montenegrin students continually scored poorly in mathematical literacy testing in the Program for International Student Assessment (PISA) (Organisation for Economic Co-operation and Development [OECD], 2018). Educators posit that mathematical instructions were insufficiently adapted to meet student needs in early childhood (Prica, Čolić & Baronijan, 2014). Montenegrin teachers were also insufficiently adapted to the new curriculum and often inflexible in implementing diverse teaching methods (Prast, Van de Weijer-Bergsma, Kroesbergen & Van Luit, 2018; Prica et al., 2014), which is important since teaching methods organise instructions and help to implement a curriculum (Abah, 2020).

Experts proposed differentiated instruction (DI) as an appropriate remedial classroom instructional strategy. DI facilitates teachers' adaption of learning content, process, and products (curricular elements) to fit students' interests, readiness, and learning profiles, i.e., their characteristics (Prast et al., 2018; Prica et al., 2014; Tomlinson, 2014). It is a student-centred approach that positions the teacher as a guide who facilitates students' participation in their learning (Hackenberg, Creager & Eker, 2021). Conversely, non-differentiated or one-size-fits-all teaching is more general (Hertberg-Davis, 2009). This traditional prescription-oriented approach to teaching (George, 2005) is difficult for students and consequently they progress slowly through the curriculum (Gamble, 2011). Tradition is not intransigence but rather adherence to long-standing practices that are familiar and comfortable (Abah, 2020). The teacher and the curriculum are the focus in traditional teacher-directed teaching (Harris & Johnson, n.d.). In traditional mathematics instruction, the instructor teaches the same content to all students who are simultaneously engaged in the same activity (Kesteloot, 2011). This method is often criticised for showing little regard for students' differences (Sammons, 2010).

DI implementation in mathematics involves adapting the curriculum, instruction, and assessment to the needs and abilities of diverse student groups (Livers, Paxton, O'Grady & Tontillo, 2018; McKeen, 2019). However, motivating teachers to implement DI strategies in elementary mathematics may prove difficult because of their poor understanding of the underlying concepts or disbelief in its effectiveness (Livers et al., 2018; Van Geel, Keuning, Frèrejean, Dolmans, Van Merriënboer & Visscher, 2019). For instance, Gaitas and Martins' (2017) evaluation of 273 Portuguese teachers found that they experienced difficulty implementing DI in regular classes because of student characteristics. Instead, it was easier for teachers to not to apply differentiation. Some authors admitted moderate success using DI in mathematics instructions. Maxey's (2013) causal-comparative design of second-graders at a United States of America (USA) military base found that high groups improved compared to average and low groups. Correspondingly, Prast et al.'s (2018) survey found positive effects in student achievement growth with DI mathematics in a large-scale teacher professional development program. Similarly, Livers et al.'s (2018) phenomenologically themed collaboration illustrated how teacher candidates gained confidence using DI to teach elementary mathematics through curriculum compacting. The latter allows teachers to bypass already mastered content, provide advanced students with more challenging content, and focus on students with less mastery.

In this paper, through the results of a pedagogical experiment involving two groups of Montenegrin fourth-graders, we assess the relative effectiveness of differentiated and non-differentiated mathematics instruction to enhance students' ability in algebra. By the fourth grade, mathematics becomes more abstract and complex so students experience difficulties when algebra is introduced (Bender, 2013). We also identify teaching challenges in implementing DI and propose recommendations.

Literature Review

Elementary Mathematics and DI

Students' success in learning mathematics depends on the way in which mathematics is taught (Boaler, 2002; Papanastasiou, 2002). Teaching mathematics at elementary school is important as basic knowledge is acquired and habits and learning styles are developed (Doubet & Hockett, 2017; Prica et al., 2014). DI gained importance when its benefits and successes for teachers and students within the classroom were highlighted (Papanastasiou, 2002; Prast et al., 2018; Sousa & Tomlinson, 2011). The approach presents a solution for teaching mathematics to students at different levels of mastery (Livers et al., 2018), as teaching and learning can be tailored to meet the students' needs, abilities, and learning profiles (Prast et al., 2018). However, Tomlinson and Imbeau (2010:15) observed that it was first necessary to differentiate the four main curriculum elements, namely "content, process, product, and affect."

Mathematics teaching should be organised based on differentiation principles if teachers want students to achieve better results (Light & Pierson, 2014; Maxey, 2013; Prast et al., 2018). Algebra and fourth-graders are a significant coupling as problem-solving becomes more complex and students mature (Bender, 2013). Therefore, applying the same approach to all students at this stage may be difficult due to the unevenness of their progression (Tomlinson, 2014). Thus, this awareness of student needs and subject knowledge are key factors to influence successful DI (Papanastasiou, 2002).

DI in elementary Mathematics and ability grouping

Ability grouping of elementary students in the learning of mathematics is a polemic issue among educators (Anthony & Hunter, 2017). Nevertheless, most would concur that students are unique and learn at different rates. Flexible grouping in DI of content areas such as algebra allows teachers to group students based on their strengths and abilities (McKeen, 2019). Some educators frown upon the practice as they contend that it promotes lower self-esteem among students of average ability and reinforces social inequality (Boaler, Wiliam & Brown, 2000). However, Boaler (2013) observed that students scored higher in international tests in countries that were more flexible about grouping, unlike the USA and European countries where students structured into rigidly fixed groups scored significantly lower in similar exams. Consequently, struggling or low achievers often developed a negative mindset towards learning and their capabilities (Boaler, 2013; Boaler et al., 2000), thus widening the achievement gap between low and high performers (Papanastasiou, 2002).

Ability grouping coupled with curricula revision or differentiation may result in substantial achievement gains for high achievers (Livers et al., 2018; Tieso, 2003). In Finland, differentiation is paired with early identification and flexible support arrangements. Using a multimethod approach that surveyed special education and mathematic teachers in 55 schools, Ekstam, Linnanmäki and Aunio (2015:75) assessed the benefits of the "pull out of the classroom" method for low-performing students to a three-tier incremental support model. This involved general support through a whole class approach at Tier I; intensified support for a limited time with periodic evaluation at Tier 2; and special support and an individual educational plan that required higher-level intervention involving teacher, parent, student, and principal at Tier 3. Schoolwide cluster groups were also perceived as a positive influence on student performance (Matthews, Ritchotte & McBee, 2013).

DI in elementary Mathematics teaching and learning

DI adapts teaching to students' abilities, knowledge, interests, and needs. The overall aim is to ensure that teachers focus on processes and procedures that ensure effective learning (Kesteloot, 2011). Other factors can affect elementary students' mathematical learning skills, such as the presence of cognitive ability like memory, reasoning, processing speed, and oral expression (Fuchs, Fuchs, Compton, Powell, Seethaler, Capizzi, Schatschneider & Fletcher, 2006). In a qualitative study of Malaysian elementary mathematics students it was found that these factors impeded their problem-solving skills and consequently students showed little interest in solving particularly challenging assignments (Tambychika & Meerah, 2010). As students mature, the mathematics curriculum becomes more complex (Bender, 2013), adding to the disinterest.

DI fosters a successful understanding of concepts such as algebra, as it considers different levels of students' knowledge (Cowan & Powell, 2014). Classroom management is key, as Hackenberg et al. (2021) found that teaching in whole-class discussions, attending to small groups, and responding during group work impeded DI. Teachers' struggles to meet all students' needs results in teaching *to the middle*, thus disenfranchising both low and high achievers (Abah, 2020). However, modelling creative mathematical in-class assignments and accounting for task complexity with students' individualised abilities develop their mathematical competencies (Ashley, 2016). Modelling enables mathematically proficient students to apply their knowledge to daily problem-solving, such as writing a multiplication equation or comparing candy prices. Differentiation in mathematics teaching leads to higher student motivation, greater academic achievements (Bal, 2016; Prast et al., 2018), and greater cooperation between students of similar abilities (Hertberg-Davis, 2009). More importantly, DI mathematics instruction improves learning outcomes (Tieso, 2003).

DI requires that teaching methods, instructions, and materials should be adapted for small group learning (Lou, Abrami, Spence, Poulsen, Chambers & D'Apollonia, 1996). Mathematics teachers must be able to differentiate between content and learning methods to respond to distinct abilities, interests, and learning styles of students in the same class (Ashley, 2016; Chamberlin & Powers, 2010). Process and grade levels are, therefore, crucial. Normally, DI research on numerical operations in elementary mathematics largely focuses on older grades (Bal, 2016), therefore, this research is significant as we sought to address this gap by investigating fourth-graders and their learning of algebra.

Research Hypothesis

The use of DI to teach algebraic content contributes to better overall task achievement than whole-class instruction.

Theoretical Framework

DI is rooted in the constructivist theory. The framework engages mathematics learners in activities that match their strengths and preferences and ensures that educators focus on effective teaching that will benefit them (Tomlinson & McTighe, 2006). DI positions the teacher as a facilitator of students' participation in their learning by redesigning and implementing the content (Hackenberg et al., 2021). The implementation of DI in the teaching of algebra can positively impact student learning (Maxey, 2013).

Methodology

In a quasi-experimental pretest-posttest design with six groups we focused specifically on algebraic content and equation solving (general form equations and equations with textual tasks), which are part of the mainstream syllabus. DI and traditional teaching approaches were tested using non-differentiated teaching methods as the **independent variable** and student achievements in solving the mathematical assignments as the **dependent variable**. Demographic variables were not included in this research design, although it is quite possible that some of them (e.g., gender) may have an influence on the students' achievement. However, the teaching approach (DI vs. traditional teaching) was the focus of this study, and since it is a very complex phenomenon, it was decided not to include demographic variables but to have full focus on the teaching approach. The within-class DI groups comprised students whose previous achievements were similar (experimental group), whereas students in the control group were heterogeneous within-class groups (with different previous achievements) who all received the same instruction. Teachers' instruction in the DI group was adapted to students' different learning approaches and needs. Both groups completed the same assignments.

Compliance with Ethical Standards

Participant volunteers were informed about the research purpose. Parents gave consent during parent/teacher meetings and the Montenegrin Bureau of Education Services approved the study (Confirmation number [no.] 01-646).

Sample

Two hundred and forty-six Montenegrin fourth-graders were recruited from four urban elementary schools with a combined population of 3,750 students. Using random sampling, we created two control and two experimental groups (three control groups [CGs] and three experimental groups [EGs]: $n \times 4 = 31$; 1 CG and 1 EG: $n \times 2 = 30$). Eight teachers participated.

Control and experimental group homogenisation were determined based on student achievements in a preliminary test. Montenegro schools require all students to undergo a psychological test on enrolment, after which they are homogenised into classes. Before the study, participants completed a pretest comprising five objective tasks. Results were compared and students were assigned to two parallel classes in each school. Students' achievements in the initial testing were similar in each parallel class. To calculate the power of the two sample tests we used the function pwr.t.test() from the PWR package in the R software. For more details we refer readers to Champely (2020). One of the input values is the so-called Cohen's *d*. According to Cohen (1988) it is the most common way to measure effect size. Therefore, for the $n = 246$, $d = 0.55$ and significance level of 5% it was found that the achieved power was 0.9.

Procedure

Two researchers trained the four EG teachers for 4 days in the planning, preparation, and teaching of algebraic content for fourth-graders based on the DI model. To maintain consistency in all teachers' approaches, the EG teachers and researchers planned the task differentiation process together. Teachers in the CG received no additional or special training because we chose to compare the effects of traditional and DI instruction, and CG teachers were asked to work in the way that they normally did. All were familiar with the experimental programme and possessed basic knowledge of differentiation, which is a compulsory didactic principle and special teaching method during initial teacher education. The teaching groups planned their classes as usual, focusing on the average group.

Differentiation was based on students' current achievement levels as determined from the pretest scores and previous mathematics grades (Roy, Guay & Valois, 2013), which provided a basis for "cognitive or readiness-based differentiation" (Prast et al., 2018:22). These two variables were used for agglomerative hierarchical clustering, where observations were initially grouped into one cluster, which was then successively partitioned. Ward's clustering with squared Euclidian distance was then conducted to create compact, even clusters (Murtagh & Legendre, 2014). The elbow step was located at the 120th observation; hence, resulting in three final clusters. After completing the analysis to determine cluster membership, we applied the k-means technique ($k = 3$) to partition each observation into the cluster with the nearest mean. This technique was applied since the application of the Shapiro Wilk's test indicates that samples are normally distributed. Table 1 shows the descriptive statistics after analysis including *p* values obtained via Shapiro-Wilk's test.

Table 1 Number of units per experimental group cluster

Cluster	*N*	*p* value (Shapiro Wilk's test)	Mean value (test)	Mean value (previous math grades)	Correlation between pretest and previous grades ($p < .001$)	95% CI for the mean value
1	43	0.7124	88.95	4.86	0.98	86.08–90.05 4.29–4.95
2	43	0.5785	77.05	3.87	0.96	75.21–78.99 3.28–3.97
3	37	0.4725	55.42	3.05	0.92	51.11–59.54 2.98–3.18

Differentiation focused on three achievement levels: high performers ($n = 43$), moderately high performers ($n = 43$), and average performers ($n = 37$). EGs and CGs were organised into four classes taught by students' regular mathematics teachers. Once parallel classes were selected, DI was provided to the EG, and non-DI was given to the CG during 12 sessions. In the DI classes, the lesson was not adapted regarding volume, depth, or planned concepts/content. The teaching method was, however, adjusted at the level of instruction and formation of small workgroups. Workgroups were homogeneous in the EGs and heterogeneous in the CGs. Students in the EGs were given special instructions to solve the assignments, whereas the CGs received identical explanations. The mathematical assignments with complex forms included several different operations. All students performed the same tasks from the common curriculum, albeit using different work methods.

Evaluation and Data Analysis

We used an assessment scale to measure the impact of the differentiated approach to solve the mathematical tasks, using knowledge testing (a series of objective tasks) to compare the effects. Ultimately, we did a posttest to assess students' ability to solve the equations. All students solved identical tasks without extra aid. The number of successfully solved tasks formed the evaluation criterion for both pretest and posttest. Students' achievements were assessed using five objective tasks. Both teachers and researchers maintained that this aptly measured students' abilities to solve general form equations and equations with textual tasks. Tests were scored on a range of 0 (unsuccessfully solved tasks) to 5 (successfully solved tasks).

We included two additional experimental techniques: observation and evaluation of direct teaching practice from which the effects of the experimental factors were determined. Indicators were developed to assess the domains of individualisation, student activity, and interest during learning. Two observers independently monitored student activities in the parallel classes and completed observation protocols based on specific indicators. The scores for each indicator were averaged, scored protocols were compared, and we agreed on a position regarding the applicability of the learning strategy for each class. The protocols provided five assessment options for each indicator; the first two categories ranged from 1 (exceptionally high) to 5 (exceptionally low). Data analysis measured objective statistical indicators.

Results

Table 2 is a summary of the pretest results. The EGs and CGs were balanced. The average achievement in the parallel groups was similar, as well as the means for both groups. However, within both groups, there were at least three distinct student levels: excellent grades, very good grades, and sufficient grades.

Table 2 Pretest results

Group	No. of students		No. of students with successfully solved tasks					*M*	*SD*
			5	4	3	2	1		
Experimental	123	*N*	43	43	25	12	0	3.95	0.94
		%	34.95	34.95	20.32	9.75	0.00		
Control	123	*N*	44	42	26	11	0	3.96	0.95
		%	35.77	34.14	21.13	8.94	0.00		
Total	246	*N*	87	85	51	23	0	3.95	0.95
		%	35.36	34.55	20.73	9.34	0.00		

In order to be sure that parametric statistics is an appropriate choice for this analysis we employed Shapiro-Wilk's test. For all groups we got $p > .05$ meaning that all samples are within normal distribution.

Table 3 Comparing the results of the initial test between control and experimental groups

Group	Control	Experimental	Levene test *p* value	*t* test (*p* value)
i.	61.24 (*N* = 30)	62.03 (*N* = 30)	0.432	-0.342 (*p* = 0.221)
ii.	58.96 (*N* = 31)	57.07 (*N* = 31)	0.336	0.954 (*p* = 0.339)
iii.	60.08 (*N* = 31)	60.78 (*N* = 31)	0.479	-0.188 (*p* = 0.785)
iv.	62.11 (*N* = 31)	61.82 (*N* = 31)	0.663	0.449 (*p* = 0.663)

Parametric statistics can now follow. In Table 3, Levene's test indicates that there was no statistical difference in variability between any pairs of CG and EG ($p > .05$). The *t*-tests indicated that there was no statistical difference in pretest scores either within or between the CGs and EGs ($p > .05$).

In Table 4, the EG demonstrates a higher proportion of successfully solved tasks than for the CG. In all four experimental subgroups, only 2.43% of those students completed more than three tasks compared to 24.37% of the CG. A line chart of the results from Table 4 shows the EG and CG differences, with the ascent line for the EG being greater than that of the CG (cf. Figure 1).

Table 4 Posttest results for the experimental and control groups

Group	No. of students		No. of students who successfully solved tasks					*M*	*SD*
			5	4	3	2	1		
Experimental	123	*N*	79	35	6	3	0	4.54	0.95
		%	64.22	28.45	4.87	2.43	0.00		
Control	123	*N*	46	47	15	9	6	3.75	0.93
		%	37.39	38.21	12.19	7.31	4.87		

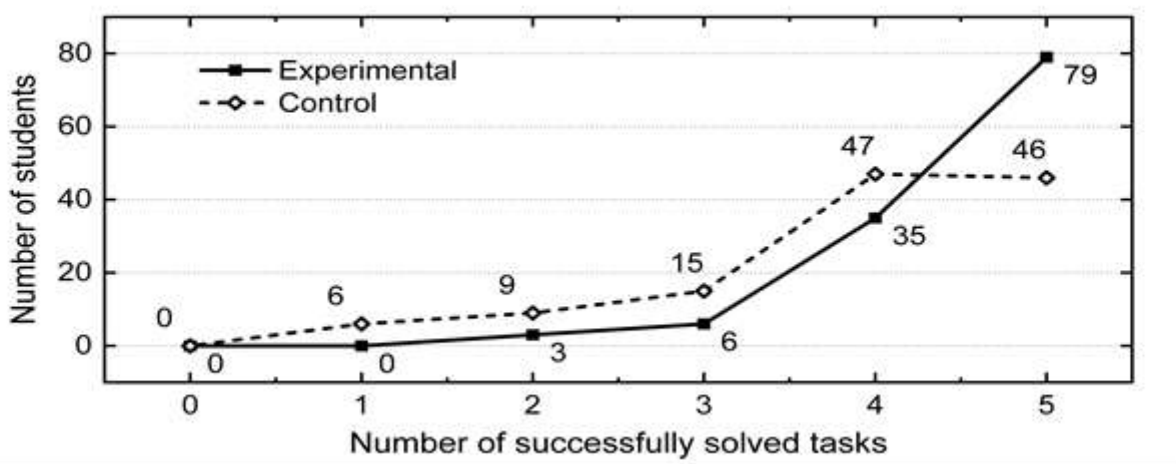

Figure 1 Student trends in the parallel groups

Welch's adjusted test was conducted because Levene's test indicated heterogeneity in-group variance. The degrees of freedom were calculated via the Welch–Satterthwaite equation. The results presented in Table 5 confirm significantly higher scores for the EG across all four schools ($p < .05$). We also calculated size effect, which was large in all four schools.

Table 5 Comparison of the final test results between control and experimental groups

	Control			Experimental			Levene's test		*t*-test for equality of *M*				
School	*N*	*M*	*SD*	*N*	*M*	*SD*	*F*	*p*	*df*	*t*	*p*	Cohen's *d*	Power
i.	30	66.98	4.79	30	71.98	5.22	6.21	0.028	48.06	−3.214	.025	0.99	0.964
ii.	31	61.42	4.94	31	68.78	5.18	5.98	0.039	42.29	−4.251	.038	1.45	0.999
iii.	31	63.25	5.04	31	69.99	5.47	4.98	0.045	56.12	−5.324	.042	1.28	0.998
iv.	31	66.18	4.87	31	70.15	5.15	6.31	0.026	58.69	−3.247	.034	0.80	0.872

The pretest and posttest results are compered in Table 6. Mean scores among the CG were slightly decreased. Dispersion of results among the students in the EG decreased, whereas the CG increased with medium size effect.

Table 6 Comparative results for the initial and final test results for the parallel groups

	Testing				Levene's test		*t*-test				
Group	*I-F	*n*	*M*	*SD*	F	**Sig	*df*	*t*	*p*	Cohen's *d*	Power
Experimental	I	123	3.95	0.94	5.924	0.027	228.5	−5.52	0.01	0.70	0.999
	F	123	4.54	0.72							
Control	I	123	3.96	0.97	1.742	0.479	244	0.0756	0.94		
	F	23	0.95	0.1							

Note. *I = initial; F = final. **Sig = *p* value of Levene's test.

With the comparative effectiveness of the differentiated approach for the teaching of mathematical content, the control, and experimental class observations we also sought to evaluate the activity and engagement levels in the teaching process regarding knowledge processing. As summarised in Table 7, we documented significantly more intense student activity and interest in the EG classes than in the CGs, and none of the students demonstrated low or extremely low activity. Overall, the percentage of students in the EG exhibiting average or lower activity was approximately 8%, compared to 33% of the CG; 18.7% demonstrated low or extremely low activity.

Table 7 Students' in-class activity and engagement levels

Group		Extremely high	High	Average	Low	Extremely low	Total	*p
Experimental	N	95	27	1	0	0	123	$p < 0.001$
	%	77.23	21.95	0.81	0.00	0.00	100	
Control	N	37	45	18	17	6	123	
	%	30.08	36.58	14.63	13.82	4.87	100	
Total	N	132	72	19	17	6	246	
	%	53.65	29.26	7.72	6.91	2.43	100	

Note. *Based on the chi-square test results.

A chi-square test with Yates' correction resulted in a p value of < 0.001. Cramér's V value was 0.72, indicating a strong and significant relationship between groups and student activity. Using R function pwr.chisq.test() we were able to calculate the power of the chi square test. It is not difficult to calculate the sample size as w = 0.31. Therefore, for a sample size = 246 and for a significance level of 5%, the achieved power was 0.89.

Additionally, the observations of the activities of the teachers and students in the CG and EG, the use of teaching materials, and the individualisation development processes were documented (Table 8). The results of the qualitative assessment were obtained by two researchers who applied systemic observation of the classes. Each of us independently completed the observation protocol that consisted of several items describing the activities of the teachers and students in the CG and EG, the use of teaching materials, and the individualisation development processes. We assessed each class activity on the scale from extremely low to extremely high presence of three selected dimensions (activities, teaching materials used, and individualisation). The final assessment (Table 8) was the result of our discussion and agreement on each dimension and its observed presence. The development of individualisation in the EG was significantly greater than in the CG; all of the students in the former group showed high or extremely high individualisation development, whereas 57.7% of the CG exhibited average or lower development in this area, including 23% who demonstrated low or extremely low individualisation. Chi-square testing resulted in a p value near 0. Cramér's V value was 0.68. There appeared to be a significant relationship between individualisation development and the study group.

Table 8 Development of individualisation in the teaching process

Group		Extremely high	High	Average	Low	Extremely low	Total	p
Experimental	N	99	24	0	0	0	123	< 0.001
	%	80.48	19.51	0.00	0.00	0.00	100	
Control	N	25	27	39	19	13	123	
	%	20.32	21.95	31.70	15.44	10.56	100	
Total	N	124	51	39	19	13	246	
	%	50.40	20.73	15.85	7.72	5.28	100	

Table 9 presents a summary of the levels at which students adapted six strategies based on teachers' and researchers' evaluations. The assessments of learning strategies were the result of systemic observations by two researchers and the teachers' self-evaluation. The final assessment (Table 9) was conducted after discussions and agreement among the observers and teachers. The EG demonstrated high use of learning adapted to student ability, assignments with adjusted requirements, tasks with work assistance, and individual work, none of which were used in the CG. The spectrum between *low* and *high* was coded in five intervals. Application of learning strategies measured students' use of six strategies (Table 8) and scored in a reverse manner, i.e., on a range between 1 (not used) and 5 (used to a great extent). Observations of established learning strategies in the CG demonstrated that whole-class instruction and solving tasks with the same requirements were used without any differentiated support. These strategies were minimally used in the EG.

Table 9 Application of learning strategies in parallel groups

Learning strategy	Experimental group	Control group
Teaching adapted to student abilities	4	1
Tasks with adapted assignments	4	1
Tasks with work assistance	5	1
Whole class instruction	2	5
Solving tasks with homogenous requirements without work assistance	1	5
Individual work	4	1

Note. 1, 0%–20% use; 2, 21%–40% use; 3, 41%–60% use; 4, 61%–80% use; 5, over 80% use.

Discussion

The pretest and posttest results show significant differences between the EG and CG. The overall ratio of successfully to unsuccessfully solved tasks demonstrated that the EG achieved a significantly higher mean score than the CG, and the EG made significant progress from the initial to the final testing. By contrast, the CG appeared to stagnate. Therefore, the research hypothesis was accepted. Additionally, based on the observed levels of individualisation, interest, and activities when solving problems in the EG and CG, we agreed that tasks given in the EG stimulated students' interest much more than in the CG.

DI in mathematics teaching increased student activity, interest and achievements, indicating that differentiation successfully catered to students' learning needs and opportunities to progress (Bal, 2016; Gamble, 2011; Tomlinson, 2014). This supports findings that DI can improve mathematics teaching and learning in younger students (Bal, 2016; Prast et al., 2018; Wilson, 2014). The results also support findings that a differentiated approach to develop students' ability to successfully solve algebraic tasks is methodologically justified for fostering achievement (Bal, 2016). Also, using a differentiated approach to teach mathematics to fourth-graders led to enhanced learning outcomes compared to non-differentiated teaching. Observations of class activities in the parallel groups identified six recognisable learning strategies for differentiated versus traditional instruction. The EG had a high degree of applicability for assignments with work support, learning adapted to student ability, assignments with adjusted requirements, and individual work, whereas the CG primarily engaged in a whole-class approach to teaching and solving tasks with homogenous requirements and no working support. These findings further support DI benefits for the successful resolution and enhancement of students' interest in mathematics tasks (Ashley, 2016; Wilson, 2014).

Although we found significant effects of DI on student achievement, we also observed that to achieve this success teachers had to have enhanced training and spend more preparation time on algebraic content. Designing differentiation for the experimental class involved a significant increase in time for teaching preparation and planning, as the EG teachers had to plan for three different instruction levels for each of the 12 sessions (N = 36 or more tasks). Conversely, teachers in the CG only had to prepare one common instruction for each task. Since tailored explanations had to be provided to average and struggling learner groups, EG teachers received additional training. We found their competencies to deliver DI and their knowledge of the learning content were weak (e.g., the gradualness of introducing and understanding equations). In addition, students had to be inspired to participate (Van Geel et al., 2019).

Although teachers in the EG were greatly interested in DI and encouraged by good results, it is uncertain whether they will continue the practice. Several studies found that differentiation was difficult to achieve (Gaitas & Martins, 2017) and teachers were often unsuccessful in its implementation (Van Tassel-Baska & Stambaugh, 2005), possibly due to the extra work involved (De Graaf, Westbroek & Janssen, 2019). This is why most teachers preferred to plan for the average student (Hertberg-Davis, 2009). Although De Graaf et al.'s (2019) investigation of the practical application of DI confirmed our attitudinal findings, DI was positively assessed by both teachers and students.

The degree to which teachers understand their students is important in DI (Tomlinson, 2008). Accordingly, the findings had implications for the ability grouping debate (Boaler & Wiliam, 2001). Although scholars continue to advocate for its abandonment (Boaler, 2013; Francis, Archer, Hodgen, Pepper, Taylor & Travers, 2017; Wiliam & Bartholomew, 2004), DI advocates found that teachers were satisfied with the effectiveness of student support and success when they differentiated mathematics teaching by ability (Ekstam et al., 2015). Steenbergen-Hu, Makel and Olszewski-Kubilius (2016) found that although between-class homogenous groupings were generally disadvantageous for students, the effects of within-class groupings were more positive. The findings indicate that besides within-class ability groupings, the teaching approach may significantly influence students' learning. If students are grouped by ability with no further differentiation, this may harm their self-esteem, particularly based on the reduced achievement levels of lower ability students (Boaler, 2013; Wiliam & Bartholomew, 2004). Within-class groupings and implementation of DI among the EGs suggest potential benefits of more flexible and mixed approaches to ability-grouping (Steenbergen-Hu et al., 2016; Tieso, 2003). One final strength of this study was its ability to observe the teaching process as well as the quantitative indicators of student performance.

This study was significant because the use of DI to teach algebraic content to lower elementary school graders has not previously been explored. There are limitations, such as the thematic scope and study duration. Therefore, future DI research should focus on other mathematics topics over longer periods. Additionally, we did not analyse factors such as students' gender or socioeconomic status, which may impact academic achievement, particularly in ability grouping (Wiliam & Bartholomew, 2004). Finally, future studies should ensure that all teachers are trained to be equally prepared across both groups so that the findings are not adversely impacted. Although no significant differences were observed in teachers' content knowledge during teaching implementation, at least some of the gaps between the EG and CGs' achievement and engagement could be attributed to variations in knowledge.

Conclusion

In this study we evaluated the efficacy of DI among learning groups in Montenegro fourth-graders and demonstrated that the teaching and learning approach optimised students' algebraic capabilities. This resulted in greater success at task resolution, greater engagement and activity, and increased persistence. Teachers have a central role in planning and implementing DI as they monitor students' work and progress. These findings support previous research that encourages more teachers to apply the DI model when teaching mathematics.

Our research also confirmed the need for additional teacher training in implementing differentiation strategies. Teachers must possess sound knowledge of pedagogical content, understand the relationships between tasks, the concepts of gradual learning, and students' cognitive characteristics in order to adequately respond to the learning needs.

Accordingly, we recommend the following:

- A differentiated model should be designed with tasks providing separate instructions tailored for the needs of advanced, in-between, and struggling learners. Teachers must have adequate time to get to know their students and other teaching materials.
- Montenegrin teachers should receive additional DI training. Teachers must become thoroughly familiar with its conceptualisation and application. This teaching model is important because diversity within classrooms should be regarded as an opportunity to assist all student groups.
- Educational decision-makers should carefully plan the introduction and assessment of DI models. Students' mathematical literacy can be successfully enhanced using DI and this should be recognised through curriculum design. It is especially crucial for systems challenged by poor results in PISA and similar tests.

Acknowledgements

We express our appreciation to all research participants. We extend special gratitude to the students who actively participated and to our teacher colleagues, whose ideas significantly improved the research.

Authors' Contributions

This article was written by the first (VM) and the second author (DV) with support from the third (BM), the fourth (NŠ) and the fifth author (TN) who also helped in data collection (designing the study, and implementing the experimental programme). The experimental programme was developed with the participation of each author, based on the ideas of the first author (VM). BM and NŠ were in charge of quantitative data processing, while the fifth author (TN) was in charge of qualitative data processing. Data discussion was created by the second (DV) and the first (VM) author. The percentage of contributions is indicated by the order of authorship provided. The authors read and approved the final manuscript.

Notes

ii. DATES: Received: 6 November 2020; Revised: 15 March 2023; Accepted: 24 June 2023; Published: 31 August 2023.

References

Abah JA 2020. An appeal in the case involving conventional teaching: Emphasizing the transformation to enhanced conventional teaching in mathematics education. *VillageMath Educational Review*, 1(1):1–10. https://doi.org/10.5281/zenodo.3860320

Anthony G & Hunter R 2017. Grouping practices in New Zealand mathematics classrooms: Where are we at and where should we be? *New Zealand Journal of Educational Studies*, 52(1):73–92. https://doi.org/10.1007/s40841-016-0054-z

Ashley LM 2016. Implementation of a math workshop model in the elementary classroom: Understanding how teachers differentiate instruction. PhD thesis. Boston, MA: Northeastern University. Available at https://repository.library.northeastern.edu/files/neu:cj82m2381/fulltext.pdf. Accessed 31 August 2023.

Bal AP 2016. The effect of the differentiated teaching approach in the algebraic learning field on students' academic achievements. *Eurasian Journal of Educational Research*, 63:185–204. Available at https://dergipark.org.tr/en/download/article-file/225130. Accessed 31 August 2023.

Bender WN 2013. *Differentiating math instruction, K-8: Common core mathematics in the 21st century classroom*. Thousand Oaks, CA: Corwin Press.

Boaler J 2002. *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning* (Rev. ed). London, England: Routledge.

Boaler J 2013. Ability and Mathematics: The mindset revolution that is reshaping education. *Forum*, 55(1):143–152. Available at http://www.youcubed.org/wp-content/uploads/14_Boaler_FORUM_55_1_web.pdf. Accessed 31 August 2023.

Boaler J & Wiliam D 2001. 'We've still got to learn!' Students' perspectives on ability grouping and mathematics achievement. In P Gates (ed). *Issues in mathematics teaching*. London, England: Routledge.

Boaler J, Wiliam D & Brown M 2000. Students' experiences of ability grouping—disaffection, polarisation and the construction of failure. *British Educational Research Journal*, 26(5):631–648. https://doi.org/10.1080/713651583

Chamberlin M & Powers A 2010. The promise of differentiated instruction for enhancing the mathematical understandings of college students. *Teaching Mathematics and its Applications: An International Journal of the IMA*, 29(3):113–139. https://doi.org/10.1093/teamat/hrq006

Champely S 2020. *pwr: Basic functions for power analysis. R package version 1.3-0.* Available at https://CRAN.R-project.org/package=pwr. Accessed 31 August 2023.

Cohen J 1988. *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum.

Cowan R & Powell D 2014. The contributions of domain-general and numerical factors to third-grade arithmetic skills and mathematical learning disability. *Journal of Educational Psychology*, 106(1):214–229. https://doi.org/10.1037/a0034097

De Graaf A, Westbroek H & Janssen F 2019. A practical approach to differentiated instruction: How biology teachers redesigned their genetics and ecology lessons. *Journal of Science Teacher Education*, 30(1):6–23. https://doi.org/10.1080/1046560X.2018.1523646

Doubet KJ & Hockett JA 2017. *Differentiation in the elementary grades: Strategies to engage and equip all learners*. Alexandria, VA: ASCD.

Ekstam U, Linnanmäki K & Aunio P 2015. Educational support for low-performing students in mathematics: The three-tier support model in Finnish lower secondary schools. *European Journal of Special Needs Education*, 30(1):75–92. https://doi.org/10.1080/08856257.2014.964578

Francis B, Archer L, Hodgen J, Pepper D, Taylor B & Travers MC 2017. Exploring the relative lack of impact of research on 'ability grouping' in England: A discourse analytic account. *Cambridge Journal of Education*, 47(1):1–17. https://doi.org/10.1080/0305764X.2015.1093095

Fuchs LS, Fuchs D, Compton DL, Powell SR, Seethaler PM, Capizzi AM, Schatschneider C & Fletcher JM 2006. The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, 98(1):29–43. https://doi.org/10.1037/0022-0663.98.1.29

Gaitas S & Martins MA 2017. Teacher perceived difficulty in implementing differentiated instructional strategies in primary school. *International Journal of Inclusive Education*, 21(5):544–556. https://doi.org/10.1080/13603116.2016.1223180

Gamble V 2011. The impact of differentiated versus traditional instruction on math achievement and student attitudes. PhD dissertation. Minneapolis, MN: Walden University. Available at https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=1922&context=dissertations. Accessed 31 August 2023.

George PS 2005. A rationale for differentiating instruction in the regular classroom. *Theory Into Practice*, 44(3):185–193. https://doi.org/10.1207/s15430421tip4403_2

Hackenberg AJ, Creager M & Eker A 2021. Teaching practices for differentiating mathematics instruction for middle school students. *Mathematical Thinking and Learning*, 23(2):95–124. https://doi.org/10.1080/10986065.2020.1731656

Harris P & Johnson R n.d. *Non-traditional teaching and learning strategies*. Available at http://www.montana.edu/facultyexcellence/Papers/activelearn2.html. Accessed 16 May 2019.

Hertberg-Davis H 2009. Myth 7: Differentiation in the regular classroom is equivalent to gifted programs and is sufficient: Classroom teachers have the time, the skill, and the will to differentiate adequately. *Gifted Child Quarterly*, 53(4):251–253. https://doi.org/10.1177/0016986209346927

Kesteloot BA 2011. Effects of differentiated mathematics instruction in a fourth grade classroom. MSc dissertation. Marshall MN: Southwest Minnesota State University. Available at https://www.proquest.com/docview/894122407?pq-origsite=gscholar&fromopenview=true. Accessed 31 August 2023.

Light D & Pierson E 2014. Increasing student engagement in math: The use of Khan Academy in Chilean classrooms. *International Journal of Education and Development using Information and Communication Technology*, 10(2):103–119. Available at https://www.learntechlib.org/p/147457/. Accessed 1 June 2019.

Livers SD, Paxton M, O'Grady N & Tontillo M 2018. Embracing curriculum compacting: Teacher candidates supporting differentiated instruction in elementary mathematics. *School-University Partnerships*, 11(1):19–25. Available at https://files.eric.ed.gov/fulltext/EJ1179969.pdf. Accessed 31 August 2023.

Lou Y, Abrami PC, Spence JC, Poulsen C, Chambers B & D'Apollonia S 1996. Within-class grouping: A meta-analysis. *Review of Educational Research*, 66(4):423–458. https://doi.org/10.2307/1170650

Matthews MS, Ritchotte JA & McBee MT 2013. Effects of schoolwide cluster grouping and within-class ability grouping on elementary school students' academic achievement growth. *High Ability Studies*, 24(2):81–97. https://doi.org/10.1080/13598139.2013.846251

Maxey KS 2013. Differentiated instruction: Effects on primary students' mathematics achievement. PhD dissertation. Scottsdale, AZ: Northcentral University. Available at https://www.proquest.com/docview/1449660560?pq-origsite=gscholar&fromopenview=true. Accessed 31 August 2023.

McKeen H 2019. The impact of grade level flexible grouping on Math Achievement Scores. *Georgia Educational Researcher*, 16(1):48–62. https://doi.org/10.20429/ger.2019.160105

Mejer L, Turchetti P & Gere E 2011. *Trends in European education during the last decade*. Available at https://ec.europa.eu/eurostat/documents/3433488/5579668/KS-SF-11-054-EN.PDF.pdf/a98c93d3-dff5-479e-8ac1-aa15975ecb37?t=1414693417000. Accessed 31 August 2023.

Murtagh F & Legendre P 2014. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification*, 31:274–295. https://doi.org/10.1007/s00357-014-9161-z

Organisation for Economic Co-operation and Development 2018. *PISA 2015: Results in focus*. Paris, France: Author. Available at https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf. Accessed 31 August 2023.

Papanastasiou E 2002. Factors that differentiate mathematics students in Cyprus, Hong Kong, and the USA. *Educational Research and Evaluation*, 8(1):129–146. https://doi.org/10.1076/edre.8.1.129.6919

Pavićević Ž, Vučeljić M, Lalić M, Pavićević I & Kostić A 2017. *Teaching subject mathematics, curriculum for grades I, II, III, IV, V, VI, VII, VIII and IX of elementary school.* Podgorica, Montenegro: Ministry of Education and Science.

Prast EJ, Van de Weijer-Bergsma E, Kroesbergen EH & Van Luit JEH 2018. Differentiated instruction in primary mathematics: Effects of teacher professional development on student achievement. *Learning and Instruction*, 54:22–34. https://doi.org/10.1016/j.learninstruc.2018.01.009

Prica I, Čolić L & Baronijan H 2014. *Studija o ulaganju u rano obrazovanje djece u Crnoj Gori* [A study on investment in early childhood education in Montenegro]. Podgorica, Montenegro: UNICEF. Available at https://www.unicef.org/montenegro/media/2781/file/MNE-media-MNEpublication34.pdf. Accessed 31 August 2023.

Roy A, Guay F & Valois P 2013. Teaching to address diverse learning needs: Development and validation of a Differentiated Instruction Scale. *International Journal of Inclusive Education*, 17(11):1186–1204. https://doi.org/10.1080/13603116.2012.743604

Sammons P 2010. The contribution of mixed methods to recent research on educational effectiveness. In A Tashakkori & C Teddlie (eds). *Sage handbook of mixed methods in social and behavioral research* (2nd ed). Thousand Oaks, CA: Sage.

Sousa DA & Tomlinson CA 2011. *Differentiation and the brain: How neuroscience supports the learner-friendly classroom*. Bloomington, IN: Solution Tree Press.

Steenbergen-Hu S, Makel MC & Olszewski-Kubilius P 2016. What one hundred years of research says about the effects of ability grouping and acceleration on K–12 Students' academic achievement: Findings of two second-order meta-analyses. *Review of Educational Research*, 86(4):849–899. https://doi.org/10.3102/0034654316675417

Tambychika T & Meerah TSM 2010. Students' difficulties in mathematics problem-solving: What do they say? *Procedia - Social and Behavioral Sciences*, 8:142–151. https://doi.org/10.1016/j.sbspro.2010.12.020

Tieso CL 2003. Ability grouping is not just tracking anymore. *Roeper Review*, 26(1):29–36. https://doi.org/10.1080/02783190309554236

Tomlinson CA 2008. The goals of differentiation: Differentiated instruction helps students not only master content, but also form their own identities as learners. *Educational Leadership*, 66(3):1–6.

Tomlinson CA 2014. *The differentiated classroom: Responding to the needs of all learners* (2nd ed). Alexandria, VA: ASCD.

Tomlinson CA & Imbeau MB 2010. *Leading and managing a differentiated classroom*. Alexandria, VA: ASCD.

Tomlinson CA & McTighe J 2006. *Integrating differentiated instruction and understanding by design: Connecting content and kids*. Alexandria, VA: Association for Supervision and Curriculum Development.

Van Geel M, Keuning T, Frèrejean J, Dolmans D, Van Merriënboer J & Visscher AJ 2019. Capturing the complexity of differentiated instruction. *School Effectiveness and School Improvement*, 30(1):51–67. https://doi.org/10.1080/09243453.2018.1539013

Van Tassel-Baska J & Stambaugh T 2005. Challenges and possibilities for serving gifted learners in the regular classroom. *Theory Into Practice,* 44(3):211–217. https://doi.org/10.1207/s15430421tip4403_5

Wiliam D & Bartholomew H 2004. It's not which school but which set you're in that matters: The influence of ability grouping practices on student progress in mathematics. *British Educational Research Journal*, 30(2):279–293. https://doi.org/10.1080/0141192042000195245

Wilson G 2014. Activating assessment for all students: Differentiated instruction and informative methods in math and science. *Mathematics Teacher*, 108(1):76–77.