



HRADEC KRÁLOVÉ JOURNAL OF ANGLOPHONE STUDIES

Published:

**Department of English Language and Literature
Faculty of Education
University of Hradec Králové
Rokitanského 62
500 03 Hradec Králové
Czech Republic**

ISSN: 2336-3347 (Print)

ISSN: 2571-032X (Online)

Registration number: MK ČR E 23933

Vol. 8

No. 1-2

2021

Web: <http://pdf.uhk.cz/hkjas/>

Volume's editor: Jan Suk

Original illustrations: Ivan Mečl

Production: Pavel Mervart Publishing

The publication of this double volume was generously supported and completed within the Specific Research Project 15/I 21, "Phenomenological Perception of Contemporary Anglophone Theatres," funded by the Faculty of Education, University of Hradec Králové

HRADEC KRALOVÉ JOURNAL OF ANGLOPHONE STUDIES

Editorial Board

Bohuslav Mánek, University of Hradec Kralové, Czech Republic

Helena Polehlová, University of Hradec Kralové, Czech Republic

Františka Schormová, University of Hradec Kralové, Czech Republic

Editor in Chief

Jan Suk, University of Hradec Kralové, Czech Republic

Advisory Board

Blanka Babická, Palacký University Olomouc, Czech Republic

Šárka Bubíková, University of Pardubice, Czech Republic

Richard Burt, University of Florida, Gainesville, USA

Petr Chalupský, Charles University, Prague, Czech Republic

Yilin Chen, Providence University, Taichung, Taiwan

Jan Comorek, Charles University, Hradec Králové, Czech Republic

James Clubb, University of Hradec Kralové, Czech Republic

Milada Franková, Masaryk University, Brno, Czech Republic

Jana Hart'anská, Constantine the Philosopher University, Nitra, Slovakia

Jakub Helvich, University of Hradec Kralové, Czech Republic

Kacie Hittel Tam, University of Georgia, Athens, USA

Mirka Horová, Charles University, Prague, Czech Republic

Vladimíra Ježdíková, University of Hradec Králové, Czech Republic

Emá Jelínková, Palacký University Olomouc, Czech Republic

Blanka Klímová, University of Hradec Králové, Czech Republic

Şevki Kömür, Muğla Sıtkı Koçman Üniversitesi, Kötekli/Muğla, Turkey

Ela Krejčová, Akcent College, Prague, Czech Republic

Stanislav Kolář, University of Ostrava, Czech Republic

Filip Krajník, Masaryk University, Brno, Czech Republic

Ivan Lacko, Comenius University, Bratislava, Slovak Republic

Marcela Malá, Technical University of Liberec, Czech Republic

Michaela Marková, Technical University Liberec, Czech Republic

Ildiko Neméthová, University of Economics in Bratislava, Slovak Republic

Ryuta Minami, Shirayuri College, Tokyo, Japan

Atilla Pató, Charles University, Prague, Czech Republic

Hana Pavelková, Czech Technical University in Prague, Czech Republic

Jozef Pecina, Comenius University, Bratislava, Slovak Republic

Ondřej Pilný, Charles University, Prague, Czech Republic

Libor Práger, Palacký University, Olomouc, Czech Republic

Štěpánka Rubešová, University of Hradec Kralové, Czech Republic

David Ryška, University of Hradec Kralové, Czech Republic

Václav Řeřicha, Palacký University, Olomouc, Czech Republic
Vladislav Smolka, University of South Bohemia, České Budějovice, Czech Republic
Alice Sukdolová, University of South Bohemia, České Budějovice, Czech Republic
Martin Štefl, Czech Technical University in Prague, Czech Republic
Věra Tauchmanová, University of Hradec Králové, Czech Republic
Ondřej Tichý, Charles University Prague, Czech Republic
Ladislav Vít, University of Pardubice, Czech Republic
Michaela Weiss, Silesian University in Opava, Czech Republic
Yukari Yoshihara, University of Tsukuba, Ibaraki, Japan
Jakub Ženíšek, Charles University, Prague, Czech Republic

Hradec Králové Journal of Anglophone Studies

CONTENTS

Jan Suk Things do not change; we change: Introduction to <i>Hradec Králové Journal of Anglophone Studies</i>	8
--	---

LINGUISTICS AND METHODOLOGY

Markéta Blažková The Distribution of Semantic Roles inside the Nominalized Projection	14
Helena Zitková Assessment and Grading - Perception of Pre-Service English Teachers	31

LITERARY AND CULTURAL STUDIES

Ahlem Fidouh, Belkacem Belmekki Purpose-Built or Accidental? Exploring the Plausibility of John Seeley's Stance Regarding the Birth of the British Empire	56
Markéta Dančová Quality Flouting as a Humorous Tool in the sitcom <i>The Big Bang Theory</i>	70
Filip Krajník, Michaela Weiss "Their hour will be his hour": Margaret Atwood's <i>Hag-Seed</i> and the Conventions of Renaissance Revenge Plays	80

REVIEWS

David Livingstone Review of Howard Jacobson's Novels in the Context of Contemporary British Jewish Literature by Petr Anténe	90
Calls	96
Notes on Contributors	97
Ethical Statement	101
Mission Statement and Guidelines for Submissions	103

Assessment and Grading - Perception of Pre-Service English Teachers

Abstract: This article aims to explore test evaluation through the eyes of future English teachers. Firstly, the topic is linked to the wider context of modern approaches to future teacher education and school evaluation and then narrowed to test evaluation itself. Secondly, a survey, whose participants were second-year undergraduate prospective teachers, is introduced. As a part of developing their assessment literacy, they were asked to assess and grade a progress test, discuss it in the class and then reflect on the experience. The analysis of the data from the respondents' handouts and open-ended questionnaire items followed a constant comparative analysis and the result is an empirically grounded theoretical framework on pre-service teacher perception of assessing and grading. The findings show that the investigated pre-service teachers find evaluation a complex and difficult process and based on their experience of assessing and grading the identical test, they feel the urge to reconsider their current perception.

Introduction

School assessment is an integral part of teaching and learning. Traditionally, assessment and grading were intertwined and inseparable parts of school practice. In recent years, however, some changes in this perception have been noticeable. With the awareness of the importance of developing assessment literacy in pre-service teachers and the need for research on pre-service teachers in the field of school assessment (Brown and Remesal 13-14), a survey of future English teachers in their second year of bachelor study program was conducted. To broaden their horizons in assessing, the pre-service teachers were asked to assess an identical test and then reflect on the experience. Since emotions and emotional states play an important role in teacher preparation and are core to reflexive processes (Holmes 147, Shoffner 783), the feelings and emotions of the pre-service teachers will be explored in the study as well. Accordingly, the aim of this paper is to first find out how the student teachers will evaluate one identical test and, secondly, to get an insight into how future teachers perceive assessment and how they think about it.

Stepping outside one's comfort zone in reflection is an inseparable part of the process of ongoing learning and plays a significant role in professional development of pre-service teachers. Through reflection, the prospective teachers are encouraged to mindfully mirror their experience, become aware of the implicit influences that affect their perceptions and make them explicit (Calderhead and Gates 3), and to "consciously choose what kind of teacher they want to be" (Korthagen 84). Thence, they take responsibility for their own professional growth, which may empower them to influence future directions in education (Calderhead and Gates 3).

In the present paper, the terms pre-service teachers, prospective teachers, student teachers and future teachers will be used interchangeably. The same applies to the terms referring to future teachers education, i.e. teacher training, future teacher training and teacher education will be used as synonyms. In the context of education, some authors distinguish between the terms "assessment" and "evaluation" (Astin and Antonio, 2012), or sometimes the terms "assessment" and "evaluation" are used together as one term "assessment and evaluation" (Gullan, D.F., 2005). However, these terms will be used interchangeably

in correspondence with, for example, Ur's (1991) or Harmer's (2007) notion within the scope of this paper. Finally, to prevent misunderstanding, the term "grade", and not "mark", will be used to refer to grades (grade '1' through '5') throughout the whole paper.

Teacher Education and the Role of Reflection

In teacher education, reflection is considered one of the key factors of professional growth. The ability to reflect on previous experience and the active role of pre-service teachers in the process of reflection are emphasized to help them become reflective practitioners (see Schön, 1983) or, in other words, grow into teachers who develop their teaching process through reflective practice and are able to consciously structure situations in the classroom and thus learn from their teaching experience in a mindful and systematic manner (DeLuca et al. 22, Spilková and Tomková 11, Korthagen et al. 138-139, Janík et al. 144).

In their reflective practice, teachers have a chance to stop and look back at the situations in the classroom, and give a deeper thought to their teaching in the broader contexts of their professional values, attitudes and intentions (Spilková and Tomková 12). Then, the process of reflection enables them to analyze, discuss, critically examine or reconsider their beliefs about good teaching (Calderhead and Gates 3). Furthermore, when reflecting, teachers may better understand their feelings and emotions related to their practice (Spilková and Tomková 12). Demetriou and Wilson are convinced that "without reflection, teachers cannot modify their practice in a controlled or deliberate way" (939). For them, reflection is "a cognitive process that helps teachers to rethink their practice, learn from their experiences and help them to cope with similar situations in the future" (938-939). Likewise, according to Zembylas (210), reflection improves the quality and effectiveness of teaching and, similarly, Korthagen et al. state that reflective practitioners are "better teachers" (149). They see a direct link between reflection and teacher behavior since "human behavior is based on mental structures which are not static, but at least in part are created or changed through experiences or confrontations with situations [...] and their reflection" (Korthagen et al. 68, 71).

In addition, the focus on the role of feelings and emotions in the process of reflection has become the centre of interest in recent years. It is argued that reflection and emotions are closely intertwined and that emotional states related to professional practice contribute to building teacher attitudes and identities (Mackenzie 186, Zembylas 210-211). Thus, the role of reflection in teacher education is crucial. If exposed to reflection in their teacher training, the pre-service teachers will get used to the process, will be aware of its importance and can become reflective practitioners.

Assessment in Education

Assessment is a natural and inseparable part of learning, thus it plays an important role in school education. Traditionally, school assessment was "used as a measure of school progress and teacher effectiveness within the accountability context of education" (DeLuca et al. 9) so the purpose and function of assessment was to grade and communicate learners' performance and achievements (Brunker 91, Spilková and Tomková 21). This approach to assessment of learning (also called summative assessment) is retrospective and refers to a concluding evaluation usually consisting of scores or grades that summarize the outcomes of learning (Green 1). However, in terms of improvement, "summative assessment does not provide opportunities to extend or enhance understanding, and is not the best representation of overall learning" (Crockett and Churches 11).

A Shift from Assessment of Learning to Assessment for Learning

In modern approaches to learning and education, a shift from focus on assessment of learning based on grades to the assessment for learning (also called formative assessment) based on meaningful and high-quality feedback may be observed. In the context of Czech school system, the principal curricular documents require “changes in the assessment of the pupils towards diagnostics on an ongoing basis, [...] and a wider use of verbal assessment”, and demand assessment built on “evaluation of the pupil’s individual development and on the capacity to make finely-tuned, positive evaluative judgements” (FEP BE 8, 10). Hence, not only is summative assessment not in correspondence with current curricular documents, but also, many scholars and educational experts recommend using formative type of assessment over the summative one to promote learning and to help learners improve (see e.g. Crockett and Churches 11, William and Leahy 5, Macpherson and Hendrick 24, Volante and Fazio 751–752, Spilková et al. 12–13, or Košťálová et al. 15). It is believed that adapting teaching to fit individual learner’s needs, providing meaningful feedback during the process of learning and involving learners into the process of assessment in a form of peer feedback or self-assessment (FEP BE 7–10, Spilková and Tomková 11, Crockett and Churches 9, William and Leahy 135) helps develop learners’ abilities and leads to personal growth (Green 1). In other words, formative assessment serving as a tool for improvement in learning is accentuated nowadays.

Furthermore, criterion-based assessment is, by some authors, incorporated into the learning oriented assessment (Brunker et al. 89) as it may support the process of learning, but only if it is realized through high-quality feedback (see e.g. William and Leahy 11). Yet, if realized by giving grades “without the opportunity to discuss, revise, and improve”, the criterion-based assessment just becomes the assessment of learning, not for learning (Crockett and Churches 11). It means that a grade, though based on clear criteria, still does not provide quality feedback for the learning process.

In Czech schools, however, summative assessment in the form of grades continues to dominate (ČŠI 113), whereas, from the learner’s point of view, getting a simple grade without having the space to think about one’s learning and its outcomes, and discuss them either with the teacher or peers, without the chance to find supporting evidence for one’s answers, analyze mistakes, and come up with the ideas to correct them, makes the assessment useless. As Crockett and Churches point out (2-3) “a number... means little and accomplishes even less”.

Nevertheless, although teachers need to adopt and implement the improvement-oriented assessment, the purpose of summative evaluation must be accomplished as well because the final results are an integral part of a learner’s academic record (DeLuca et al. 9). In the Czech Republic, final summative evaluation is even prescribed by law. The school reports that learners get after each semester are legally binding documents, where the evaluation of the outcomes can be expressed either by a grade (on a scale from 1 to 5), a form of verbal assessment, or by a combination of both (školský zákon § 51, vyhláška § 15).

Assessment criteria and subjectivity in school assessment

To meet the legal requirements, every school has its own classification set of criteria that are to be followed by teachers when assessing. However, they are formulated too generally to fit the widest possible content of several school subjects and so vaguely that everyone can perceive them in their own way as shown in the example of formulated assessment criteria for foreign languages for grade ‘3’:

The learner shows shortcomings in performing the required intellectual and motor activities. The learner makes mistakes when applying the acquired knowledge and skills. The learner's thinking is quite correct, but not very creative; there are mistakes in his logic. The written outcomes are less aesthetic and show some imperfections. (Classification set of criteria of a randomly selected primary school in the CR)

It is not possible for all teachers to perceive these criteria in the same way. Everyone has a different aesthetic feeling, everyone judges creativity in thinking and imperfections differently, and everyone can have different requirements for written and graphic outcomes. Hence, every teacher can manipulate the criteria freely, which results in low objectivity of the assessment. The teacher's perception of the meaning of the criteria and the 'value' of grades is then just a matter of the individual and personal point of view, consequently assessment is "intuitive and subjective" (Slavík 62), and therefore, the grade given to the learner depends solely on the teacher. The causes of low objectivity of evaluation in schools were defined by Slavík (62) who refers to the values of the teachers that can be considered relative and subjective, unreliability in data collection, unintentional or even intentional inaccuracy of the assessors, or the inappropriate and unsuitable criteria selected for assessing.

To conclude, the same learner's performance may be assessed differently by different teachers because different teachers will have a differing conception of assessment. If the exam or the test is not standardized, which is not common in the daily practice in schools (Slavík 62), the sets of criteria, distribution of points in assessment scales, and grading scales can be perceived and operated diversely. When assessing, different teachers may take different aspects of the subject matter into consideration, they may have different sets of criteria and different grading scales. Then, the evaluation procedures depend only on the decision and subjective perception of each teacher.

Developing Assessment Literacy of Pre-Service Teachers

Learning to assess (assessment literacy) is one of the competencies pre-service teachers need to develop in their teacher training. Assessment literacy is defined as "the knowledge of how to assess what students know and can do, interpret the results from these assessments, and apply these results to improve student learning and program effectiveness" (Webb 1). DeLuca, et al. (22) recognise the importance of teacher education as "primary sites for developing teachers' initial beliefs about assessment" (22).

It is believed that the initial professional conceptions and knowledge of prospective teachers arise primarily from their experiences as students (Pajares in Brown and Remesal 3, Harrison 256), however, it is likely that their perceptions, attitudes, beliefs and behavior may change in teaching practice (Brown and Remesal 12). In the context of teacher education, to develop pre-service teachers' assessment literacy, they need to be exposed to topics and course work focused on assessment and learn through practice-based assignments (Kahl et al. 3). It means that pre-service teachers need to be provided with real-life opportunities to challenge their conceptions and to enable them to apply what they have learned (ibid.).

Moreover, since the actual behavior can be influenced by reflection (Korthagen et al. 68, 71), pre-service teachers may start questioning their initial attitudes and beliefs about teaching and learning and reconsider them through involvement in assessing experiences and their reflection. Price et al. assume that "strategies that require students to actively engage with assessment will result in deeper and longer-term development of assessment literacy" (10).

In agreement with the theoretical background, the following methodological part of this paper presents a research focused on developing pre-service teachers' assessment literacy through real-world based assignment and its reflection.

Methodology

The present study is based on exploring pre-service teachers' sole experience of assessing and grading a progress test (Harmer 380) and investigating their feelings, emotions and reasoning related to it. Hence, the aim of the study is to examine the actual perception of pre-service teachers of assessment and grading.

The following research questions were formulated accordingly:

1. How do the future English teachers assess and grade the identical test?
2. How do the future English teachers perceive the process of assessing and grading a test?

To get a profound insight into the perceptions of pre-service teachers a qualitative research design, specifically the grounded theory approach, was applied.

Context of the study and its participants

The survey was conducted in November 2020. The respondents in this study were students in the second year of bachelor's study program English for Education preparing them to become English teachers. None of them had any school teaching experience, therefore no practice with evaluating and grading tests. The group of respondents was ethnically and racially homogenous since they all were of the Czech origin. In total, 19 students in their third semester of teacher training voluntarily participated in the study.

In their prior semesters, the investigated pre-service candidates had completed several courses where they came across the topic of assessment and its implications for learning. Specifically, all the participants of the study took introductory courses of Psychology and Educational Psychology, which are offered in the first and second semester of the teacher education program respectively. The syllabi of these courses cover the topic of learning and teaching, intrinsic and extrinsic motivation, and include grading as a learning topic as well. In the first year of their study program, all the participants were also introduced to Education Sciences with a special focus on modern approaches to teaching and learning as well as the issue of rewards and punishments and their influence on learning. At the end of this two-term course, they were asked to elaborate on the possibilities of assessment in schooling contexts in their final critical essay the topic of which was "The ideal assessment at school".

All the mentioned courses in the first year of future teacher training aim to intentionally challenge the student teachers reasoning about the purpose of school assessment and shift their perspective towards improvement oriented assessment, that is to consider assessment as an effective tool for learning (Crockett and Churches 20, William and Leahy 9, Košťálová et al. 14, FEP BE 8). To develop assessment literacy of pre-service teachers even more, the second-year students enroll in an ELT methodology course that consists, next to the others, of the topics related to assessment, such as learner differences, evaluation and assessment and giving feedback. Since the participants of the study were recently students, they may still be influenced by their previous experience as learners. Nevertheless, there is an assumption that during their teacher education, the prospective teachers will get out of their comfort zone and develop their own assessment philosophy within their evolving teacher identity based on professional knowledge and teaching practice (Calderhead and Gates 3, Brown and Remesal 12, Korthagen et al. 68). Thus, in compliance with modern approaches to future teacher education (see

e.g. Schön, Spilková and Tomková, Korthagen et al.), not only are they provided with the theoretical input to construct their knowledge but they also build their beliefs and attitudes towards assessment through practical tasks and self-reflection.

Further information about the participants, such as their age or gender are considered irrelevant for this study.

Data Collection

The pre-service teachers were surveyed within the ELT methodology course in the third semester of their teacher training. The survey consisted of two phases.

The first phase took place in class as a part of the regular classroom instruction and was focused on experiential test evaluating and grading. All 30 students who signed up for the course were present and actively participated. Each student teacher was given a handout with the assigned task together with a sample of a short completed progress test. All students got the same handout with the same test and were asked to assess it.

The progress test (inspired by Ur 251) to be assessed and graded looked like this:

Test on vocabulary and relative clauses, written during the lesson.

Define the following words, using *who / which / that / whose / when / where*.

For example: a deserted house = a house where nobody lives

- 1 courage: a man who not have any fear
- 2 an illusion: a false sight.
- 3 sweat: its like terrible but more than this
- 4 a virus: a thing which make people sick
- 5 a paw: a part of a animal
- 6 a temple: a house where religious people lives in

The task for the students was formulated as follows: There is a revision test that was completed by a ninth-grader (approximately A2 level) in front of you. It is aimed at determining the knowledge of the meaning of particular words and the ability to use relative pronouns correctly in the definition of the words. Your task is

1. to create a set of criteria for assessing the test
2. to determine the number of points for fulfilling each criterion
3. to design a grading scale
4. to assess the test and grade it (by grade from 1 to 5)

The sets of criteria with corresponding points, the grading scales with given points and grades and the whole class discussion about them were the basis for subsequent reflection.

In the collected handouts, some attempts to describe and explain the decision-making process behind awarding the points were indicated, but not all the respondents included them. Therefore these data were not considered relevant for the study and were omitted from the analysis.

In the second phase, the participants were asked to complete a self-reflective questionnaire. Notwithstanding that the responses in a questionnaire are subjective and the researcher cannot ask any additional questions, still this research tool has a benefit in addressing a large number of respondents and gaining a substantial amount of data in a short period of time (Chráska 158). The questionnaire was administered online after the class, after the first phase of the survey. It was sent to all 30 students enrolled in the course and present in the class, and only those who were willing to, filled it in.

The questionnaire was anonymous but each completed questionnaire was encrypted with a secret code by the respondent, which made it possible to match it with the lists of criteria and grading scales accordingly. Predominantly, it comprised a series of open-ended questions through which the student teachers reflected on the experience they had in the process of evaluating and grading the test in the class. Only one question was closed but served only as an opening to the topic and was followed by an open-ended question to extend the space for reflection. Since open-ended questions promote the motivation to reveal respondents' ideas freely in their own words (Züll 2), they allow them to explore their actual feelings, emotions, thoughts and opinions.

The respondents were asked the following questions:

What feelings and emotions did I have when evaluating and grading the test?

What influenced me when evaluating and grading the test?

How did I feel and what did I think when the criteria, points, grading scales, grades and individual test questions were analyzed in the class?

During or after the analysis in the class, did I feel that I should evaluate the test differently? And if so, how and why?

What did I realize about assessment?

The reflective questions were designed to evoke students' retrospective feelings, emotions and thoughts both about their recent experience with the process of assessing and grading the test and the subsequent discussion about it in the class. On top of that, the set of questions had a potential to induce and influence the students' prospective standpoints and attitudes towards assessment and grading. Such a strategy corresponds with a modern paradigm in teacher education, i.e. the importance of reflection in constructing and reconstructing mental structures and attitudes of pre-service teachers (see Korthagen et al. 2011).

The questionnaire did not include questions regarding the respondents' age, sex or gender, as they were not considered relevant for the study.

Data Analysis

In initial purposive sampling (Chun Tie et al. 3), the materials from respondents (handouts and questionnaires) were collected. In the first phase of the survey, all 30 students present in the class were willing to share their handouts with a set of criteria, a final number of given points and grading scales with the final grade. In the second phase, only 19 respondents out of 30 completed the self-reflective questionnaire. Subsequently, the materials were sorted out by the encryption in order to gain the ultimate sample of data. It means that each handout with a set of criteria and a grading scale was put

together with the fitting questionnaire according to the same secret code. The handouts that did not fit any questionnaire secret code were not used for further analysis. Hence, the data from 19 respondents were the subject of the successive analysis.

After the initial purposive sampling, first the data from handouts were analyzed and described using frequencies and percentages to get an answer for the first research question. Then, the constant comparative analysis based on the principle of repeated comparing (Švaříček et al. 207; Chun Tie et al. 5) was used to find similarities and differences in the data gained from the questionnaires using the open, axial and selective coding within the applied grounded theory approach (Corbin and Strauss 373) to answer the second research question.

In the open coding, the responses were systematically analyzed and coded with the aim to develop thematically relevant categories. The codes in a form of short labels were assigned directly in the margins of each questionnaire, comparing incident to incident. This initial analysis allowed me to detect the thematic similarities across the student teachers' responses. In the advanced axial coding, the categories were interconnected and built into new meaningful groups through exploring the relationships between them (Corbin and Strauss 373). Then, the core categories were identified to form a structure for building a theoretical explanation in the selective coding phase (Corbin and Strauss 15).

Findings and Discussion

In this part of the paper, the findings from the analysis of the data gained from both phases of the survey will be presented and discussed.

Findings from the first survey phase

Sets of Criteria and Corresponding Points

In respondents' handouts, various combinations of the following criteria appeared: "a correct relative pronoun used in the definition", "a meaningful definition", "an accurate and meaningful definition", "grammatical accuracy" and "correct spelling". Each criterion was allocated either a half point or one point.

All sets of criteria designed by the respondents and allocated number of points are presented in the table below:

	Sets of criteria	Points / criterion	Points/ test item	Maximum test points
1.	o correct relative pronoun	0.5	1	6
	o accurate and meaningful definition	0.5		
2.	o correct relative pronoun	1	2	12
	o accurate and meaningful definition	1		
3.	o meaningful definition	1	2	
	o correct relative pronoun	0.5		
	o grammatical accuracy	0.5		
4.	o correct relative pronoun	1	2.5	15
	o meaningful definition	1		
	o grammatical accuracy and correct spelling	0.5		
5.	o correct relative pronoun	1	3	18
	o meaningful definition	1		
	o grammatical accuracy	1		

Table 1: Sets of criteria with corresponding points

As Table 1 shows, the respondents designed five different sets of criteria for the identical test. The first two sets displayed are actually analogous because they consist of two identical criteria and just differ in the number of points allocated to each criterion. In the first set, a half point is assigned to each criterion whereas in the second set, one point is allocated to each criterion, ergo the maximum score (number of points) for the whole test is doubled from six points to twelve in the latter set. The third and fifth set comprise three criteria formulated in the same way but the points allocated to each criterion are different. In the third set of criteria, “meaningful definition” is awarded one point and the other two, “correct relative pronoun” and “grammatical accuracy”, are assigned just a half point, while the fifth set operates simply with one point for each criterion. Therefore, the score for a test item is also different together with the maximum number of points for the whole test, which is twelve, respectively eighteen points. The fourth set of criteria contains the same criteria listed in the third and fifth set, but “grammatical accuracy” is further extended by “correct spelling”. The number of points allocated to one test item is two and a half, so the whole test can be awarded fifteen points at maximum.

It is obvious from the evidence above that the same performance could be awarded different scores depending on the assessor’s subjective perception and decision. With no doubt, “the correct relative pronoun used in a definition” and “correct spelling” are the criteria that can be measured precisely. On the other hand, we can only assume what „a meaningful definition“, “an accurate and meaningful definition” and “grammatical accuracy” mean since no measurable descriptions were given. Moreover, it is not clear whether or not the criterion of “grammatical accuracy” includes “correct spelling”. That is why the identified sets of criteria imply that an identical test can be awarded different scores, not only due to different maximum number of points for the whole test, but also due to different criteria as well as due to the arrangement of points for each criterion. What is more, different pre-service teachers may perceive some of the criteria, for instance “meaningfulness of the definition” or “grammatical accuracy”, differently, because they are not specific and measurable enough. As a matter of fact, they can use some of the criteria more intuitively than based on precise objective “measurement”, which, however corresponds to common practice in schools where intuitive and subjective evaluation takes place (Slavík 62).

Grading Scales and Grades

From the whole survey sample, 12 different grading scales emerged. They are all displayed in three tables below according to the number of maximum possible points. In Table 1, there are grading scales with 6 maximum points, Table 2 presents grading scales with 12 maximum points, and in Table 3 grading scales with 15 and 18 maximum points are displayed.

The most common grading scale was designed by 6 student teachers. It is marked dark grey in Table 2. Grading scales marked light grey in Table 1 and Table 2 were both represented twice, i.e. in both cases 2 student teachers designed them that way. Grading scales marked white in all tables were represented just once.

In addition, not only do the grading scales differ in the number of maximum possible points that can be given to the test, they also diverge in the distribution of points to grades from 1 to 5. In the following tables, all the grading scales designed by the respondents and corresponding grades are presented:

Grade	Grading Scales – 6 points		
1	6	6-5.5	6-5
2	5.5-5	5-4	4.5 – 3.5
3	4.5-4	3.5-2	3-2.5
4	3.5-3	1.5-0.5	2-1
5	2.5-0	0	0.5 - 0

Table 2: Grading scales with 6 points at maximum and corresponding grades

Table 2 demonstrates the differences of point distribution within grading scales with 6 maximum points. As can be seen, the identical test can be awarded a different grade depending on the grading scale used. For example, a test that would be awarded 2 points would get three different grades (5, 3 and 4) from three different assessors. In fact, it means that a learner could get three different grades from three different teachers for exactly the same performance.

Grade	Grading Scales – 12 points					
1	12	12	12-11	12-10	12-10	12-10.5
2	11	11-10	10-9	9-7	9.5-8.5	10-8.5
3	10	9-8	8-7	6-4	8-7	8-6.5
4	9	7-6	6-5	3-2	6.5-5.5	6-4
5	8-0	5-0	4-0	1-0	5-0	3.5-0

Table 3: Grading scales with 12 points at maximum and corresponding grades

As can be seen from Table 3, the same principle about awarding different grades to the identical test applies in the grading scales with 12 maximum points. For instance, a test with 8 points would be given three different grades (5, 3 and 2).

Grade	Grading Scales – 15 points		Grading Scale – 18 points
1	15-14	15-12	18-16
2	13.5-10.5	11.5-9.5	15-12
3	10-8	9-6	11-8
4	7.5-5.5	5.5-2.5	7-4
5	5-0	2-0	3-0

Table 4: Grading scales with maximum 15 and 18 points and corresponding grades

Table 4 shows that when using the grading scales with 15 maximum points, the identical test would not be assigned the same grade either. In this case, the difference would always be just one grade. Compared to the grading scale with 18 maximum points, it is noticeable that a test with 8 points would be given the same grade regardless of whether the maximum number of points for the whole test was 15 or 18.

To sum up, the following table demonstrates the percentage range for individual grades:

Grade	Range in %
1	100 - 80
2	92 - 58
3	83 - 33
4	75 - 8
5	67 - 0

Table 5: **Percentage Range for Individual Grades**

As illustrated in Table 5, the percentage ranges of individual grades overlap diversely and there is a remarkable disproportion in the range of the percentages for each grade.

The rich diversity in the percentage range for individual grades also results in a variety of grades given to one identical test. The following table presents how the respondents rated the test on a scale from 0 to 100% and which grades they awarded.

Points in %	33	40	42	44	50	53	54	58	67
Given Grades	5,5,5	3	4,5	3	3,3,3, 4,4,4,4	3	4	2,3	2

Table 6: **Variety of percentage range and grades given to one identical test**

The findings presented in Table 6 uncovered that when assessed by 19 respondents, the identical test falls into the wide range from 33% to 67% and is awarded four different grades. A distribution of given grades in total for one identical test among the respondents is demonstrated in the figure below:

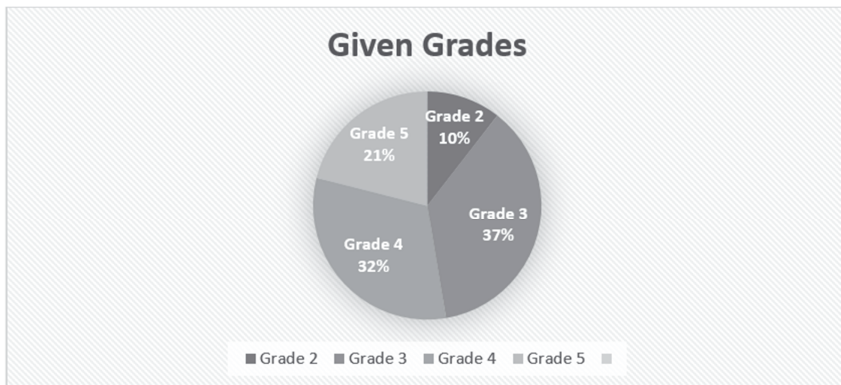


Figure 1: **The distribution of given grades**

As Figure 1 displays, the distribution of all grades given to the same test by the respondents varied. Seven respondents (which is 37%) awarded the test grade 3, six respondents (32%) gave it grade 4, four respondents (21%) offered grade 5 and two respondents (10%) offered grade 2. Grade 1 did not

appear in the responses at all. It means that one identical test was most often given grade 3, however, it was also awarded all the other grades except grade 1.

Findings from the second survey phase

Through the inductive coding, the thematic similarities across the respondents' responses were detected. Subsequently, thematically relevant categories were formed. Thus, the findings related to the reflection of the respondents' sole experience in assessing and grading a progress test can be categorized into: (i) Aspects influencing the perception of assessing and grading (ii) Feelings and emotions connected to assessing and grading a test, (iii) Reconsideration and reflection-based realization about assessing and grading.

Aspects Influencing the Perception of Assessing and Grading

The findings related to the influential aspects of assessing and grading were connected predominantly to the process of designing assessment criteria and distribution of points within the grading scales. The prospective teachers' responses concerning the influential aspects were thematically in concord. All the respondents claimed they were influenced by their previous school experience as learners:

- (1) *I was probably influenced by my own experience from school, when the range of points for grade 1 was usually smaller than for other grades.*
- (2) *This type of test evoked memories from school in me. When asked to define words or terms in a test, the answers were always awarded 2 points - 1 point for a meaningful definition and 1 point for correct grammar.*
- (3) *I was mainly influenced by my own experience from school. I tried to evaluate the test with regard to how my assessment system would be understood by my "student-self".*
- (4) *It was the first time I have ever graded a test and I was worried about doing it properly, so I applied the same grading system my teachers did at secondary school.*

The respondents' statements indicate that, in the middle of their regular bachelor studies, they are still principally influenced by their learner experience when assessing and grading. This phenomenon is understandable since they have never assessed and graded a test before. These findings correspond with Harrison's viewpoint (256). Yet, an exclusive influential aspect occurred among the answers when one of the respondents considered a learner's perspective when grading the test:

- (5) *I tried to get into the learner's shoes and see the mistakes with their eyes. For example, it could have been just a slip when the learner used the word "think" instead of "thing".*

The response may imply the respondent's reasoning in terms of learners' differences and needs. This attitude is closely related to the importance of individualization in school assessment described by many authors (see e.g. Spilková and Tomková, Helus, or Harmer) and demanded by current curricular documents (see e.g. FEP BE).

Feelings and Emotions Connected to Assessing and Grading a Test

Feelings and emotions played an important role both in the process of assessing and grading the test, as well as in the follow-up reflection in the class discussion.

The respondents revealed the following feelings and emotions they encountered during assessing and grading the test:

- (6) I was **nervous**. I **speculated** whether I should add or subtract some points and **couldn't decide**.
- (7) I **felt under pressure** when designing the set of assessment criteria because it is very difficult to create measurable and specific enough descriptors to match all possible answers in the test.
- (8) When assessing I was **in a mess and worried**. I didn't want to be unfair towards students.
- (9) I was **stressed** because I wanted to be fair and do it right.
- (10) In some cases, it happened that I understood what the learner wanted to say in the answer, but I had to follow the criteria I'd set. **It was difficult and I didn't like this feeling at all**.
- (11) When evaluating I became **frustrated and uneasy**. It was **hard to decide** how many points I should give to an answer even though the criteria I set seemed explicit and unambiguous.
- (12) I **felt very uncomfortable** because I wasn't sure at all how to evaluate the test. The learner generally seemed to understand the words, but the answer didn't meet the criteria I set.
- (13) I know for sure that I felt **unsure, anxious and strange** because I didn't know how to proceed properly. I **hesitated** when giving the points - is it ok to give only one point, or should I add half a point more?

These answers show that the respondents' prevailing feelings and emotions during assessing and grading the test were described as unpleasant and can be summarized into these labels: Doubts and Hesitation - Stress and Frustration - Discomfort and Uneasiness - and Fear and Anxiety. The last three labels can be further gathered into a hypernymous label Mental Strain.

Apart from the unpleasant feelings, nine respondents felt enthusiastic (label: Enthusiasm), yet only at the beginning of the test evaluation process; and one respondent claimed she felt neutral and unbiased because she knew the test and the evaluation of it was not real, as shown in the following quotes:

- (14) At first, I was **excited and looking forward** to a new experience.
- (15) To be honest, I felt **relaxed and fair-minded** because it was just "a mock test". However, if it was **a real test** of a real learner and me a real evaluator, I **wouldn't feel comfortable** about giving the learner a bad grade.

Besides the feelings and emotions experienced during the process of assessing and grading the test, the respondents expressed their feelings and emotions that appeared in the follow-up whole class discussion. These feelings and emotions were different from those the respondents experienced when assessing and grading the test. At the beginning of sharing their lists of criteria and assessing first test items, some respondents felt relieved and satisfied as they perceived similarities in their attitudes and assessment systems:

- (16) I was **glad** that some people created nearly the same criteria as I did.
- (17) At first, I thought that assessing a test would be a "**piece of cake**" because I evaluated the first two test answers in the same way as my classmates.

After the first differences in awarding the points to the test answers appeared in the discussion, the respondents were surprised:

- (18) I was **shocked** there were so many grading scales!
- (19) I was **surprised** that everyone had a different idea of which mistake is more or less "serious". But then, when our viewpoints varied I got to the point where I **began to doubt** my strategy and started **feeling uncertain**.

(20) I felt like a fool! It seemed everyone was used to a different system of assessing and grading and I appeared to be the most strict evaluator!

As the responses indicate, variations in the assessment and grading scales contributed to the feelings of uncertainty and puzzlement among the respondents. They described they were mainly puzzled about distinguishing between mistakes in terms of a meaningful definition, i.e. they were not sure what mistakes they could or should ignore as they did not prevent understanding.

(21) When reflecting the feelings and emotions they had when discussing the points and grades they awarded the identical test, the respondents admitted that they often hesitated. They were not sure about *either* the number of points or the proper grade to be given:

(22) I really did not know whether I was too strict or too forgiving. Which mistakes should I have overlooked because the definition still made sense?!

(23) Being in the middle of the discussion about the grades, I felt I should find at least half a point to be able to give the “learner” a better grade.

(24) When I heard the others talking about the number of points they gave to individual test items, I re-evaluated my assessing strategy at least a thousand times!

These doubts led the respondents to the need of revision of their assessment and grading strategies, as described in these respondents' quotes:

(25) I was confused because I realized my mistake in the logic of assessing according to my set of criteria. I should have designed them differently.

(26) Based on our discussion I came to a conclusion that if I were to assess the test again I would reconsider some issues, for example, what mistakes are still acceptable for me in terms of the meaning.

Out of nineteen respondents, only one did not feel the urge to reconsider anything in the process of assessment or grading. She said:

(27) I know that some teachers wouldn't even grade the test, but I definitely would. And I wouldn't change anything in my evaluation system and would stick to what I know from my teachers at secondary school.

However, on the basis of the similarities detected in respondents' answers in the analysis, the following thematically linked labels can be identified: Satisfaction and Relief - Surprise - Uncertainty and Puzzlement - Doubts and Hesitation - and the Urge to Revise the Assessment and Grading Process.

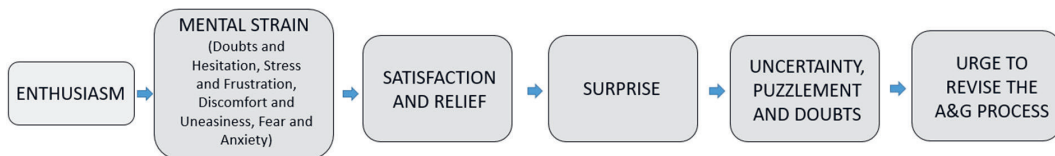


Figure 2: The Development of Feelings and Emotions during Assessing and Grading the Test and the Follow-Up Class Discussion

To sum it up (see Figure 2), during assessing and grading the test and the follow-up whole class discussion there was a noticeable progress and development in the respondents' feelings and emotions. Their feelings and emotions evolved from less represented Enthusiasm and largely represented Doubts and Hesitation through Mental Strain followed by Satisfaction and Relief, then Surprise and Uncertainty, Puzzlement and Doubts again to the final Urge to Revise the Assessment and Grading (A&G) Process.

Reconsideration and Reflection-Based Realization about Assessing and Grading Reconsideration about Assessing and Grading

The prevailing respondents' need for revision of their assessment and grading strategies resulted in potential changes in assessment and grading scales. All the respondents but one agreed that they would reconsider their set of criteria and the process of assessing and grading the test after the experience they had. They stated that mental strain, uncertainty and frustration in particular, was the main reason for the change of their prior attitude:

- (28) *Several times, when assessing and, especially during our discussion, I had this "urgent" feeling that I should have **evaluated the test differently**.*
- (29) *I was **really frustrated** when I saw that I **couldn't decide** which mistake was just a slip and which mistake was "serious". At the end, I had a feeling of **complete futility**.*
- (30) *If I evaluated the test again, I would **definitely change** the set of criteria (I would add some more) and I would give more points to each test answer.*
- (31) *I would be **more forgiving** in terms of grammar accuracy. After the discussion, I concluded that accuracy in grammar was not the main aim and that's why I would just focus on the message in the definitions.*
- (32) *In relation to my frustration when evaluating the test, I would **change my set of criteria, for sure**. I would **think** about the actual aim of the test **more**. And probably, I would make **some amendments** in my grading scale as well. When grading the test, I **wasn't sure** about the "right level" of strictness, but now I think I was far too strict.*

The conscious reconsideration of assessing and grading the test on the basis of their experience and its reflection brought the respondents a new view they probably had never come across before. They had to deal with their negative feelings of uncertainty and frustration, and tried to improve their assessment strategies. In their questionnaire answers, five of the respondents even analyzed and criticized the form of the test they were supposed to assess. They mentioned they had found the test poorly designed, as too many aspects of language were tested at once. The following quote summarizes their thoughts and speculations about it:

- (33) *If I were the teacher, I would **design the entire test differently**. Cutting of the points for a grammatically incorrect definition is wrong in this case, I think, because it was obvious that the learner knew the meaning of most words and probably had his/her own definition in his/her head. Therefore, I wouldn't push the learners into creating relative clauses when defining the words, but I would probably test this particular grammar separately. In addition, the words seemed completely unrelated and too abstract for a 9th grader. And what is more, I realized that I had experienced a lot of such meaningless tests at school.*

The respondents' answers suggest that the whole experience opened up a space for reflecting their feelings related to assessing and grading the test. This finding corresponds to the notion of Korthagen

et al. who mention that “the process of reflection can be activated by emotional signals, such as a feeling of frustration” (75). Subsequently, the respondents’ answers uncovered a possibility of potential changes in their attitudes and perception of school evaluation.

Reflection-Based Realization about Assessing and Grading

The findings in this part of the paper will be discussed in terms of the identified thematic labels, which are Complexity and Difficulty, Subjectivity, and the Power of the Teacher and Responsibility.

The investigated student teachers found the process of assessing and grading the test complex and difficult. In their answers, they mentioned difficulty of deciding on the correctness of test answers, the appropriate number of points as well as the corresponding grade, as can be seen in the following examples:

- (34) *I realized **how difficult it is to assess and grade one short and simple test. Even with the set of criteria, it was hard to decide which test answer was correct and which mistakes didn't influence the meaning of the definition.***
- (35) *I realized that **the whole issue of assessment is even more complex than I thought. It is not only about giving “a tick” or “a cross” next to the right, respectively wrong answer, and assigning points and giving grades. It's far more complex. Deciding on the points to give was the most difficult for me in the whole process of evaluation.***
- (36) *I think evaluating is extremely important and **extremely difficult** at the same time, and I still have to learn a lot about this process.*

Furthermore, it appears the respondents, to their surprise, realized that assessing and grading depends on individual teachers’ point of view and therefore is rather subjective:

- (37) *Surprisingly for me, I understood that evaluating a test by assigning grades **can never be fair and objective. Everyone can have a different point of view of what is right or wrong, especially when it comes to the meaningfulness of a definition, but also when distinguishing between “small” and “serious” grammar mistakes.***
- (38) *I got the idea of **how subjective the assessment is. It never occurred to me before! There were so many differences in assessment and grading scales among us.***
- (39) *Since I myself would probably evaluate this test in **another way a little later, just as well someone else could evaluate it completely differently. It's so subjective!***
- (40) *The criteria for assessing can be **very subjective** and, given the variety of mistakes, it is **difficult for a teacher to design a logical system that can be applied to all of them.***
- (41) *I got the feeling that assessing and grading a test is **not objective at all! Every teacher can have his/her own assessment scale and nobody cares.***

In addition to the subjectivity in assessing and grading, the power of the teacher and his/her responsibility related to the process of evaluation turned out to be a shared topic for the respondents. They pointed out several aspects that may influence the grade a test would be awarded, such as awareness of the aim; specific and measurable criteria set and communicated to the learners; the design of the grading scale; and personality of the teacher and other subjective determinants influencing the teacher when evaluating. Implicitly stated in the respondents’ answers, these aspects may have consequences for the learner and affect his/her life:

- (42) *In particular, I realized that assessing is first and foremost **a matter of responsibility. It is necessary to set accurate and measurable criteria, however, I am not sure whether this is possible in all cases.***

- (43) *I realized that the evaluation system and the design of both the list of criteria and grading scale is in full **control of the teacher**.*
- (44) *Now I know that it is **very important to think** about the aim of the test and the assessment criteria **in advance**. A teacher needs to **determine what s/he actually wants to test and evaluate**, also needs to think about the weight of individual criteria, and make them clear for the learners.*
- (45) *I realized that assessing a test **depends on many factors**. It **primarily depends on the teacher**, on his/her knowledge, attitude, personality and even mood. And this is what scares me the most - a teacher in a bad mood would/could give bad grades!*
- (46) *I realized that **a teacher actually has a great power**, which is manifested, for example, in evaluation. As a teacher, I can decide whether to give one point, a half point, or zero points for a definition with a wrong word. It's completely up to me, but it **influences the life of a learner**. And that is disturbing.*

It seems that assessing and grading the identical test together with the follow-up whole class sharing and discussion served as "an eye-opener" for the student teachers. Through the reflection, they appear to realize the complexity and difficulty of the evaluation process and the power together with the responsibility of the teacher stemming from the subjectivity of assessing and grading.

Moreover, out of the nineteen respondents, five revealed that the experience brought back their memories from primary school and led to further doubts about testing and assessment. According to these respondents, the grade seems to be a very poor form of feedback with uninformative value for the learner, and even the question of preferring and using other forms of assessment has been raised in their responses:

- (47) *I was **wondering** how many test results and grades were actually somehow **affected by the mood or other aspects influencing my teachers**, and how many grades really corresponded to my knowledge at that time?*
- (48) *It made me **speculate whether I would just repeat the same "patterns" as my school teachers without giving it more thoughts**. Now I know, I wouldn't.*
- (49) *Since the evaluation was limited to points and grades only, it **didn't provide the learner with the necessary and meaningful feedback**.*
- (50) *In my opinion, we, as teachers, should **help the learners in learning rather than just assess the correctness of test answers and give them "some meaningless number"**.*
- (51) *I think it is very important **not to end the evaluation process at a grade**. It is necessary to **go through the test together** with the learner and focus on the parts that were wrong. It would be even better, if the learners could **compare and discuss their answers** in pairs or groups so they can learn from their mistakes. In this case, I think that **giving grades wouldn't be a good idea**. All in all, from the learning point of view, grades are not necessary.*

In summary, for the investigated pre-service teachers, assessing and grading a test seems a challenging and demanding process in terms of setting the goal, criteria and grading scale; diversity of potential mistakes a learner can make; difficulty in deciding on the correctness of test answers, the appropriate number of points and the fitting grade; the influence of subjective factors affecting the teacher and related responsibility and power of the teacher. Some of the respondents considered peer and formative assessment to provide better feedback than grades.

Building-up a theory

In the process of constant comparative analysis (Švařiček et al. 2017; Chun Tie et al. 3), all the three categories characterized above were further examined, interconnected and built into new meaningful groups through exploring the relationships between them. With the help of the axial coding paradigm (Hendl 252-254), *A model of the process of change in the perception of assessment and grading in pre-service teachers* was designed (see Figure 3). School Assessment was identified as the core category that interconnects all the other categories: Teacher Education; Learner Experience; Assessing and Grading a Test; Sharing, Discussion and Reflection; and Change in Perception.

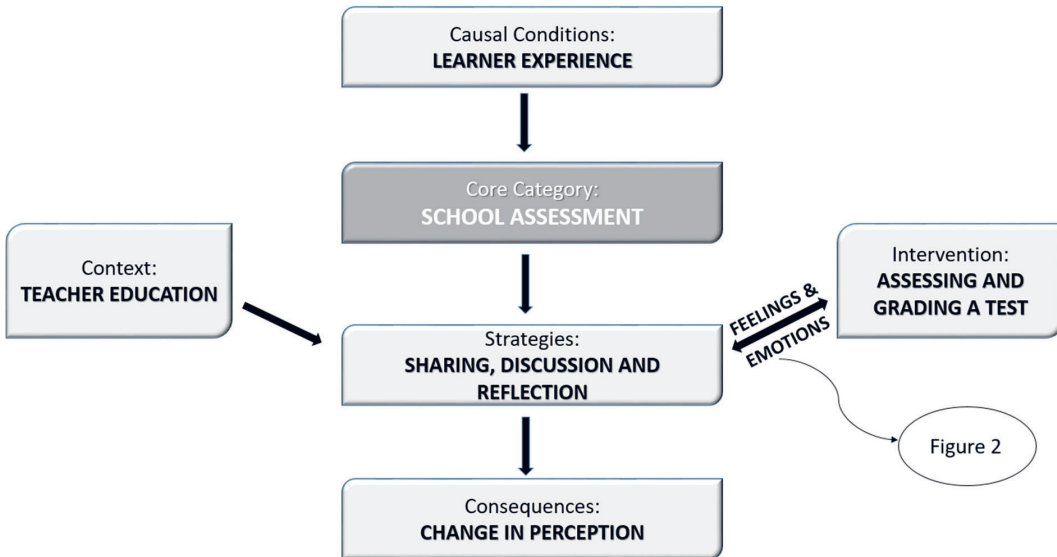


Figure 3: A model of the process of change in the perception of assessment and grading in pre-service teachers

The Model presented in the diagram provides a structure for building up a theoretical explanation of the phenomenon of school assessment through the eyes of the investigated pre-service teachers. Based on the analysis, a theory⁽¹⁾ emerging from the gathered data was constructed:

When assessing and grading a test for the first time, the investigated pre-service teachers rely on their prior school experience as learners. In a situation that is new for them, assessing and grading a test in this case, they are likely to follow and stick to procedures they know. This may be considered a preconcept in their perception of school assessment. When confronted with their own feelings and emotions related to assessing and grading the test in a class discussion and reflection, the pre-service teachers are likely to step out of their comfort zone. Through their conscious realization based on reflection, they may start a process of rethinking and subsequent rebuilding of their preconcepts, and thus consider some changes in their perception and attitudes.

Conclusion

This grounded theory study attempted to explore the pre-service teachers' perception of assessing and grading in the early stages of their teacher training. Two research questions were formulated at the beginning of the survey.

In response to the first research question, the outcomes of the study uncovered variations in the pre-service teachers' perceptions and conceptions. The future English teachers assessed and graded one identical test in different ways. The diversity was identified in different sets of criteria and distribution of points in grading scales, as well as in a different number of points given to the test. As a result of this diversification the same test was awarded four different grades by the pre-service teachers.

One possible explanation for the differences could be the lack of experience of the respondents. Another source of the differences, however, could be that evaluation is a highly subjective process even when carried out by experienced teachers (Slavík 62). Since the assessment criteria stated in curricular documents (classification rules) are too vague and hard to measure (Zitková, Hezká), the decision about the criteria, corresponding points, grading scales and the distribution of points in the grading scale is utterly up to the assessor. Although some recommendations may be given by the school principal or the head of the subject commission at school, still the evaluation is subjective. and just a matter of agreement among the teachers, but not compulsory.

As for the second research question, it may be concluded that the future English teachers in the present study perceive the process of assessing and grading a test as a complex and difficult process. Based on their self-reflection, the respondents realized that they were "trapped" by their past experience as learners and that it is quite difficult to get out of it. Although at first some respondents felt enthusiastic about a new experience, during the task of assessing and grading, their feelings changed to less pleasant. They felt stressed, uneasy, puzzled, anxious, even frustrated. The results of the study showed that their experience and related feelings and emotions were a starting point for speculating about school evaluation and realizing that there are more possibilities of assessing and grading, not just the one they were used to at their primary school.

Besides, there were respondents who proceeded even further in their thoughts when pondering about the design of a test and applying other forms of assessment. It is worth noting that some of them came to a concluding impression that a test in this particular form is pointless and grading it has no particular learning value, and more than grades demands formative assessment in the form of a meaningful feedback and discussion about the answers. This belief corresponds to current concepts of evaluation that consider summative feedback in a form of a grade not beneficial to learning because it prevents learners from participating in the feedback process, which is valued in learning (Volante and Fazio 751, William and Leahy 103, Crockett and Churches 2-3, Kolář and Šikulová 54).

The final developed theory grounded in the data gained from the student teachers is a dynamic outcome since stepping outside their comfort zones and subsequent changes in their perception, attitudes and beliefs are likely to occur during their further professional development.

Some questions may be raised about how the students would approach assessment when being in-service once since various factors, such as the school policy, professional philosophies of their future colleagues, or even the attitudes of learners' parents (Kolář and Šikulová 54, Slavík 87, Crockett and Churches 3), may influence them.

Yet, due to the fact that the respondents were pre-service teachers in their third semester, further investigations at the end of their teacher training would be beneficial. To gain even deeper insight into the student teachers' perception - to survey the decision making process behind awarding points in

the test items – supposedly even more differences in the perception of assessment would be revealed. It would be also interesting and beneficial to survey in-service teachers, explore their perception of the phenomenon of assessing and grading, and compare it with the result of pre-service teachers to get a more complex picture of the phenomenon.

It is important to be aware of the limitations of small-scale research, thus, the conclusions drawn in this study cannot be generalized. A small number of respondents and the applied research tool unquestionably limit the findings. Since the survey was done by means of a questionnaire, it was impossible to ask for further clarification in case of any misunderstanding. Nevertheless, the results illustrate that, while extremely speculative, given the lack of longitudinal data from more respondents, the possibility that perception about assessing and grading might change depending on reflection is quite possible.

Notes

(1) The theory was constructed in consistency with Birks and Mills who refer to grounded theory as a “research with an overtly interpretative component” where the researcher constructs the theory according to his own underlying philosophical assumptions about the world and in turn personal methodological position (4-8).

References

- Astin, Alexander W. and Antony L. Antonio. *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education*. American Council of Education and Rowman and Littlefield Publishers, Inc., 2012.
- Birks Melanie, and Jane Mills. *Grounded theory: a practical guide*. London: SAGE, 2015.
- Brown, Gavin and Ana Remesal. “Prospective Teachers’ Conceptions of Assessment: A Cross-Cultural Comparison.” *The Spanish Journal of Psychology*. 15(1), March 2012, pp. 75-89. Online: doi: 10.5209/rev_SJOP2012.v15.n1.37286
- Brunker, Nicole, Spandagou, Ilectra and Christine Grice. “Assessment for Learning while Learning to Assess: Assessment in Initial Teacher Education Through the Eyes of Pre-Service Teachers and Teacher Educators.” *Australian Journal of Teacher Education*, 44(9), 2019, pp 89-109. Online: <http://dx.doi.org/10.14221/ajte.2019v44n9.6>
- Calderhead, James and Peter Gates (Eds.) *Conceptualizing reflection in teacher development*. London: Falmer Press, 1993.
- Chráška, M. *Metody pedagogického výzkumu*. Grada, 2016.
- Chun Tie, Y., Birks, Melanie and Karen Francis. “Grounded theory research: A design framework for novice researchers.” *SAGE open medicine*, 7, 2 Jan. 2019, pp. 1-8, 2019. Online: <https://doi.org/10.1177/2050312118822927>
- Corbin, Juliet and Anselm Strauss. *Basics of qualitative research: techniques and procedures for developing grounded theory*. Thousand Oaks, CA: SAGE, 2014.
- Croketř, Watanabe.L, and Andrew Churches. *Mindful Assessment. The 6 Essential Fluencies of Innovative Learning*. Solution Tree Press, 2017.
- ČŠI. *Kvalita a efektivita vzdělávání a vzdělávací soustavy ve školním roce 2019/2020. Výroční zpráva České školní inspekce*. 2020. Online: http://www.csicr.cz/html/2020/Vyrocní_zpráva_CSI_2019_2020/html5/index.html?&locale=CSY

- DeLuca, Christopher, Klinger, Don A., Searle, and Shulha, Lyn M. "Developing a curriculum for assessment education". *Assessment Matters*, 2, Jan. 2010, pp. 20–42.
- Demetriou, Helen and Elaine Wilson. "A Return to the Use of Emotion and Reflection." *The Psychologist*. 21(11), Nov. 2008, pp. 938-941. Online: <https://thepsychologist.bps.org.uk/volume-21/edition-11/return-use-emotion-and-reflection>
- Green, Anthony. Assessment of Learning and Assessment on Learning. *Assessment and Evaluation. Basic Concepts in Testing and Assessment*. Wiley Online Library. 18 Jan. 2018. Online: <https://doi.org/10.1002/9781118784235.eelt0352>
- Harmer, Jeremy. *The Practice of English Language Teaching*. Pearson Longman, 2007.
- Harrison, Christine. "Teachers developing assessment for learning: Mapping teacher change." *Teacher Development*, 9, July 2005. pp. 255–263. Online: doi:10.1080/13664530500200251
- Helus, Zdeněk. *Dítě v osobnostním pojetí: Obrat k dítěti jako výzva a úkol pro učitele i rodiče*. Praha: Portál, 2004.
- Hendl, Jan. *Kvalitativní výzkum: základní metody a aplikace*. Praha: Portál, 2016.
- Holmes, Mary. "The Emotionalization of Reflexivity." *Sociology*, 44(1), Feb. 2010. 139-154. Online: <http://dx.doi.org/10.1177/0038038509351616>
- Kahl, Stuart R., Hofman, Peter and Sara Bryant. *Assessment Literacy Standards and Performance Measures for Teacher Candidates and Practicing Teachers*. Measured Progress, 2013.
- Korthagen, Fred, Kessels, Jos, Koster, Bob, Lagerwerf, Bram and Theo Wubbels. *Jak spojit praxi s teorií: Didaktika realistického vzdělávání učitelů*. Brno: Paido, 2011.
- Kolář, Zdeněk and Renata Šikulová. *Hodnocení žáků*. Grada, 2005.
- Košťálová, Hana, Miková, Šárka and Jiřina Stang. *Školní hodnocení žáků a studentů*. Portál, 2007.
- Mackenzie, Catriona. "Critical Reflection, Self-Knowledge, and the Emotions." *Philosophical Explorations, An International Journal for the Philosophy of Mind and Action*. 5(3), 2002, pp. 186-206. Online: <http://dx.doi.org/10.1080/10002002108538732>:
- Macpherson, Robin and Carl Hendrick. *Co funguje ve třídě?* Universum, 2019.
- Ministry of Education. *Framework Educational Programme for Basic Education (FEP BE)*. VÚP, 2007.
- Price, Margaret, Rust, Chris, O'Donovan, Berry, Handley, Karen and Rebecca Bryant. *Assessment Literacy: The Foundation for Improving Student Learning*. Wheatley: Oxford Brookes University, 2012.
- Shoffner, Melanie. "The Place of the Personal: exploring the affective domain through reflection in teacher preparation." *Teaching and Teacher Education*, 25(6), Aug. 2009. pp.783-789. Online: <http://dx.doi.org/10.1016/j.tate.2008.11.012>
- Schön, Donald A. *The reflective practitioner: How professionals think in action*. London: Temple Smith, 1983.
- Slavík, Jan. *Hodnocení v současné škole*. Portál, 1999.
- Spilková Vladimíra, Tomková Anna et al. *Kvalita učitele a profesní standard*. Univerzita Karlova, 2010.
- Spilková Vladimíra and Anna Tomková. "Profesionalizace vzdělávání učitelů pro primární školu – možnosti a meze." *Orbis Scholae*, 13(1), 2019, pp. 9-29. Online: https://karolinum.cz/data/clanek/7121/OS_13_1_0009.pdf
- Švaříček, Roman, Šed'ová, Klára et al. *Kvalitativní výzkum v pedagogických vědách*. Portál, 2007.
- Ur, Penny. *A Course in Language Teaching. Practice and Theory*. CUP, 1991.
- Volante, Louis and Xavier Fazio. "Exploring Teacher Candidates' Assessment Literacy: Implications for Teacher Education Reform and Professional Development." *Canadian Journal of Education*, 30(3), Jan. 2007, pp. 749-770. doi: 10.2307/20466661

- vyhláška č. 48/2005 Sb. Vyhláška o základním vzdělávání a některých náležitostech plnění povinné školní docházky. *Sbírka zákonů*, částka 11/2005, 25 Jan. 2005. Online: <https://www.msmt.cz/file/38827/>
- Webb, Norman L. *Assessment literacy in a standard-based urban education setting*. Paper presented to the American Educational Research Association Annual Meeting in New Orleans, Louisiana, Jan. 2002, pp. 1-19. Online: https://www.researchgate.net/publication/251923627_Assessment_Literacy_in_a_Standards-Based_Education_Setting#fullTextFileContent
- William, Dylan and Siobhán Leahy. 2020. *Zavádění formativního hodnocení*. EduLab, 2020.
- zákon č. 561/2004 Sb., zákon o předškolním, základním, středním, vyšším odborném a jiném vzdělávání (školský zákon). *Sbírka zákonů*, částka 190/2004. 10 Nov. 2004. Online: <https://www.msmt.cz/file/38970>
- Zembylas, Michalinos. "The Place of Emotion in Teacher Reflection: Elias, Foucault and 'critical emotional reflexivity'." *Power and Education*. 6(2), 2014. Online: <https://journals.sagepub.com/doi/pdf/10.2304/power.2014.6.2.210>
- Zítková, Helena and Hezká, Martina. *Smysluplné hodnocení pro školu 21. století*. Svět vzdělání. 2021.
- Züll, Cornelia. *Open-Ended Questions. GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. 2016. doi: 10.15465/gesis-sg_en_002