

Actor and Sapper in the Classroom: A Method of Evaluation and Students' Results

DOI: 10.15804/tner.2022.68.2.15

Abstract

In the presented study, researchers consider if the evaluation method impacts early education students' results and their beliefs about themselves. The results prove significant relation between the method of evaluation and students' results (with the same questions). The reliability of evaluation depends on the type of diagnosed skills (language or mathematical skills). In an oral examination in the presence of the class, children achieve better results in the linguistic test. Higher scores on mathematical tests were noted in the case of written exams. Research conclusions could be used as a hint for early education teachers and an interesting area for other studies and consideration.

Keywords: *early education, examination methods, oral exam, written test*

Introduction

In modern school, the child's activity in situations that diagnose their knowledge and skills is a series of behaviours that can be determined using two analogies: „Student-actor” and „student-sapper”. Both expressions refer to the child's actions, who is assessed in the school reality, and what determines the adopted strategy is an educational situation related to the evaluation. In the case of verbal evaluation, one can notice a certain similarity to a theatrical performance. The student-actor is a child who stands on the blackboard, and the subject of evaluation is the messages they utter to the teacher-viewer in the presence of the class – audience. The student-actor remains in dialogue with the audience. If he is a vigilant observer, he

reads the viewer's reactions and modifies his own game in response to them. He plays a forgotten fragment of a text and guides the audience with phrases created for the script.

The second analogy invoked is the student-sapper. The situation of assessing the written test is compared with the „disarming” of explosives in a minefield. A student-sapper is a child who is taking a written test. In a problematic situation, they are left alone in the face of a problem to be solved. They make decisions without anyone's help, take responsibility for them and have one approach to the task at hand. They will be unable to correct a mistake made, and they may take a risk in situations of ignorance. They take action because failure to act is also against them.

The everyday life of the school environment forces the child to take the role of an actor and a sapper alternately. Written and oral tests are commonly used methods of controlling student achievement. When thinking about a child's functioning at the initial education stage, it is worth considering both of these methods and whether any of them is more reliable in assessing a student's knowledge and skills and how students perceive these methods.

Can Social Exposure Affect Student Performance?

One of the key points to consider is the relationship between students attending the same class in a situation of being evaluated by a teacher. It is worth noting that the mere presence of other people is enough to significantly modify an individual's behaviour and the effect of actions taken by him (Aronson et al., 2010, pp. 323–325). At this point, it is worth mentioning social facilitation, which can be defined as the impact of someone else's physical presence on the performance of an individual, his level, and the result (Geen, 1989, pp. 15–51). The opposite effect of social facilitation is social inhibition, which is the deterioration of the performed task's level due to other people's presence (Turner & Richard, 2014, pp. 151–160). Due to the public nature of the situation, work in front of the „audience”, certain tasks are performed with a much weaker effect than without the audience. Currently, people talk more often about the audience or the coaction effect (Colman, 2015, p. 64), but the phenomenon's essence remains constant.

One of the theories explaining the effect of social facilitation is the concept of fear of a grade¹, which is based on the assumption that it is not the mere presence of the

¹ This theory has been questioned as the only explanation of the facilitation effect, but in the context of the conducted research it seems to be of particular importance.

audience that determines the effect. Physiological agitation is a consequence of the fear of being judged by others who witness the task being performed. In other words, the individual in front of the audience is afraid of being judged when he does not present himself best. The result is agitation – fear of bad grades (Cottrell, 1972, pp. 185–236). Research on social facilitation among children has been conducted since the end of the 20th century. The conclusions from the research seem to be irrefutable: the phenomenon occurs in children's groups, but it is the strongest among adolescents (Baumeister, 1985, pp. 48–52; Aiello & Douthitt, 2001, pp. 163–180).

Another aspect worth paying attention to is the intensification of impression management mechanisms. According to social psychologists, people work most intensively on their image when the focus is on them, when interaction partners can help them achieve their goals, or when they realise that other people in the social situation perceive them differently than they might have wished (Kenrick et al., 2005). These situations are part of everyday functioning in the classroom. A good example that meets all the mentioned criteria is the assessment of students' knowledge and skills, especially in the form of an oral examination.

An oral and written assessment is controversial among researchers and practising teachers in the context of the accuracy and reliability of assessment and the emotions associated with the process itself. Moreover, although in most cases, teachers declare that they care about the accuracy of school evaluation, the results of the research show that when giving the grade, teachers may take into account factors other than the actual achievements of the student (Randall & Engelhard, 2010, pp. 1372–1380).

According to a study conducted in China in 2010, teacher evaluation generates stress and anxiety in up to 80% of early school students in the study groups (Hesketh et al., 2010, pp. 136–140). Emotional arousal of students, referred to as worry or anxiety, also depends on the form of assessment (oral or written), where social exposure or lack of knowledge increase unpleasant emotional experiences (Geen, 1989, pp. 15–51). Among the factors generating stress among students, Fairbrother and Warn (2003, pp. 8–21) mention competition with peers or failures (in the context of exams/tests). Both of these situations may involve oral and written evaluation.

According to the research conducted in 2011, stress and anxiety experienced by students are realistically related not only to their mental well-being but also to their achievements (Agarwal, 2011, pp. 79–87). Therefore, it is not excluded that similar relationships can be found in early childhood education.

So, if different evaluation situations can generate different emotional states, and these are related to students' results, it is worth asking whether the effect of students' work will be different in different assessment situations.

Methods

At the beginning of the research procedure, the following questions were formulated: what is the significance of the situation of social exposure to the student's performance and internal experiences? This issue has become the main research problem in the conducted investigations. The essence of the research was to establish whether the written test enables obtaining the same results as the analogous oral test and, if not – which method is more reliable and gives a complete picture of the students' knowledge and skills.

Therefore, the present research procedure may concentrate on the hypothesis that **the situation of social exposure (the situation of evaluation) is important for the results of the student's work**. Following the adopted hypothesis, we assume that in different situations of evaluating students' work, the effect of this work may differ.

The basic research method was the experiment, the technique of two independent groups without the initial measurement². 70 students from the 2nd grade and 68 from the 3rd grade (N = 138) participated in the experiment, a comparable percentage of boys and girls.

The main selection of the sample was random, and Silesian primary schools were the sampling frame. Due to the preservation of natural conditions for the study participants, the researchers did not interfere with the composition of the class team (the size of individual groups was dictated). One of the 2nd grades was selected randomly as the experimental group (verbally evaluated), and the other as the control group (test writers). The same algorithm was used for 3rd-grade students.

The Experiment Procedure:

In the 2nd and 3rd grades, students from the control groups received a written language and mathematical competence test. The control group testing was carried out through standard school written examination procedures. The students (control groups) were asked to write a test. Everyone worked on their permanent place at the desk, using only the pen and test. The entire class, as usual, was writing the test at the same time, and the teacher informed the group that the work was independent and supervised the test. The students' completed tests were collected by the teacher and handed over to the researchers, who checked all the tests,

² The initial measurement was abandoned to eliminate the learning effect.

determining the number of tasks completed by the students in both the linguistic and math parts.

Students from experimental groups answered the same questions during the oral test on the blackboard in the presence of their peers. The procedure itself was identical to the standard polling on the blackboard. The questions were asked by the teacher (instructed by the researchers), who works in a given class on a daily basis. Students' responses were recorded and written down literally (a record of the child's words). Based on this record, the effects of students' work were assessed, which were compared with the results of written tests in control groups.

What may raise doubts when designing the research procedure is the discrepancy in the level of knowledge between the studied groups (experimental and control). In order to control this issue, a specially chosen tool was used in the study, the questions which are not closely related to the current teaching curriculum in the studied classes, and the results can be related to the developmental norm.

The tool applied was the mathematical and linguistic competence test prepared based on D. Wechsler's Intelligence Scale for children (2003, 4th ed.) – „Dictionary” and „Arithmetic reasoning” tests, respectively. In the mathematical part, the children's task was to perform arithmetic calculations by memory, while in the linguistic part, the respondents had to construct a definition of the indicated words (answer the question „what is ...?”).

The data collected in the experimental part was compiled and compared using appropriate statistical procedures (described below).

The method supporting the experiment was a diagnostic survey (auditorium questionnaire). 100 students from the 2nd and 3rd grades participated in the survey. The questionnaire included questions about the emotions and internal experiences of the students in two different assessment situations (written and oral).

Results

When analysing the survey results, it can be concluded that early childhood education students rarely perceive the two assessment situations (oral and written) as stressful circumstances that generate nervousness or anxiety. The respondents described themselves as people motivated to act, confident, and convinced of their success, regardless of the exam form. Students declare a high level of faith in their abilities and skills. They almost unanimously admit that their motivation and commitment are always high, and during the examination, they give their best. The students did not indicate that any of the evaluation forms was associated with

much stronger anxiety or stress. Although the oral test was associated with greater intensity of unpleasant sensations and difficulty defining and naming one's feelings and emotions, this relationship was not significant.

The results of the experimental part were converted separately for the 2nd and 3rd grades (due to the natural changes in linguistic and mathematical competencies during the entire school year). The linguistic and arithmetic parts are also included. Table 1 shows the mean values of all tested variables for each group.

Table 1. Summary of the mean results for all study groups

Grade	Group	Mean result – mathe- matics	Mean result – language	Mean result – test in total	Mean miss- ing answer – mathe- matics	Mean miss- ing answer – language
II	Experimen- tal (oral exam)	4.43*	9.03*	13.46	0.91*	1.49*
II	Control (written test)	5.19*	7.34*	13.26	1.51*	2.74*
III	Experimen- tal (oral exam)	6.25*	10.18	16.43	0.47	1
III	Control (written test)	7.08*	9.19	16.28	0.64	1.53

* statistically significant ($\alpha=0,05$)

MEAN RESULT – MATHEMATICS – mean result of the math competencies test

MEAN RESULT – LANGUAGE – mean result of the language test

MEAN RESULT – TEST IN TOTAL – mean result of both competency tests

MEAN MISSING ANSWER MATHEMATICS – mean number of unanswered questions in the maths test

MEAN MISSING ANSWER LANGUAGE – mean number of unanswered questions in language test

The significance of differences in the studied groups was verified based on tests of significance of differences for independent groups in the parametric or non-parametric version. The choice of the statistical test was determined by fulfilling the assumptions of the t-test, i.e. the normality of the distribution of the studied variables (the normality of the distribution was checked using the W Shapiro-Wilk Test) and the homogeneity of variance.

In cases where the distribution of the studied variables differed from the normal distribution, the non-parametric version of the t-test was used in subsequent analyses, i.e., the U Mann-Whitney test. In the groups where the assumption of

homogeneity of variance was not met, a revised version of the t-test was used – the t-test with independent variance estimation.

Based on the conducted experiment, it can be assumed that the effectiveness of the assessment method varies depending on the area of the diagnosed competencies (mathematics/language). The test results are presented in Table 1. Students scored higher on the language test when they took it orally in both age groups. In the 2nd grades, the difference turned out to be statistically significant, as shown in Table 2. There was no statistically significant difference in the results of the language skills test in the 3rd grades (Table 3).

Table 2. The result of the t-test for language competencies in 2nd grades (significance of differences between the groups)

Variable	t-test results	p
result – language	-2.07*	0.04

* statistically significant ($\alpha=0,05$)

Source: own research

Table 3. The result of the t-test for language competencies in 3rd grades (significance of differences between the groups)

Variable	t-test result with independent variance estimation	p
result – language	-1.11	0.27

Source: own research

A different result was obtained in the case of mathematics. Students who solved arithmetic problems in the classroom obtained lower results than those who solved them in writing in the classroom. This relationship is statistically significant in all groups, and presented in Tables 4 and 5.

Table 4. The result of the t-test for mathematical competencies in 2nd grades (significance of differences between the groups)

Variable	t-test results	p
result – mathematics	4.22*	0.00

* statistically significant ($\alpha=0,05$)

Source: own research

Table 5. The result of the U Mann-Whitney test for mathematical competencies in 3rd grades (significance of differences between the groups)

Variable	U test result	p
result – mathematics	2.34*	0.02

* statistically significant ($\alpha=0,05$)

Source: own research

During the analysis of the research results, another variable, tentatively defined as “student involvement” (in the table “mean missing answer”), turned out to be significant. In all variants of the research, students had the option not to answer. All groups were informed (during the initial activities) about the possibility of not answering. Students who wrote the test could leave a blank space. Those answering on the blackboard could say “I don’t know” and move on to the next question or end the test.

Table 6. Avoiding responses in all research groups

Grade	Group	Mean missing answer – mathematics	Mean missing answer – language
II	Experimental (oral exam)	0.91*	1.49
II	Control (written test)	1.51*	2.74
III	Experimental (oral exam)	0.47	1
III	Control (written test)	0.64	1.53

* statistically significant ($\alpha=0,05$) Source: own research

MEAN MISSING ANSWER MATHEMATICS

– mean number of unanswered questions in the maths test

MEAN MISSING ANSWER LANGUAGE

– mean number of unanswered questions in language test

In all research groups, students more often gave up answering the question in the case of the written test than the oral test. This relationship is statistically significant in the case of 2nd-grade students. The results of the significance of differences tests for the variable “no response” are included in Table 7.

Table 7. The result of the U Mann-Whitney test for the variable “no response” in all research groups (significance of differences between the compared experimental and control groups)

Grade	Variable	U test result	p
II	no response – mathematics	2.34*	0.01
III	no response – mathematics	0.59	0.55
II	no response – language	2.93*	0.00
III	no response – language	1.07	0.28

* statistically significant ($\alpha=0,05$)

Source: own research

An interesting aspect of children’s answers turned out to be improvising, making up the results: students who answered at the blackboard when they did not know the answer, made it up. They preferred to say anything rather than admit their ignorance. A similar situation practically did not take place in the case of written tests. Students writing the test did not create new solutions – when they did not know the answer to the question, they left a blank space.

From the results, it can be stated that **the evaluation situation (test on the blackboard or a written test) impacted the effects of students’ activities**. Therefore, the true hypothesis adopted in the research is that the situation of social exposure (the situation of evaluation) is vital for the results of student work. In different situations of evaluating students’ work, the effect of this work may differ. The present research procedure showed that students achieve a lower score in mathematical skills in the situation of social exposure and a higher score in language skills. Regardless of the studied area, children more often give up answering a question when they write a test at the desk.

Discussion

Research on oral and written evaluation effectiveness was conducted in various age groups. The reports from the 2012 research (Huxham et al., 2012, pp. 125–136) on a sample of 128 students seem to be significant. The subjects were divided into two groups. The first was tested in biology (oral and written), and the second took the test for “scientific” and personal development questions. All study groups fared better in verbal responses. The researchers emphasised an additional factor was the examiner’s “professionalism”. The respondents noticed great potential in an

oral examination, even though it seemed more stressful to them, how the question is framed, and if additional leading questions may help build the oral answer. The aspect of learning and analysing the material during the interview with the examiner is also essential.

Thorburn and Collins (2006) also emphasised the importance of verbal assessment. Their research indicates that oral interviews could provide a constructive atmosphere for ongoing assessment and informative feedback for teachers to utilise as part of their reflections on practice.

When discussing the obtained results, it is worth paying attention to the specificity of the studied age group. One should consider whether the essence of the described variability in the results of the child's work was not the difficulty in writing down one's thoughts, characteristic for the age. Perhaps it was not the social exposure that modified the effect of the student's work, but the priority of the activity of speaking over the activity of writing? Not quite. Apart from the content of information in the child's statement (which was evaluated), other properties of the statement should also be taken into account: coherence, clarity, and correctness of the wording. These fared better on the written responses. Thus, depending on what becomes the subject of evaluation (information capacity of statements or their precision and formulation), one or the other method of checking students' competencies can be used. In the case of the mathematical competence test, it can be stated that writing down the result of an arithmetic operation was not a new and challenging activity for students (in terms of plotting numbers), so it did not translate into the test result. In this aspect, the difference between speaking and writing will therefore not be taken into account.

A very important aspect of the conducted research, which is its additional value, is paying attention to avoiding answers by students. It is a situation significant for the description of the effect of the student's work in different evaluation situations. The students who answered the blackboard did not allow themselves to give up. They tried to create their own ideas if they did not know the answer, just like an improvising actor. Due to this dependence, an oral examination may be crucial for discovering students' potential and their involvement in the search for new solutions. It should be considered that problem-solving skills and creative thinking are often more important than recalling previously accumulated conceptual knowledge. After all, education is more about developing thinking, not remembering facts. The quality of students' ideas when they do not know the correct answer to a question could be the subject of other significant research.

Not only problem-solving but also understanding the issue turns out to be better during oral exams. The research conducted by Rushton and Eggett (2003) showed

that oral examinations can effectively evaluate the student's comprehension and application of information in an important situation. Nursing students gained higher scores in the oral examination than written tests (oral examination can be as effective or more effective in evaluating student understanding of medical/surgical content and its application in clinical situations).

Conclusion

Based on the conducted research, it is possible to formulate several indications for pedagogical practice. The most important of them is the statement that in the educational environment, it is impossible to limit ourselves to one preferred method of working with a child, also in checking and evaluating students' knowledge and skills. What works for the evaluation of mathematical competencies may be less effective for language skills and vice versa. According to the presented research results, one of the significant factors that should influence the choice of the method of assessing students' work is the subject of evaluation (type of competence, the scope of material). Therefore, it is impossible to resign from any of the evaluation situations described in the introduction. So, sometimes the classroom must become a stage where the student-actor will appear, answering at the blackboard, to become a "minefield", where the student-sapper will "disarm" the written test questions. A good teacher is the one who creates opportunities for both a talented actor and an efficient sapper to prove himself. He can look at the effects of their actions through the prism of difficulties related to a specific evaluation situation and (and perhaps above all) individual talents, limitations, preferences, and progress of his actors and sappers. This teacher is aware of the phenomena and knows every student and the class as a group.

When formulating research conclusions, it is worth considering two aspects. The first one concerns the correctness of taking up similar issues in subsequent studies. The presented topic is essential as an explanation and detailing of certain phenomena and as a real problem in the school environment and everyday lives of students and teachers. Evaluation is an integral part of the teaching process, so understanding its regularities and limitations benefits researchers and practitioners alike.

The second thought concerns the students. In the self-characteristics, the respondents described themselves as active, committed, willing to work, open to action, certain of their abilities and convinced of their success, motivated but not stressed. This attitude towards oneself and one's own learning becomes a great

treasure in the hands of teachers. They are faced with a very important task so that a self-confident, motivated child becomes a fulfilled, successful adult.

Acknowledgements

The paper's authors would like to thank the students and teachers of the Tysiąclecie Państwa Polskiego No. 7 primary school in Czeladź for enabling them to conduct research.

References:

- Agarwal, A. (2011). Impact of Academic Stress upon Academic Achievement and Mental Health of the Adolescents. *International Journal of Management and Social Sciences*, 1(1), 79–87.
- Aiello J. R., & Douthitt E. A. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice*, 5(3), 163–180.
- Aronson, E., Wilson, T. D., & Akert, R. M. (2010). *Social Psychology* (4th ed.). Pearson Prentice Hall.
- Baumeister, R. F. (1985). The championship choke. *Psychology Today*, 19(4), 48–52.
- Colman, A. M. (2015). *A Dictionary of Psychology* (4th ed.). Oxford University Press.
- Cottrell, N. B. (1972). Social facilitation. In C. G. McClintock (Ed.), *Experimental social psychology* (pp. 185–236). New York.
- Fairbrother, K., & Warn, J. (2003). Workplace Dimensions, Stress and Job Satisfaction. *Journal of Managerial Psychology*, 18(1), 8–21. DOI: 10.1108/02683940310459565
- Geen, R. G. (1989). Alternative conceptions of social facilitation. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed.) (pp. 15–51). Lawrence Erlbaum Association.
- Hesketh, T., Zhen, Y., Lu, L., Zhou, X. D., Ye, X. J., & Zhu, W. X. (2010). Stress and psychosomatic symptoms in Chinese school children: cross-sectional survey. *Archives of Disease in Childhood*, 95, 136–140. DOI: 10.1136/adc.2009.171660
- Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: a test of student performance and attitudes. *Assessment & Evaluation in Higher Education*, 37(1), 125–136. DOI: 10.1080/02602938.2010.515012
- Kenrick, D., Neuberg, S., & Cialdini, R. (2005). *Social psychology: Unraveling the mystery* (3rd ed.). Allyn and Bacon.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372–1380. DOI: 10.1016/j.tate.2010.03.008
- Richard, J. C., & Turner, R. (2014). *Essential Social Psychology* (3rd ed.). SAGE Publications.
- Rushton, P., & Eggett, D. L. (2003). Comparison of written and oral examinations in a baccalaureate medical-surgical nursing course. *Journal of professional nursing: Official Journal of the American Association of Colleges of Nursing*, 19(3), 142–148. DOI: 10.1016/S8755-7223(03)00049-8
- Sparfeldt, J. R., Rost, D. H., Baumeister, U. M., & Christ, O. (2013). Test anxiety in written

and oral examinations. *Learning and Individual Differences*, 24, 198–203. DOI: 10.1016/j.lindif.2012.12.010

Thorburn, M., & Collins, D. (2006). Accuracy and authenticity of oral and written assessments in high-stakes school examinations. *The Curriculum Journal*, 17(25), 3. DOI: 10.1080/09585170600682491