

Developing oral presentation-related self-assessment among prospective teachers of German as a foreign language: Analysis of self-assessment comments¹

Jana Veličková

Masaryk university, Faculty of Education, Department of German language and literature

Received 28th May 2023 / final version received 12th February 2024 /
/ accepted 14th February 2024

Abstract: Learner self-assessment is a significant predictor of learning outcomes (Hattie, 2018). However, it is insufficiently implemented in Czech secondary schools (Czech School Inspectorate, 2021). One of the reasons for this may be the lack of teachers' experience in developing their own self-assessment skills. This paper presents a study framed by a 12-week intervention programme to develop self-assessment skills focused on presenting in German among prospective teachers of German as a foreign language (n=15). The study examined the content (characteristics) of the participants' self-assessment comments collected before and after the intervention program. A total of 25 self-assessment comments were collected using the "Lautes Erinnern" method (13 before the intervention, 12 after the intervention) and analysed using the inductive category formation of the qualitative content analysis (Mayring, 2014). The analysis revealed three main characteristics of the development of self-assessment: increasing evidence in the self-assessment comments, a shift in focus from the predominance of non-language-specific to language-specific assessment, and a shift in focus from mostly negative to also positive aspects of performance. The study concludes with a discussion of the implications for better teacher education that develops their self-assessment skills appropriately.

Keywords: student self-assessment, development of self-assessment skills, content analysis, German as a foreign language, teacher education

Research has shown that student self-assessment increases student motivation (Benson, 2001), contributes to the development of learner autonomy (Tassinari, 2010), and has a positive impact on the quality of

¹ This study is part of the project MUNI/A/1335/2022 – *Research in foreign language didactics*.

student learning and learning outcomes (Hattie, 2018; McDonald & Boud, 2003). Furthermore, education should not only provide learners with knowledge and skills but also teach them how to assess and manage their learning so that it (ideally) becomes a lifelong process (Boud, 2003, p. 13).

In many countries, however, student self-assessment is not widely used in the classroom², although teachers are encouraged to promote self-assessment by professional frameworks³ or sometimes even by curriculum and legal documents⁴.

Not surprisingly, the lack of guided opportunities for self-assessment in the classroom can prevent learners from adequately developing their self-assessment skills (Apeltauer, 2010, pp. 22–27). Possible reasons for teachers not developing self-assessment may be related to busy lesson plans, a low belief in the effectiveness of self-assessment (e.g., Mäkipää, 2021), or a lack of experience and knowledge of self-assessment implementation (e.g., Volante & Beckett, 2011), as teachers may not be sufficiently trained to promote assessment for learning, not just of learning (McMillan, 2013, p. 5).

Building on the premise articulated by Raya (2014, p. 149), this article argues that prospective teachers need to gain experience in developing their self-assessment skills during teacher education programs. However, in order to design teacher education programmes that enable prospective teachers to (more effectively) implement student self-assessment in their (later) classroom practice, teacher educators need to understand, how the process of developing self-assessment skills occurs in prospective teachers.

Therefore, this article presents an exploratory study focusing on the content of self-assessment comments collected during a facilitated process to develop self-assessment skills. The comments are seen as a manifestation of the development of self-assessment skills.

The present study is based on a specially designed intervention to promote the self-assessment skills of Czech prospective teachers of German as a foreign language in the specific area of giving a short presentation. The aim

² For the Czech Republic, see Czech School Inspectorate (2021); for Canada, see Hunter et al. (2006); for Finland see Lasonen (1995).

³ For the Czech Republic, see *The framework of professional teacher qualities of a foreign language teacher* (Klečková et al., 2019).

⁴ For the Czech Republic see e. g. the *Elementary Education Act 561/2004 Sb.* (MŠMT, 2004) and *Framework education programme for elementary education* (MŠMT, 2017).

is to describe the content characteristics of the self-assessment comments collected at the beginning and at the end of this intervention. The results will shed light on how self-assessment is altered by scaffolding and serve as an empirically supported example for working with the development of self-assessment in teacher education.

First, student self-assessment is conceptualized in the context of education and foreign language learning. Then, in section 2, the research design of the study is presented (research question, participants, data collection, data analysis), followed by the analysis of the results (section 3). The paper ends with a discussion of the results and a conclusion.

1 Theoretical framework: Defining self-assessment in an educational and foreign language learning context

In the educational context, self-assessment has received more attention since the 1990s in relation to its conceptualisation as an essential aspect of formative assessment and the assessment-for/as-learning approach (Brown & Harris, 2013, p. 367; McMillan, 2013, pp. 4–6). In foreign language teaching, self-assessment has been of interest since the 1970s, as foreign language didactics has tended towards the constructivist paradigm (Weskamp, 2007).

To conceptualise student self-assessment for research and teaching purposes, authors have created taxonomies, typologies (e.g. Panadero, Brown & Strijbos, 2016; Taras, 2010, among others), or categorisations (Boud & Brew, 1995) of student self-assessment. However, there is no generally accepted definition. In summary, student self-assessment can first be conceptualised as a process of assessing the quality of one's abilities (skills, competences), processes or products related to learning (Andrade, 2018, p. 377), "based on evidence and explicit criteria, for the purpose of doing better work in the future" (Rolheise & Ross, 2001). This process usually takes place cyclically over time in relation to a particular task or performance and involves the use of various self-assessment tools (happy/sad face, rubrics, reflective journals, portfolios etc.; see e. g. Schneider, 1996; Wilkening, 2013) or implementation of self-assessment methods (Dochy et al., 1999, p. 335), practices or techniques such as "self-ratings, self-estimates of performance, and criterion- or rubric-based assessments" (Brown & Harris, 2013, p. 369). These tools or techniques represent the second conceptualisation. Using them, "students

describe (i.e., assess) and possibly assign merit or worth to (i.e., evaluate) the qualities of their own learning” (Panadero et al., 2016, p. 804).

However, self-assessment should not be limited to assessing the quality of learning. A higher level of cognitive engagement in self-assessment involves “deep engagement with the processes affiliated with self-regulation (i.e., goal setting, self-monitoring, and evaluation against valid, objective standards)” (Brown & Harris, 2013, p. 386). These cognitive processes can be referred to as metacognition (Belgrad, 2013, p. 335), i.e. the cognitive essence of self-assessment, thanks to which one can become aware of and reflect on one’s own actions (Krykorová, 2010, pp. 27–28).⁵ Thus, thirdly, self-assessment can be conceptualised as a self-regulatory ability (Brown & Harris, 2014).

The ambiguity in the conceptualisation of self-assessment is also reflected in the ambiguity in the terminology, as the term self-assessment is sometimes used as a synonym for self-evaluation (Boud, 2003, p. 13). The term self-assessment emphasizes the procedural understanding of self-assessment, is associated with a formative understanding of assessment and “involves students collecting data to evaluate their own progress” (Brown & Harris, 2013, p. 368). In contrast, self-evaluation refers to a summative understanding of self-assessment (Brown & Harris, 2013, p. 369) and can be conceptualised as one of the sub-components or phases of the self-assessment process (McMillan & Hearn, 2008, p. 41).

In this study, self-assessment is understood as a cyclical/iterative process of metacognitive operations that includes awareness of the goals of the activity, focusing on the object of evaluation (monitoring), evaluating the quality of this component, and formulating alterations to improve the quality. Therefore, the term “self-assessment” is used here.

According to a review study by Andrade (2018, pp. 309–401), the process of developing self-assessment in the learning context seems to be an under-researched area, as research tends to focus on the accuracy and consistency of self-assessment. The findings of these studies suggest that problematic self-assessment accuracy or consistency can be eliminated through appropriate and scaffolded self-assessment development (see, for example, Brown & Harris, 2013, p. 384; Ross, 2006; Ross et al., 1998).

⁵ Therefore, self-assessment is rightly referred to as a metacognitive learning strategy by the authors of various classifications (see Janíková, 2007, p. 95–106).

In order to develop the most accurate selfassessment possible, certain factors should be considered. The first factor to mention is the assessment criteria and descriptors, which are important in obtaining the most accurate self-assessment (Andrade & Valtcheva, 2009). The most effective way is to negotiate them directly with the self-assessors, or at least ensure that they understand them properly (Rolheiser & Ross, 2001, p. 7). Predictors of an accurate self-assessment also include learners' prior experiences with (Ross, 1998, p. 17) and perceptions of self-assessment (Brown & Harris, 2013, pp. 383–384), as well as various intrinsic factors such as fear of making mistakes or self-efficacy (Blanche, 1988, pp. 84–85). As for sociocultural factors, the influence of a culture's attitude towards self-criticism and self-praise (Hosseini & Nimehchisalem, 2021, p. 858) should be taken into account. In the context of foreign language learning, another factor that needs to be considered is the influence of learners' L2/L3 proficiency. Individuals with lower levels of language proficiency, and especially younger individuals, are more likely to overestimate themselves (Butler, 2023, p. 44).

The present intervention study takes the above points into account. Its design is also inspired by studies such as Léger (2009) and her self-assessment forms; Chen (2008), who developed a self-assessment through peer feedback; and Gil-Salomov and Benloch-Dualde (2016), who investigated self-assessment through peer feedback. Peer feedback is an important part of the presented intervention as it provides an additional incentive to reflect on one's own performance. It has also been argued that peer feedback is more acceptable and uses more natural and understandable language than teacher feedback (Black et al., 2004, p. 14). In addition to the feedback recipient, the feedback giver also benefits from the feedback process (Nicol et al., 2014). This is because it leads to a more intensive and deeper processing of the learning process and also serves as a stimulus for reflecting on one's own performance (i.e. for self-assessment) (Grotjahn & Kleppin, 2015, p. 145). Last but not least, the importance of peer feedback is also reflected in the design of some models for the development of self-assessment – for example in the model by Rolheiser and Ross (2001), which was an important starting point for the intervention in this study.

2 Methodology

This study examines the process of developing self-assessment skills of Czech pre-service teachers of German as a foreign language in the context of a specially designed intervention to promote self-assessment skills with a focus on giving a presentation. The aim of the study is to investigate the characteristics of self-assessment comments in the first and last phase of the intervention.

2.1 Sample

The participants were 15 prospective teachers in the second year of a bachelor's degree programme for teaching German as a foreign language at the secondary level (ISCED 2) who attended a one-semester German course at B2 level as part of their studies⁶. A purposive sampling strategy was used. As mentioned in the introduction, the first selection criterion was that they were prospective teachers. The second criterion was the level of German language proficiency. It was assumed that participants with advanced language proficiency (B1+ or higher) would have a deeper understanding of language structure, a higher level of language awareness, and would be better able to provide detailed selfassessments than participants with less advanced language skills.

A total of 17 students took part in the course and 15 of them were included in the study.⁷ All participants had passed a language exam at B1+ level in the previous semester and therefore met the language level requirements. German was their second foreign language and they had been learning it for 5–10 years.

The intervention consisted of six phases, but the reported data refer to the first phase (phase 1) and the last phase (phase 5). The first data collection was conducted with 13 participants and the second with 12 participants.⁸ Table 1 summarises the experience of all participants with selfassessment.

⁶ The language levels are based on the *Common European framework of reference for languages* (see Council of Europe, 2001).

⁷ The study did not include one student who had gone abroad during the course and one student who was studying a different programme from the other participants and whose language level could not be verified either.

⁸ They are not the same participants - three participants from data collection 1 did not take part in data collection 2 and two participants from data collection 2 did not take part in data collection 1.

Table 1*Previous experience of the participants (n=15) with self-assessment⁹*

How often have you experienced self-assessment	All the time	Often	Sometimes	Rarely	Never	I don't know
in the context of learning German as a foreign language?	1	1	2	10	1	0
out of the context of learning German as a foreign language?	0	2	3	8	2	0

2.2 Procedure

In line with the stated aim, a qualitative intervention research design according to Krainer and Lerchster (2012) was chosen. The intervention was based on the models of Rolheiser and Ross (2001) and Zimmerman (2002) and followed the logic of the transition from an object of assessment (other-regulated) to an active agent of assessment (self-regulated). This process is divided into six phases (phase 0 – phase 5) reflecting different degrees of dependence on external assessment. The goal was to bring the participants as close as possible to the stage of independence, in which they should already be able to evaluate their performance without external help (independent stage; Oscarson, 1997, cited in Poehner, 2012, p. 612) – see Table 2.

⁹ Respondents were asked to tick one option, but not everyone actually ticked one of the answers.

Table 2*Description of the intervention*

Phase 0	• Clarification of self-assessment, its relevance for learning	in-class
Phase 1	<ul style="list-style-type: none"> • Collaborative negotiation between the learners about the (partial) goals of the task and their concretization based on success criteria describing the characteristics of a good presentation and their indicators.¹⁰ • Presentation • Oral self-assessment (data source 1) • Oral peer feedback • Written teacher feedback (added after a few days) 	in-class
Phase 2	<ul style="list-style-type: none"> • Presentation • Oral self-assessment + Peer feedback (receiving and giving to another peer) 	in-class
Phase 3	<ul style="list-style-type: none"> • Presentation • Written self-assessment 	in-class
Phase 4	<ul style="list-style-type: none"> • Presentation • Written self-assessment 	in-class
Phase 5	<ul style="list-style-type: none"> • Setting goals for the presentation (to provide the participants with further internalization of the objective criteria for a successful presentation) • Presentation • Written self-assessment (data source 2) 	individually, out-of-class (online)

This 6-phase intervention (phase 0 – phase 5) was implemented in a 12-week (one-semester) German language course for prospective teachers of German as a foreign language. One of the aims of the course was to develop learners' ability to prepare and give a short oral presentation (about 3–5 minutes) on a selected topic. These presentations were the subject of the self-assessment tasks. The course took place for 90 minutes each week. The intervention was realised every one to three weeks and lasted about 30 minutes. The specific intervals between the phases were as follows: phase 0–1 one week, phase 1–2 two weeks, phase 2–3 three weeks, phase 3–4 one week, phase 4–5 two weeks. The intervals resulted from the inclusion of supportive interim phases between phases 1–4. These phases were completed by the participants

¹⁰ Learners formulated success criteria first in small groups and then under the supervision of an experienced assessor (course teacher). The negotiated rubrics, in the form of a mind map, was available to the learners throughout the whole intervention. The categorical system was used during the collaborative evaluation of a video of a foreign person giving a presentation in order to refine the categorical system and ensure a shared understanding of each criterion.

individually online via the Moodle platform. The aim of these phases was to support the development of self-assessment.¹¹

The self-assessment and the feedback from peers and the teacher were conducted in Czech, the participants' mother tongue. It can be assumed that a lack of foreign language proficiency would hinder the verbalisation of more complex cognitive content and lead to shorter self-assessment comments of poorer quality, which would impair the validity of the study (Seliger & Shomamy, 1989, p. 170).

2.3 Data collection, research instruments and data sources

The intervention included six phases. However, this text focuses only on the data from phase 1 (the first self-assessment in the intervention) and phase 5 (the last self-assessment in the intervention). The data was collected using forms with open-ended questions, which were answered verbally (data source 1) and in writing (data source 2). In order to anonymise the data, participants marked their self-assessments with unique codes that were assigned to them at the beginning of each data collection. In the oral self-assessment, participants said the code at the beginning of the recording.

The two data sources are:

- Data source 1 (DS 1) is the first self-assessment of the intervention (from phase 1), i.e., before peer and teacher feedback. The research instrument was an oral self-assessment in class, formulated on the basis of the following questions: How did it go? What went well? What could be done better?¹²

¹¹ Participants watched the recording of their presentation and completed an additional self-assessment of this recording. This additional self-assessment was intended to simulate the evaluation of others' performance, which may seem easier than evaluating one's own performance. However, at the same time, the evaluating/assessing others contributes to the development of effective self-assessment (Hattie, 2020, p. 146). Some of the supportive phases also included reflection on peer and teacher feedback.

¹² By adding the two more specific questions, it was assumed that a self-assessment structured with additional questions would contribute to a higher quality and range of self-assessment, similar to the study by Gan and Hattie (2014) on peer feedback.

- Data source 2 (DS 2) is the final self-assessment of the intervention (from phase 5). The research instrument was an out-of-class¹³ written self-assessment formulated on the basis of the following questions: How did it go? What went well? What could have been better and how could it be improved?

The oral self-assessments (from data source 1) were recorded using iPads and transcribed as pure verbatim protocols (Mayring, 2014, p. 45).

2.4 Data analysis

A total of 26 self-assessment comments from both data sources (13 from data source 1, 12 from data source 2) were analysed using qualitative content analysis according to Mayring (2010 in German; 2014 in English), in particular with inductive category formation (Mayring, 2014, p. 79). Due to the different nature of data sources 1 and 2 (see section 2.3), each data source was analysed separately, and the comparison of the results is only made at the level of discussion. The aim was to examine the content (characteristics) of the self-assessments from the first and last phase of an intervention to promote self-assessment skills. The coding was carried out using the software MAXQDA. Although the author analysed the data in Czech, the selected excerpts were subsequently translated into English for this article.

3 Results

First, an overview of the assigned codes and the inductively formed categories is given. The categories represent the content (characteristics) of the student self-assessments in the first and last phase of the intervention – see Table 3. This table also provides an overview of the frequency of occurrence of the most common characteristics of the student self-assessments. However, as the categories are not disjunctive, the frequencies only provide a rough overview. Next, the content of the self-assessments from data sources 1 and 2 is described and illustrated with examples from the data. The chapter concludes with a discussion of the results, including the limitations of the study.

¹³ Participants recorded their presentations and uploaded the recordings to the Moodle platform. It can therefore not be completely ruled out that they did not play back the recording of their presentation before the self-assessment.

Table 3*Codebook: Description of the categories; description and frequency of the allocated codes*

Categories	Description of the categories	Allocated codes	Description of the codes	Frequency in DS 1	Frequency in DS 2
1. Focus	Responses focusing on the whole performance (i.e., presentation) or the sub-aspects of the performance.	1.1. Global assessment	Responses focusing on the whole performance (i.e., presentation).	8	12
		1.2. Specific assessment	Responses focusing on sub-aspects of the performance (fluency, structure, content, grammar, vocabulary etc.).	61	223
2. Quality of the performance how well a task is performed.	Judgments about how well a task is performed.	2.1. Negative judgments	Judgments expressing negative, low, or undesirable quality of the performance or its sub-aspects.	44	60
		2.2. Positive judgments	Judgments expressing positive or good quality of the performance or its sub-aspects.	9	61
		2.3. Unclear or unspecified judgments	Judgments expressing unclear or unspecified quality of the performance or its sub-aspects.	4	2

Categories	Description of the categories	Allocated codes	Description of the codes	Frequency in DS 1	Frequency in DS 2
3. Reference norms for the quality	Benchmark against which participants compared the quality of their performance/task completion.	3.1. Criterion-referenced norm	Comparing one's performance to established standards represented by criteria (Brookhart, 2013, p. 258) – e.g., fluency, structure, grammar; vocabulary, presentation skills, pronunciation, ...	52	100 ¹⁴
		3.2. Individual goal-referenced norm	Comparing one's performance to individual goals of the performance or desirable outcomes expected by the participants themselves.	3	
		3.3. Social-referenced norm	Comparing one's performance to that of another person, e.g., a classmate (William, 2013, p. 207).	1	0
		3.4. Self-referenced norm	Comparison of current performance with some of one's own past performances (Brookhart, 2013, p. 258).	1	23
4. Attributions	Explanations of the cause of the positive or negative performance quality.	4.2. Attributions for the negative quality of the performance.	Explanation of the cause of the negative, lower, or undesirable performance quality.	12	14

¹⁴ As the self-assessment task in Phase 5 also includes a request to assess goals, it is not possible to distinguish clearly in the data between criterion-referenced and individual goal-referenced norms. Therefore, both types of norms are reported together.

Categories	Description of the categories	Allocated codes	Description of the codes	Frequency in DS 1	Frequency in DS 2
		4.3. Attributions for the positive or high quality of the performance.	Explanation of what caused the positive or higher performance quality.	3	14
5. Alterations for improving the quality of the performance	Suggestions for alterations that could improve the quality of the performance.	5.1. Presentation-related alterations	Suggestions for alterations that could have contributed to a better quality of the specific performance (presentation) or could contribute to a better quality of the next specific performance (presentation) in the future.	7	14
		5.2. General alterations	Suggestions for alterations that might help improve the quality of general language performance in the future. Expressed in general terms or as commitments.	3	18
6. Task-related responses	Reflection on the contextual aspects of the task outside of the performance (the task, the topic, the preparation, other settings – recording, place – at school / at home, the role of peers).	6.1. Task assignment-related responses	Reflection on the assignment, including the presentation topic.	1	7

Categories	Description of the categories	Allocated codes	Description of the codes	Frequency in DS 1	Frequency in DS 2
		6.2. Task-preparation-related responses	Reflection on the preparation of the presentation.	14	24
		6.3. Task-settings-related responses	Reflection on the task settings: influence of the recording, place - at school / at home, the peers).	0	12
7. Process-related responses	Description and feedback on the effort and processes involved in completing a task.	7.1. Process-related responses	Description and feedback on the processes involved in completing a task.	2	4
		7.2. Effort	Description of the effort involved in completing a task.	1	16

3.1 *Characteristics of self-assessment comments at the beginning of the intervention (from data source 1)*

The first self-assessment as part of the intervention comprised 7–159 words. The main focus here was on the sub-aspects of their performance (presentation). In these aspects, the participants primarily commented on their quality, which they evaluated by comparison to reference norms. Some participants also formulated attributions about the perceived quality of the sub-aspects and alterations to improve them. The category *focus of the self-assessment* is therefore cross-cutting and is presented separately for each area (quality assessment, attribution, alterations). In addition, some comments related to the task (e. g., the task assignment and the preparation of the presentation) as well as the processes and efforts involved in completing the task.

Quality of the performance

Participants expressed the perceived quality of their performance in the form of judgments about how well the task was performed. These judgments mainly referred to sub-aspects of the performance and not to the overall performance. More often, the judgments commented on non-language-specific aspects such as presentation skills (“Of course, the contact with the listener could be better. And not looking so much at the notes.” 1B_09), fluency (“I was actually stuttering.” 1B_14), or presentation structure (“The structure could be better.” 1B_03). Less evaluated were the language-specific aspects of performance such as grammar (“I know there were definitely some grammatical mistakes.” 1B_02) or vocabulary (“I definitely think the vocabulary could be better because I don’t think I have enough vocabulary knowledge for this topic.” 1B_11).

We can assume that the participants either do not consider the linguistic aspects of performance (grammatical and lexical correctness) to be important or they find it more difficult to evaluate them. Although the participants have language proficiency at the B1+/B2 level, they may not have sufficient knowledge of the language system and the terms used to describe the linguistic phenomena.

The perceived quality of their performance was verbalised by the participants as positive (focusing on strengths) or negative (focusing on weaknesses). Although the self-assessment task included the question “What did I do well?”, negative evaluations predominated. The focus of the negative

evaluations was mainly on the overall performance (“My presentation was horrible.” 1B_07). For the sub-aspects, the participants tended to give negative evaluations for fluency of the speech (“I stuttered and did not finish my sentences.” 1B_10) and presentation skills (“The contact with the listener could of course be better.” 1B_09). The positive evaluations occurred only marginally and focused on presentation skills (“What did I do well? [...] Greeting you nicely.” 1B_07) or the elaboration of the topic (“But yes, I talked about the topic, that was good.” 1B_03).

The predominant focus on negative aspects of one’s performance is usually discussed in the context of individuals who have a higher level of competence or knowledge and tend to underestimate themselves (Oscarson, 2009; Kruger & Dunning, 1999). Moreover, the focus on mistakes can be seen as a concomitant factor in the development of self-regulation (Keith & Frese, 2005). It can therefore be argued that developing a positive view of one’s performance is also essential, as self-efficacy of one’s actions forms the basis for the perception of one’s self-efficacy in future actions, as well as for the self-assessment itself (McMillan & Hearn, 2008, p. 44). However, given the superficial view of one’s own performance described above, it is questionable whether dealing with mistakes in this case can be seen as promoting the effective development of self-regulation.

Reference norms for the quality of the performance

The quality of performance was assessed by comparison with four norms. The criterion-referenced norm was predominant, as reflected in the use of labels for the evaluation criteria – e.g., “What could be better – probably everything, e.g., vocabulary, grammar.” (1B_07). Given that the success criteria were negotiated with the participants at the beginning of the intervention, it is not surprising that they refer to them in their self-assessments. Although the criteria seem to be the most comprehensible reference norm, in most cases, the self-assessments remained only at the rather general level of these criteria (“I am sure the structure could be better.” 1B_03) and more specific self-assessments were rare. Thus, this initial self-assessment can be seen as a somewhat superficial consideration of one’s performance, as it consisted only of mentioning selected or tangible aspects of the performance without delving deeper into the specific evidence.

As this was the first self-assessment conducted, participants may also have been overwhelmed by the complexity of the task. Furthermore, it can be assumed that the participants based their assessment on general impressions rather than specific evidence due to a lack of knowledge in the specific disciplines (such as phonetics, phonology, morphology, syntax, semantics, and pragmatics).

Although they had not been instructed to set task goals, some of the participants also compared their performance with their own expectations (code goal-referenced norm): “Better than what I prepared.” (1B_10).

To a lesser extent, there was also the self-referenced norm (“DSD¹⁵ made me a hundred times more nervous.” 1B_17) and the social-referenced norm (“This girl here was excellent, a hundred times better than me, but she also prepared two weeks in advance.” 1B_12).

The lack of a self-referenced norm is probably related to the fact that this was the very first presentation and the associated self-assessment in the context of the intervention and the participants were therefore not explicitly offered their own similar performance for comparison with the current one.

However, the low incidence of verbalization of social-referenced norms is surprising. Since the self-assessment in this phase was conducted orally in groups of two or three, one might expect participants to compare their performance with that of their classmates. There are two possible interpretations. The self-assessors were likely so focused on their own performance at that moment that a more tangible assessment framework for them was exactly the criteria discussed at the beginning of the intervention, and further comparison was already beyond their current cognitive capacity. An explanation based on the interaction between the individual’s so-called academic self-concept and the social norm is also offered (see Stiensmeier-Pelster & Schöne, 2008, pp. 66–67). The participants were only able to make the comparison in their minds and concluded from the prevailing negative self-assessment of their communication partner(s) that they also had a problem with the task and therefore possibly no longer considered it important or appropriate to articulate the comparison out loud.

¹⁵ DSD stands for the language exam for the certificate *Deutsches Sprachdiplom*.

Attributions

In addition to evaluating the quality of the performance, participants also commented on possible reasons for the perceived quality of the performance. Among these comments, those referring to possible reasons for a lower quality of the performance predominated. The participants attributed the low quality mainly to external factors, in particular, the limited preparation time (“If I had prepared it, it would not be a problem to cover it, but as it is, it’s almost on the spot.” 1B_12) and the given topic of the presentation. Among the external factors, they also mention poor language skills, especially in speaking: “I think my performance was very poor because I have a problem expressing myself unless it’s in writing, and communication is just a big problem for me. I can’t express myself, I can’t respond that quickly.” (1B_22)

Three justifications referred to the positive quality, that participants attributed to good preparation and prepared notes: “Maybe it was better when I didn’t have that support, and then when I was just at the end and I hadn’t written anything yet, I talked as if I was just thinking about myself and not sticking to what I had written because I was so lost in it.” (1B_10).

The predominant justification for the negative quality of the performance could be related to the fact that the self-assessments focus almost exclusively on the negative aspects of the performance. It is interesting to note that when it comes to attributing failures or negative characteristics of their own performance, participants tend to attribute these to external factors. This tendency is referred to as ego-defensive or self-protective attribution and is associated with a reduced willingness to take responsibility for failures or negative consequences of one’s actions (see Miller & Ross, 1975; Weiner & Kukla, 1970).

At the same time, attribution to external factors may indicate a preference for causes that are easier to infer (cf. the principle of cognitive economy – Vašátková, 1995, p. 11). Given the cognitively demanding nature of self-assessment, which takes place as self-monitoring during the performance itself, this explanation seems logical. Due to the cognitively demanding nature of self-assessment, participants likely no longer have the cognitive capacity to search for deeper causes for the quality of their performance.

Alterations

In addition to the negative evaluation of performance by pointing out performance weaknesses, only a few participants also formulated suggestions for their improvement. These alterations were more often specific to a presentation format, i.e. they related to how to improve the quality of the presentation given or how to improve the quality of the next presentation. The focus of these alterations was mainly on better structuring the presentation: “Firstly, I would definitely separate the individual parts so that I can formulate this in German. I would separate the fast food and the preparation of food at home and maybe focus a bit more on the disadvantages of eating at home or preparing food at home.” (1B_21). Preparing a presentation and working with notes was also mentioned: “And don’t look at the papers so much” (1B_09).

On a side note, there were also general alterations, i.e. not related to the presentation format, but to improving language skills in general: “Overall, I should probably learn to communicate better when it comes to oral communication.” (1B_22).

The focus of the presentation-related alterations corresponds to the focus of the evaluations, i.e. mainly on non-language-specific aspects. These alterations are relatively specific, so it can be expected that participants are more likely to improve the non-language-specific aspects of their presentation than the language-specific aspects. The alterations that focus on linguistic aspects tend to be very vague, so it can be expected that they are more difficult to implement and less likely to lead to improvements.

Task- and process-related comments

In the initial self-assessment of the intervention, there were also some task-related comments. They mainly concerned the inappropriateness of the presentation topic (“And overall, I found it really difficult to talk about this topic.” 1B_01) and the short preparation time (“My presentation was terrible because I didn’t have enough time [to prepare].” 1B_07)

One comment described the effort made to complete the task: “I tried to speak slowly.” (2B_02). Also only marginally represented were descriptions of the processes involved in completing the task, which can be illustrated by the following comment: “[...] I was just thinking about it and didn’t stick to what I had written.” (1B_10)

The presence of this marginal category illustrates that although the participants focused primarily on their own performance in their self-assessment, there are indications that they also take into account the situational context (preparation, task, topic) and are able to perceive the procedural level of performance to a certain extent.

3.2 *Characteristics of self-assessment comments at the end of the intervention (from data source 2)*

The final self-assessment in the intervention was 59–226 words long and contained mainly comments on the sub-aspects of one's performance (presentation) – e.g., their quality and the reference norms for the evaluation, attributions to justify the perceived quality of the performance and suggestions for its improvement (alterations). In a few cases, process- and task-related comments were also found in the data. The cross-sectional category *focus of self-assessment* is again reported as part of the other thematic categories.

Quality of the performance

Most of the comments on the quality of performance related to partial aspects of performance. Predominant were comments on fluency (“I was speaking fluently and I think I managed not to repeat myself.” 1H_14), grammar (“I think I managed to minimise my grammatical mistakes today.” 1H_13), presentation skills (“I also managed to be on time because my presentation is three and a half minutes long.” 1H_20) and the structure of the presentation (“I followed the structure of the presentation.” 1H_02). Among the evaluations, both the positive (i.e., referring to the strengths of the performance) and the negative (referring to the weaknesses) are almost equally represented. The focus of the *negative evaluations* is primarily on fluency (“In particular, the omission of the parasitic or filler sounds (“ehm” etc.) could have been better.” 1H_21). There were also some negative evaluations of the linguistic aspects of the performance in terms of grammar (“I made a lot of grammar mistakes.” 1H_20) and vocabulary (“I didn't know the vocabulary.” 1H_15).

The *positive judgments* include comments that focused primarily on the structure (“The presentation had structure. I mentioned the conclusion, advantages, disadvantages, general information and my own opinion.” 1H_10). The overall performance was also often rated positively (“I think it

was better than the very first presentation.” 1H_15). The linguistic aspects (grammar and vocabulary), on the other hand, were rarely rated positively (e. g. “I used more connecting expressions – einerseits, trotzdem, weder – noch etc.” 1H_01).

The high frequency of positive judgments could be related to the use of the self-referenced norm, because when using this norm two thirds of the participants formulated positive judgments, i.e. they focused on the improvement of the quality of their performance compared to the previous presentation(s). The increase in positive self-assessment due to the influence of an individual reference norm is attributed in particular to the fact that it strengthens confidence in one’s own abilities, weakens fear of failure and increases motivation (Rheinberg, 1980, cited in Rheinberg, 2008, pp. 183–184). The influence of considering oneself as successful on increasing learners’ self-efficacy (beliefs) is also confirmed by various studies conducted in the context of foreign language learning (e.g. Baleghizadeh & Masoun, 2013).

Reference norms for the quality of the performance

When evaluating the quality of the presentation, the participants compared their performance primarily with the *criterion-referenced norm* (“I managed to keep to the structure of the presentation.” 1H_03) and with the individual *goal-referenced norm* (“I managed to keep to the structure of the presentation, which was one of my goals.” 2H_20). Since the self-assessment task in this phase also contains the request to evaluate goals, it is not possible to make a clear distinction between criterion-referenced and individual goal-referenced norms. Therefore, both types of norms are reported together.

The evaluations of the sub-aspects were often quite specific, i.e. they did not remain at the superficial level of the evaluation criteria negotiated with the participants at the beginning of the intervention. This is illustrated by the following comments – one referred to the presentation skills and preparation of the presentation: “This time I managed to give a long presentation, over 4 minutes. And the preparation also only took 10 minutes, I managed to write in paragraphs.” (1H_15), the other on structure: “I think I did it right because I had an introduction in which I introduced the potential audience to the topic, then I explained the advantages, disadvantages and my own point of view and thanked them for their attention, so I think I did all the steps right.” (1H_20).

The comments with the poorest evidence were related to grammar. These comments often remained at the level of the assessment criteria discussed at the beginning of the intervention, as in the following comment: “[What could be done better?] The grammar.” (1H_05). A rare evidence-based self-assessment targeting grammar is the comment: “Sometimes I replaced a verb with a noun.” (1H_02)

At this point, we can discuss the possible influence of participation in the intervention, i.e. the repeated reflection on one’s performance together with peer feedback on the performance of others, which may have contributed to the internalisation of the content of the negotiated success criteria and a more detailed view of the performance. As the self-assessment in this phase took place outside the classroom, the unlimited time for writing the self-assessment may also have contributed to a more specific and comprehensive self-assessment. The specific case of grammar-focused self-assessment is discussed in section 4.

Very often there was also a *self-referenced norm*, i.e. participants compared their current performance with previous presentations. Most frequently, participants compared the overall presentation with all previous presentations (“That was my best attempt.” 1H_01) or with a specific presentation (“But I think it was better than the very first presentation.” 1H_15). Individual aspects were only marginally compared with previous presentations (“Relatively fluent presentation compared to my other presentations.” 1H_22).

The increase in the individual reference standard could be related to the repeated performance of the task, in which participants can compare their current performance with a similar previous performance. At the same time, the increase in the individual reference norm seems to have been reflected in an increase in the positive evaluation – see above.

Attributions

The participants formulated the same explanations for the causes of positive (higher) and negative (lower or undesirable) performance quality. They attributed the lower quality of their performance to various factors, with a particular emphasis on task-related circumstances (“[...], but when I started filming, I got quite nervous.” 1H_20) and poorer language skills (“On the other hand, I used listed phrases – that could also be because they were

new to me and I did not know the vocabulary for them, which is why I got stuck.” 1H_03). The higher quality was mainly attributed to the task settings, which was related to the fact that the self-assessment was conducted on an optional topic (“The topic was close to my heart and I had something to say about it.” 1H_15) and in conditions outside the classroom (“It was relatively easy. When you are alone, you are less nervous and can concentrate better.” 1H_03).

In the self-assessment after the intervention, a balanced reflection on possible causes for both the positive and negative quality of one’s performance can be recognised. This could indicate a tendency for self-efficacy to improve in similar tasks in the future – i.e., to be aware not only of the aspects that need to be eliminated in order to achieve better quality but also of those that need to be strengthened. Considering that participants attributed both the positive and negative quality of their performance mainly to external factors (situational and task-related conditions), it can be concluded that participants view the quality of their performance to a certain extent as an interplay of coincidences or circumstances beyond their control.

Alterations

Most participants formulated more general, not just presentation-related alterations. These focused mainly on speaking and vocabulary, such as in the following comment:

Include new words in your vocabulary because they are very familiar vocabulary that you will encounter throughout your life. Practice speaking more, for example by standing in front of a mirror and trying to speak or asking someone if you can try to present in front of him/her. Make pauses when speaking. (1H_02)

There were also alterations in terms of how the presentation could have been improved (“Maybe I could have talked about more areas that are relevant to the topic.” 1H_05) or could be improved in the future (“Be more natural. Don’t stick too much to predetermined points. Speak fluently.” 1H_22). As this example shows, the focus of these alterations was mostly on fluency. Interestingly, although grammar was one of the most negatively evaluated aspects in this phase, there was only one alteration that focused on improving grammar (“I will pay attention to some grammatical phenomena when I write notes, but I’m afraid that when I start speaking, I won’t be able to focus on grammar anymore.” 1H_20).

Overall, it can be said that both types of alterations (general and presentation-related alterations) are formulated quite clearly and with a fairly high level of evidence, as illustrated, for example, by the following example: “For the presentations to get better, it’s probably important to rehearse at least once a week, to record myself, to listen to myself and see if I’m making progress. And I think it then tends to get better.” (1H_14).

Both types of alterations are quite specifically formulated and show that participants have gained a deeper understanding of the desired performance and knowledge of specific strategies for improvement. The increase in cognitive capacity through repeated self-assessment, which became more routinized and therefore required less cognitive load, may also have played a role and enabled an enhancement of cognitive processes.

Task-related comments

The task-related comments primarily referred to the preparation of the task and the lack of preparation time (“There could have been a little more time for preparation.” 1H_02). The second most common were task *settings-related comments*. They reflected a shift in presentation and self-assessment situation from in-class to out-of-class presentations:

My performance was definitely influenced by the topic – I was allowed to choose a topic that I enjoyed and was interested in. And also the home environment. I was alone at home; no one was looking at me or listening to me. I wasn’t stressed that I might say something wrong. (1H_01)

Finally, the appropriateness of the presentation topic was also discussed:

I found the task assignment clear and the topic interesting. The holidays in our country and the holidays in Germany are very topical and it’s not bad to know something about them, whether there are differences, etc. (1H_02).

Overall, we can see that the participants talk quite a lot about the situational aspects of the performance, including its preparation. So, they do not limit their self-assessment to the presentation itself but perceive the performance in a much broader framework.

Process-related comments

This marginal category consists of comments describing the processes involved in carrying out the task. The main focus here was on the effort involved in completing the task. The *effort* was either only described: “That’s why I only chose the two best-known holidays (Christmas and Easter) and tried to describe them.” 1H_02), or more often its effectiveness (or efficacy) was also evaluated (positively and negatively in equal measure): “I tried to speak slowly and clearly – with occasional stuttering or longer pauses, I succeeded.” 1H_02). Exceptionally, there was also a justification for the efforts made: “In terms of grammatical correctness and word order, I tried not to make my sentences too long so that I wouldn’t get stuck and say stupid things.” 1H_02). Participants reported that they made an effort to elaborate on the topic well (“I tried to explain my point of view.” 1H_05) and to speak fluently (“I paid attention to speak fluently, but sometimes a new idea came to my mind and I wanted to say it, and then I realized I didn’t know one word of the sentence I wanted to say and I got thrown off track.” 1H_15).

Participants mainly described their efforts, which can be interpreted as “the effort is appreciated” or “this is my merit”, depending on the context. Participants also often directly evaluated whether their efforts led to success. This is likely to have a more positive impact on future goal setting and achievement, as participants can refer back to what worked for them and what did not. Describing and evaluating one’s own efforts can therefore be seen as a desirable feature of self-assessment. In the literature, it is linked to the concept of self-efficacy (Bandura, 1997).

4 Discussion

This study aimed to investigate the characteristics of the self-assessment comments from prospective teachers of German as a foreign language in the first and last phase of the intervention for developing self-assessment of speaking (giving a short presentation in German as a foreign language). The results of this study suggested three main characteristics of self-assessment comments that need to be considered in developing self-assessment skills: level of evidence, focus on strengths and weaknesses, and focus on language- and non-language-specific aspects of performance (giving a presentation).

On the one hand, three characteristics changed during the intervention. Firstly, there was a shift from an almost exclusive focus on weaknesses (negative evaluations) to an equal representation of positive and negative evaluations. Secondly, comments on non-language-specific aspects of performance predominated at the beginning of the intervention, while comments at the end also included language-specific aspects (grammar, vocabulary). Thirdly, the evidence in the self-assessment comments changed from poor to stronger over time. These aspects have already been discussed in the results section. On the other hand, one aspect of the self-assessment – the non-language-specific aspect of grammar – did not correspond to this trend. Although the participants frequently commented on grammar, the comments on grammar tended to have poor evidence and be evaluated negatively. Therefore, the self-assessment of grammar is discussed below. The discussion concludes with the limitations of the study.

The finding that the self-assessment of grammar tended to be negative and with poor evidence can primarily be attributed to the fact that the criterion of grammatical correctness in oral production is not so easy to define in terms of what it entails. This raises the question of how (and whether at all) learners should make and develop a self-assessment of grammar when giving a presentation or speaking in general.

There are not many studies on self-assessment of grammar in foreign language learning, which might illustrate the difficulty of this task. One of the few studies was conducted by Nurov (2000) in EFL settings. The study investigated the correlations between grammar-focused teachers' evaluations, students' self-assessments, and a test. The results showed a low correlation between the students' self-assessments and the other types of evaluations. Therefore, the question arises: What accuracy in self-assessment of grammar can be achieved?

Accuracy in self-assessment (not just) of grammar is thorny, claim Brown, Andrade and Chen (2015, p. 445). A research review by these authors suggests that a simple and concrete task and specific and concrete reference criteria promote accuracy. They also note that „more accurate self-assessors tend to be less optimistic than more inaccurate self-assessors” (Brown, Andrade & Chen, 2015, p. 446) and conclude their directions and cautions for research on student self-assessment by asking, “Does it matter if students are inaccurate in their self-assessments, so long as they are engaged in thinking

about the quality of their work?" (p. 445). In answering this question, we can rely on Brown, Andrade and Chen's (2015) argument that effective self-assessment doesn't necessarily have to be only accurate (p. 448).

In our study, self-assessment of grammar had relatively poor evidence, mainly in the form of vague phrases such as "The grammar could be better" or "There were a lot of grammar mistakes." If effective self-assessment serves "the purpose of doing better work" (Rolheiser & Ross, 2001, p. 8) and evidence-based assessment is essential for formative (self-)assessment (Brown & Harris, 2013, p. 368), it is not surprising that insufficient evidence does not serve this purpose of self-assessment well. Nevertheless, where is the desired level of evidence in a self-assessment focusing on the broad area of grammar? In the study by Pereira, Bermúde and Medina (2018), participants used video recordings of their speech to assess grammatical accuracy and range. In particular, they were able to focus on the confidence and clarity of grammatical structures, error-free sentences, and verb forms. However, they faced challenges in widening their grammatical structures. Despite the relatively poor evidence in the self-assessment of grammar, the participants improved their grammatical accuracy. The authors, therefore, conclude that the goal of self-assessment in grammar should not only be to improve awareness of correct grammar use but also to compare individual performance to get a sense of their improvement and support their motivational potential. Regarding the cognitive demands of self-assessment focused on speaking, self-assessment using recordings of one's performance seems to be an essential training tool. When high cognitive demands are combined with a difficult-to-delimit reference level of the criteria to be assessed, such as grammar, learners easily slip into "The grammar could be better." or "There were a lot of mistakes.". Or they do not pay attention to the grammatical level of the production in their self-assessment, as the following statement illustrates: "I pay attention to some grammatical phenomena when I write notes, but I'm afraid that when I start speaking, I won't be able to focus on the grammar anymore." 1H_20).

Limitations

Due to the qualitative research design and small sample size, the study's main limitation is that the results cannot be generalised. The research attempts to compensate for this by providing a greater depth of data.

Another limitation relates to the different forms of self-assessment prompts in each intervention phase. For this reason, data from the first and last phases are reported separately, and a comparison of the results is only discussed.

Finally, it should be noted that self-assessment processes have been examined on the basis of verbalized cognition, which may not correspond to fully realized cognition.

5 Conclusion

The study presented has provided valuable insights into the content of self-assessments, which otherwise often remains hidden, and their characteristics before and after participation in an intervention to develop self-assessment skills.

We can conclude that although there are many manuals for teachers on developing learners' self-assessment, the findings underline the importance of prospective teachers of German as a foreign language gaining experience in developing their own self-assessment skills during teacher education. The findings showed that prospective teachers do not necessarily know how to carry out self-assessment effectively, i.e., making it evidence-based and focusing on different levels of quality (positive and negative) of both language-specific and non-language-specific aspects. It can therefore be assumed that they would not develop this effectively with their students either. However, the completion of the intervention appears to contribute to the effectiveness and therefore validity of the self-assessment skills, thus enhancing the impact on teachers' ability to develop learners' self-assessment skills in their subsequent teaching practice. Based on these findings, the following implications can be drawn, which relate primarily to foreign language teacher education

At a general level of teacher education, the aim should be for future teachers to develop the habit of self-assessment or, more generally, of self-reflection on their actions – whether about their learning or, later, about their teaching. Of course, practice alone is not a sufficient condition. What is important is to develop self-assessment through reflective and structured work with goals and criteria at the outset and on an ongoing basis. Regardless of the area or focus of self-assessment, the goal is to adopt some kind of universal practice: If someone wants to evaluate the quality of his/her actions, he/she

needs criteria and indicators that represent the desired outcome and help him/her to find the evidence in their performance. In addition, he/she should proactively calibrate his/her own self-assessment with external feedback and, for example, work with video recordings of performance that he/she cannot get back to (e.g. an oral speech). Subsequently, working with goal setting, evaluating the effectiveness of one's efforts (attribution), and formulating alterations, and their implications. The intervention-based research design presented can serve as an empirically supported example of working with self-assessment development in teacher education that incorporates these aspects and through which problematic areas of self-assessment development can be identified and further addressed.

At the same time, it should be assumed that the development of self-assessment is highly individualized due to its interaction with various individual-specific variables, which also underlines the importance of an interdisciplinary approach to the development of self-assessment (and research on it). For curriculum development, it would be desirable to link the different approaches to the development of students' and prospective teachers' self-assessment within the pedagogical-psychological, domain-specific, and field-didactic dimensions of the studies and to take a more integrative approach to this topic. While the pedagogical and psychological components of the studies can effectively contribute to the individual-specific level of the self-assessment, the field didactics can contribute to the specifics of the subject of the self-assessment.

Specifically for the learning and teaching of foreign languages, the results suggest that a strong emphasis should be placed on negotiating criteria and then working with them. One particular area is the language-specific aspects of grammar and vocabulary and the associated criteria, where not everything can be covered. However, at least the grammatical phenomena that are addressed at a given language level can be clearly defined. In terms of vocabulary, it is also possible to focus on specific areas relating to the curriculum and the use of associated phrases.

A related point is that prospective teachers should have adequate knowledge of linguistic disciplines such as phonetics, phonology, morphology, syntax, semantics, and pragmatics in order to be able to name individual linguistic structures at an appropriate level of concretisation and performance. At the same time, it should be borne in mind that teachers should be expected to have a different knowledge of terminology and theory than learners.

Based on the above, the process of negotiating criteria can be considered as a research desideratum for an area where further data is needed, as it shows the importance of this phase for the subsequent process of developing effective self-assessment.

I would like to thank to the reviewers, as well as to the guest editor, Prof. Janíková, and Dr. Minaříková, the editor-in-chief of the Pedagogical Orientation journal, for their invaluable contributions and guidance throughout the publication process.

References

- Andrade, H. L. (2018). Feedback in the context of self-assessment. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 376–408). Cambridge University Press.
- Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice*, 48(1), 12–19.
- Apeltauer, E. (2010). Lernerautonomie, Lehrerautonomie und Deutsch als Fremdsprache. In Y. Eđit (Hrsg.), *XI. Türkischer Internationaler Germanistik Kongress. Globalisierte Germanistik: Sprache- Literatur-Kultur. Tagungsbeiträge* (pp. 15–34). Ege Üniversitesi.
- Baleghizadeh, S., & Masoun, A. (2013). The Effect of Self-Assessment on EFL Learners' Self-Efficacy. *TESL CANADA JOURNAL/REVUE TESL DU CANADA*, 31(1), 42–58.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Belgrad, S. F. (2013). Portfolios and E-Portfolios: Student reflection, self-assessment, and goal setting in the learning process. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 331–346). Sage.
- Benson, P. (2001). *Teaching and researching autonomy in language learning*. Longman/Pearson Education.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappa*, 86(1), 8–21.
- Blanche, P. (1988). Self-assessment of foreign language skills: Implications for teachers and researchers. *RELC Journal*, 19(1), 75–93.
- Boud, D. (2003). *Enhancing learning through self-assessment*. RoutledgeFarmer.
- Boud, D., & Brew, A. (1995). Developing a typology for learner self-assessment practices. *Research and Development in Higher Education*, 18, 130–135.
- Brookhart, S. M. (2013). Grading. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 257–271). Sage.
- Brown, G. T., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 367–393). Sage.
- Brown, G. T., & Harris, L. R. (2014). The future of self-assessment in classroom practice: reframing self-assessment as a core competency. *Frontline Learning Research*, 3, 22–30.

- Brown, G. T. L., Andrade, H. L., & Chen, F. (2015). Accuracy in student self-assessment: directions and cautions for research. *Assessment in Education: Principles, Policy & Practice*, 22(4), 444–457.
- Butler, Y. G. (2023). Self-assessment in second language learning. *Language Teaching*, 57(1), 42–56.
- Chen, Y. M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, 12(2), 235–262.
- Council of Europe (2001). *The common European framework of reference for languages*. Cambridge: Cambridge University Press.
- Czech School Inspectorate (2021). *Kvalita vzdělávání v České republice: Výroční zpráva 2020–2021*. Praha: Česká školní inspekce. Available from https://csicr.cz/CSICR/media/Prilohy/2021_p%c5%99%c3%adlohy/Dokumenty/VZ_CSI_2021_e-verze_22_11.pdf (01.11.2022)
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331–350.
- Gan, M. J. S., & Hattie, J. (2014). Prompting secondary students' use of criteria, feedback specificity and feedback levels during an investigative task. *Instructional Science*, 42(6), 861–878.
- Grotjahn, R., & Kleppin, K. (2015). *Prüfen, Testen, Evaluieren*. Klett-Langenscheidt.
- Hattie, J. (2018). *Hattie Ranking: 252 Influences And Effect Sizes Related To Student Achievement*. <https://visible-learning.org/hattie-ranking-influences-effect-sizes-learning-achievement/>.
- Hattie, J. (2020). *Lernen sichtbar machen für Lehrpersonen*. Schneider Verlag Hohengehren.
- Hosseini, M., & Nimehchisalem, V. (2021). Self-assessment in English language teaching and learning in the current decade (2010-2020): A systematic review. *Open Journal of Modern Linguistics*, 11, 854–872.
- Hunter, D., Mayenga, C., & Gambell, T. (2006). Classroom assessment tools and uses: Canadian English teachers' practices for writing. *Assessing Writing*, 11(1), 42–65.
- Janíková, V. (2007). *Autonomní učení a lexikální strategie při osvojování cizích jazyků*. MU.
- Keith, N., & Frese, M. (2005). Self-regulation in error management training: Emotion control and metacognition as mediators of performance effects. *Journal of Applied Psychology*, 90(4), 677–691.
- Klečková, G., & S., Hanušová et al. (2019). *Rámec profesních kvalit učitele cizího jazyka*. NÚV.
- Krainer, L., & Lerchster, R. E. (Hrsg.) (2012). *Interventionsforschung*. Band 1: Paradigmen, Methoden, Reflexionen. Springer.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134.
- Krykorková, H. (2010). Autoregulace a metakognice v kontextu psychologie učení. In K. Hrbáčková et al. (Eds.), *Kognitivní a nonkognitivní determinanty rozvoje autoregulace učení studentů* (pp. 27–44). Paido.
- Lasonen, J. (1995). A case study of student self-assessment in upper secondary education. In J. Lasonen & M.-L. Stenstrom (Eds.), *Contemporary issues of occupational education in Finland* (pp. 199–215). University of Jyväskylä, Institute for Educational Research.

- Léger, de S. (2009). Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals*, 42(1), 158-177.
- Mäkipää, T. (2021). Students' and teachers' perceptions of self-assessment and teacher feedback in foreign language teaching in general upper secondary education – A case study in Finland. *Cogent Education*, 8(1), 1–19.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Beltz Verlag.
- Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. SSOAR Open Access.
- McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy & Practice*, 10(2), 209–220.
- McMillan, J. H. (2013). Why we need research on classroom assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 3–16). Sage.
- McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons*, 87(1), 40-49.
- Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 82(2), 213–225.
- MŠMT (2004). *Školský zákon 561/2004 Sb. MŠMT*. http://www.msmt.cz/uploads/skolsky_zakon.pdf.
- MŠMT (2017). *Rámcový vzdělávací program pro základní vzdělávání*. MŠMT. https://www.msmt.cz/file/41216_1_1/.
- Nicol, D. J., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review Perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102–122.
- Nurov, A. (2000). *Self-Assessment of foreign language achievement: The relationship between students' self-assessment, teachers' estimates and achievement test*. Master's Thesis. Bilkent University.
- Oscarson, A.-D. (2009). *Self-assessment of writing in learning English as a foreign language*. ACTA UNIVERSITATIS GOTHOBURGENSIS. <http://files.eric.ed.gov/fulltext/ED505960.pdf>.
- Panadero, E., Brown, G. T. L., & Strijbos, J.-W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, 28, 803–830.
- Pereira, M. J. C., Bermúdez, J. A. L., & Medina, L. A. V. (2018). Improving L2 oral accuracy and grammatical range through self-assessment of video speech drafts. *Profile: Issues in Teachers' Professional Development*, 20(2), 127–142.
- Poehner, M. E. (2012). The zone of proximal development and the genesis of self-assessment. *The Modern Language Journal*, 96(4), 610–622.
- Raya, M. J. (2014). Lehrerautonomie: Auf dem Weg zu Freiheit und Eigenverantwortung. *Jahrbuch Deutsch als Fremdsprache*, 40, 147–162.
- Rheinberg, F. (2008). Bezugsnormen und die Beurteilung von Lernleistung. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (pp. 178–186). Hogrefe.
- Rolheiser, C., & J. A., Ross (2001). *Student self-evaluation: What research says and what practice shows*. <http://eloewen.pbworks.com/w/file/118309278/Student%20Self%20Evaluation%20What%20Research%20Says%20and%20What%20Practice%20Shows.pdf>

- Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experimental factors. *Language Testing*, 15(1), 1–20.
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Evaluation*, 11(10), 1–12.
- Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (1998). Skills training versus action research in-service: Impact on student attitudes to self evaluation. *Teaching and Teacher Education*, 14(5), 463–477.
- Schneider, G. (1996). Selbstevaluation lernen lassen. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschsprachenunterrichts Sondernummer 1996*, 16-23.
- Seliger, H. W., & Shohamy, E. (1989). *Second language research methods*. Oxford University Press.
- Stiensmeier-Pelster, J., & Schöne, C. (2008). Fähigkeitsselbstkonzept. Self-concept of Ability. In W. Schneider & M. Hasselhorn (Eds.), *Handbuch der Pädagogischen Psychologie* (pp. 63–73). Hogrefe.
- Taras, M. (2010). Student self-assessment: Processes and consequences. *Teaching in Higher Education*, 15(2), 199–209.
- Tassinari, M. G. (2010). *Autonomes Fremdsprachenlernen: Komponenten, Kompetenzen, Strategien*. Lang.
- Vašátková, D. (1995). *Atribuční procesy: Studie*. Gaudeamus.
- Volante, L., & Beckett, D. (2011). Formative assessment and the contemporary classroom: Synergies and tensions between research and practice. *Canadian Journal of Education*, 34(2), 239–255.
- Weiner B., & Kukla A. (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, 15(1), 1–20.
- Weskamp, R. (2007). Self-assessment/Selbstkontrolle, Selbsteinschätzung und -einstufung. In K. R. Bausch, H. Christ, & H. J. Krumm (Hrsg.), *Handbuch Fremdsprachenunterricht*. (pp. 382-384). Francke Verlag.
- Wilkening, M. (2013). *Selbst- und Partnerevaluation unter Schülern*. Beltz Verlag.
- William, D. (2013). Feedback and Instructional Correctives. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 197–214). Sage.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70.

Author

Mgr. Jana Veličková, Ph.D., Masaryk university, Faculty of Education, Department of German language and literature, Poříčí 9, 603 00 Brno, e-mail: velickova@ped.muni.cz

Rozvíjení sebehodnocení zaměřeného na ústní prezentace u budoucích učitelů němčiny jako cizího jazyka: Analýza sebehodnoticích výroků

Abstrakt: Sebehodnocení učícího se se jeví jako významný prediktor učebního úspěchu (Hattie, 2018), přesto je v českých základních školách realizováno nedostatečně (ČŠI, 2021). Jednou z příčin může být chybějící zkušenost učitelů s rozvíjením vlastního sebehodnocení. Předložený text představuje studii, jejímž rámcem byl 12týdenní intervenční program zaměřený na rozvoj sebehodnocení v oblasti prezentování v němčině u budoucích učitelů němčiny jako cizího jazyka (n = 15). Cílem studie bylo zjistit, jaké obsahové charakteristiky vykazují sebehodnoticích výroky participantů před a po absolvování intervenčního programu. Za použití metody *Lauter Erinnern* (vzpomínání nahlas) bylo získáno 25 sebehodnoticích výpovědí (13 před intervencí, 12 po intervenci), které byly analyzovány pomocí induktivní tvorby kategorií kvalitativní obsahové analýzy (Mayring, 2014). Analýza ukázala tři hlavní charakteristiky rozvoje sebehodnocení: nárůst evidence sebehodnocení, přesun od zaměření primárně na jazykově nespécifické aspekty výkonu i k jazykově specifickým a vývoj od převažujícího negativního hodnocení k zastoupení také pozitivního hodnocení. V závěru studie jsou diskutovány implikace pro kvalitnější vzdělávání učitelů v oblasti adekvátního rozvíjení jejich sebehodnocení.

Klíčová slova: sebehodnocení učícího se, rozvoj sebehodnoticích dovedností, obsahová analýza, němčina jako cizí jazyk, vzdělávání učitelů