education sciences

MDPI

Article

Quality of Reflections on Teaching: Approaches to Its Measurement and Low-Threshold Promotion

Katrin Arendt *, Lisa Stark, Anja Friedrich, Roland Brünken and Robin Stark

Department of Education, Saarland University, 66123 Saarbrücken, Germany; lisa.stark@uni-saarland.de (L.S.); a.friedrich@mx.uni-saarland.de (A.F.); r.bruenken@mx.uni-saarland.de (R.B.); r.stark@mx.uni-saarland.de (R.S.)
* Correspondence: katrin.arendt@uni-saarland.de

Abstract

The present study introduces a multidimensional approach to describing the composition of reflections, addressing previous inconsistencies in operationalization. The multidimensional approach was empirically explored in comparison to unidimensional reflection quality measures and examined in relation to reflection-related dispositions. Given the challenges of promoting reflection in teachers' daily practice, low-threshold interventions—repeated practice and structuring prompts—were investigated regarding their potential to foster reflective competence. Using an integrated model of reflection, five proposed dimensions—describing, evaluating, naming alternatives, justification, and self-reference—were identified and assessed alongside the quality measures reflection depth and breadth, and holistic grading. $N = 29$ teachers reflected verbally on two teaching videos, first openly and then with structuring prompts. The transcribed reflections underwent qualitative and evaluative content analysis to extract quantitative data. Compared to existing quality measures, the multidimensional approach provided deeper insights into the complexity of the reflections. The correlations between different approaches to reflection measurement and the reflection-related dispositions were only partially as expected. While repeated practice did not yield a training effect, structuring prompts improved reflection quality. Overall, the findings confirm the usefulness of the multidimensional approach as a valuable means to describe the composition of reflections and highlight its potential for quality measurement.

Keywords: teacher reflection; reflection quality; promotion of reflection; reflective competence

Academic Editors: Kira Elena Weber and Hendrik Lohse-Bossenz

Received: 30 May 2025
Revised: 4 July 2025
Accepted: 8 July 2025
Published: 10 July 2025

Citation: Arendt, K., Stark, L., Friedrich, A., Brünken, R., & Stark, R. (2025). Quality of Reflections on Teaching: Approaches to Its Measurement and Low-Threshold Promotion. *Education Sciences*, *15*(7), 884. https://doi.org/10.3390/educsci15070884

1. Introduction

The construct of reflection is becoming increasingly important in current research on teacher professionalization (Arendt et al., 2025; Lenske & Lohse-Bossenz, 2023). Reflection is regarded as a key component of professional competence and its lifelong development, as it can facilitate the alignment of theoretical knowledge and its practical application (Wyss & Mahler, 2021). Measuring the quality of reflection is an important prerequisite for further investigating how reflective competence can be promoted. Nevertheless, a pivotal aspect for evaluating reflection quality is the structure of the reflection process. Even though several approaches to assessing reflection quality have already been introduced (i.e., by using indicators of the depth or breadth of reflection), conceptualization of the reflection process remains ambiguous (Göbel et al., 2022). This process has previously been conceptualized as typically entailing a series of reflection steps (e.g., Hatton & Smith,

1995) that must be completed. However, other pivotal components of the reflection process (e.g., self-reference; Lenske & Lohse-Bossenz, 2023) are frequently disregarded. To date, multidimensional perspectives on reflection have not been applied. Therefore, the first step in advancing this field involves deriving a multidimensional approach to describing the reflection process. To this end, this study developed a multidimensional approach based on the integrated model of reflection proposed by Arendt et al. (2025), conducted an exploratory analysis of the resulting quantification of the proposed structure, and compared the data with existing measures of reflection quality.

As reflection can be regarded as an engine of lifelong teacher professionalization (Gruber, 2021), the question arises of how reflective competence can be promoted for in-service teachers. Low-threshold interventions such as structured assistance and repetition could be a means for accomplishing this objective, as they can be implemented efficiently in practice. Therefore, the present study also aimed to investigate the effects of low-threshold interventions, namely structured assistance and repetition, and how they impact reflection quality.

Thus, the aim of this study was twofold: First, a multidimensional approach to describing the reflection process was explored by comparing it to different reflection quality measures. In addition, the relationships between reflection quality and other constructs were investigated. Second, the effects of two low-threshold interventions—structured assistance and repetition—were examined regarding their effects on reflection quality and the reflection process. In this way, the present study attempted to gain, in an exploratory manner, new insights into the composition of the reflection process and reflection quality.

1.1. Conceptualization of the Reflection Process as a Multidimensional Construct

The concept of reflection is used inconsistently in research on teaching and teacher education (Aeppli & Lötscher, 2016; Clarà, 2015). In line with a scoping review by Arendt et al. (2025), reflection on teaching in teacher education can be defined as a "cognitive process of structured and critical examination of (one's own and others') pedagogical experiences in relation to oneself by referring to pedagogical knowledge, pedagogical content knowledge, and content knowledge as well as to skills and beliefs with the aims of developing oneself and improving professional practice" (p. 24). The integrated model of reflection on teaching (Figure 1) conceptualizes this concept in the form of an input–output structure: one's own and others' pedagogical experiences form the input as a prerequisite for the reflection process. It is important to distinguish between one's own experiences and those of others as input for the reflection process, as this can influence how teachers reflect. Reflecting on others' experiences seems to favor the reflection process (Seidel et al., 2013) compared to reflecting on one's own lived experiences, while self-referential elements (such as thoughts, feelings, and identity) tend to be neglected (Kleinknecht & Gröschner, 2016).

The aims of reflection, which are to develop oneself and improve professional practice, form the output of the reflection process. The reflection process itself is divided into reflection-related dispositions, the reflection process, and reflection performance (cf. von Aufschnaiter et al., 2019). The core reflection process comprises three sub-processes: describing, evaluating, and naming alternatives (Arendt et al., 2025). These processes can be arranged in a hierarchical order as they build upon each other. In addition, two further components—justification and self-reference—facilitate a decisive configuration of the individual sub-processes of reflection. On the one hand, pedagogical knowledge, pedagogical content knowledge, and content knowledge, as well as skills and beliefs, are used to justify the sub-processes of reflection (justification). On the other hand, the critical examination of teaching situations is governed by the construction of self-reference, which denotes the active integration of experiences, thoughts, or emotions into the reflection

process (Merkert et al., 2023). Justification refers to the explanations behind one's own judgments, while in an educational context, self-reference involves relating the teaching situation to one's own experiences, feelings, or practices. Although personal experiences can also be used as justifications, self-references should clearly refer to oneself or one's own practice. As demonstrated by the existing research, self-reference can be regarded as a crucial component of reflection, as it differentiates reflection processes from analysis processes (Arendt et al., 2025; Lenske & Lohse-Bossenz, 2023). Self-reference is also characterized by an explicit personal connection between the observed teaching situation and the observer's own professional identity. When making self-references in reflection, teachers do not merely analyze and evaluate the teaching practices of others. They also engage in a critical and reflexive examination of their own background and professional practice, classifying their observations against this background. The prerequisites for successful reflection are reflection-related dispositions, such as attitudes towards reflection, and reflection-related self-efficacy beliefs (Arendt et al., 2025). Reflection competence is indicated by reflection performance, i.e., written or oral verbalizations of the whole reflection process (Arendt et al., 2025; Stender et al., 2021).

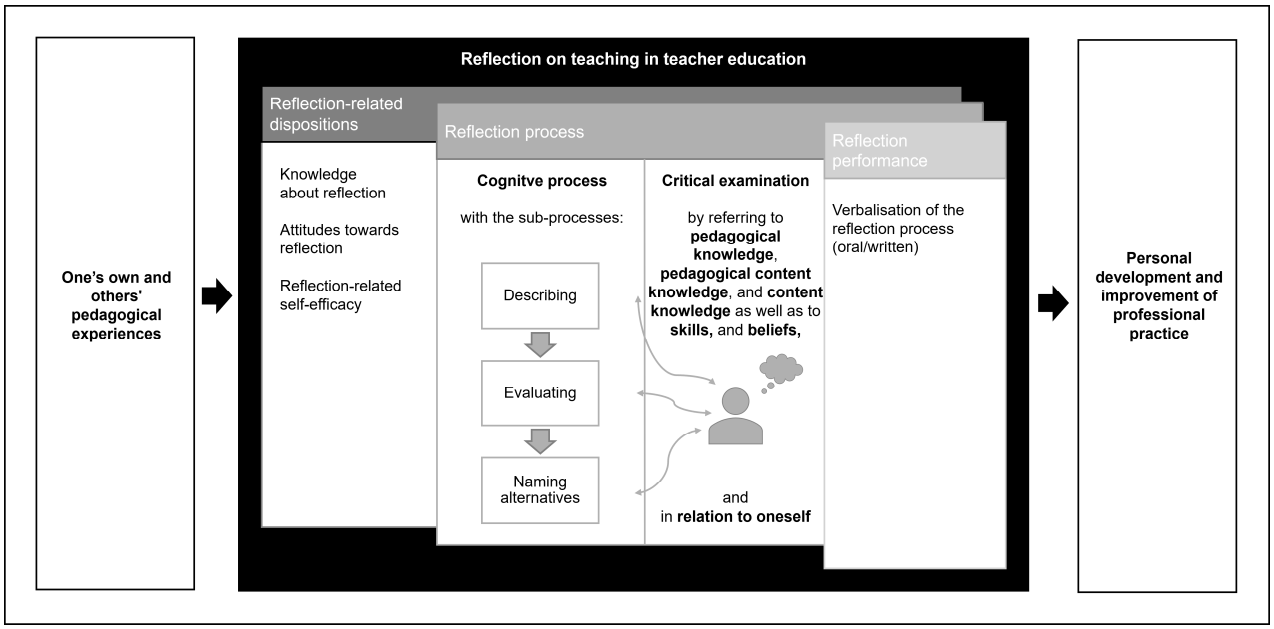

Figure 1. Integrated model of reflection in teaching and teacher education (based on Arendt et al., 2025).

Existing measures for the conceptualization of the reflection process have been criticized, particularly with regard to their failure to consider the reflection process in a sufficiently comprehensive manner. In response to these shortcomings, in this study, an exploratory multidimensional approach was developed based on the integrated model of reflection (Arendt et al., 2025). This multidimensional reflection process encompasses the following five sub-processes (or dimensions) of reflection: *describing* (1), *evaluating* (2), *naming alternatives* (3), *justification* by referring to knowledge, skills, and beliefs (4), and *self-reference* (5). This approach complements existing approaches to conceptualizing the reflection process, incorporating further crucial elements that facilitate a more comprehensive description of this process.

1.2. Reflection Quality in Current Research on Teaching and Teacher Education

Although already a focus of research, the above-mentioned scoping review (Arendt et al., 2025) revealed that reflection quality had not, at the time of writing, been explicitly included in models of reflection. Consequently, reflection quality has been operationalized

using a variety of approaches. Frequently, the *depth of reflection* and the *breadth of reflection* have been applied as criteria for assessing reflection quality. However, current empirical evidence suggests that these approaches do not provide a comprehensive assessment of reflection quality (Göbel & Neuber, 2022; Petko et al., 2019). These quality measures can be understood as unidimensional, as a single value is assigned to a reflection, which results in a loss of information. Consequently, the insights regarding the composition of the reflection process are limited, and the ability to develop nuanced approaches to promoting reflection quality is also restricted. Thus, the multidimensional approach (see Section 1.1), in line with the assumptions of the integrated model of reflection (Figure 1), might compensate for these limitations and bring together existing research. Taking a multidimensional approach to describing the composition of the reflection process—by incorporating various quality measures and reflection-related dispositions—can provide insight into the advantages and suitability of such an approach for educational research and practice.

1.2.1. Measures on Depth of Reflection

Indicators for the depth of reflection are grounded in level models of reflective competence. In such models, higher competence levels are associated with deeper reflections, and the highest levels are assumed to represent deeper reflection (e.g., Hatton & Smith, 1995).

The most commonly used level model for assessing reflection quality is that proposed by Hatton and Smith (1995). These authors defined four levels of reflection—descriptive writing, descriptive reflection, dialogic reflection, and critical reflection—which are conceptualized as a hierarchical series, with the highest level (critical reflection) being associated with the highest reflection quality, i.e., the deepest reflection. However, it has to be noted that Hatton and Smith (1995) excluded the first level from their quality assessment, which contains only descriptive elements, arguing that these elements did not technically constitute reflection.

In a similar way, other level models of reflection have also been employed to assess the depth of reflection (e.g., Kulgemeyer et al., 2021; Nowak et al., 2019; Wulff et al., 2021, 2023). The majority of studies that have assessed reflection quality in teacher education on this basis revealed that reflections mostly remained at the lowest level and largely contained only descriptive elements (Fränkel et al., 2022; Leonhard & Rihm, 2011; Weber et al., 2022; Wulff et al., 2021, 2023).

This hierarchical interpretation of levels of reflection is, therefore, open to criticism. Further information on the reflection process and its sub-processes is often not taken into account (Aeppli & Lötscher, 2016; Leijen et al., 2012), and if only the highest level reached in some parts of a verbalized reflection process is considered to be a quantitative indicator of reflection quality, it is not possible to ascertain at which stage other components of the reflection process are carried out. This leads to a reduction in the spectrum of observed phenomena, as the lowest level in particular does not represent the phenomenon of reflection.

1.2.2. Measures on the Breadth of Reflection

A second approach to reflection quality refers to the breadth of reflection. However, researchers refer to a wide range of elements when using the term breadth of reflection; for instance, the quantity and variety of content reflected upon have been frequently considered (Stender et al., 2021). It is imperative to examine the content of the reflection. However, the content criterion alone is not sufficient for evaluating the breadth of reflection, especially if quality is measured only by the quantity and variety of content. A more profound examination of the interconnections between reflection content and, for instance,

the sub-processes of reflection is frequently not conducted. Furthermore, the relevance of the content addressed in the reflection is often not assessed.

In other studies, the term breadth of reflection has been used to denote whether the reflection was focused on the actions of the teacher or those of the students (Fränkel et al., 2022; Leonhard et al., 2010), while Ullmann (2019) presented a further set of elements pertaining to the breadth of reflection, including feelings, personal beliefs, and awareness of difficulties. Thus, it appears that the concept of the breadth of reflection encompasses a multitude of disparate elements.

This leads to the question of whether, in order to prevent an arbitrary choice of reflection elements, all the previously studied elements should be viewed as genuinely integral components of the reflection process. Although these previously applied breadth-of-reflection elements might offer a comprehensive basis for assessing the specific composition of reflection quality, they seem to suit only very specific and limited research contexts. As research should also be driven by theoretical assumptions, from our point of view, the breadth of reflection components should be derived from theory. We argue that the core elements of the breadth of reflection should be assessed by referring to the integrated model of reflection (Arendt et al., 2025), especially regarding the justifications and self-references that are addressed during the reflection process.

1.3. Correlates of Reflection Quality in the Context of Teaching and Teacher Education

The integrated model of reflection presented above (Arendt et al., 2025; see Section 1.1) postulates that reflection-related dispositions, namely, reflection-related self-efficacy beliefs and attitudes towards reflection, are correlates of reflection quality; thus, they are components of the reflection process. In addition, general dispositions (e.g., the need for cognitive closure) and professional experience can supplement the set of reflection quality correlates.

Reflection-related self-efficacy expectations. Reflection-related self-efficacy expectations describe the beliefs that individuals hold regarding their capacity to overcome the challenges associated with reflection, based on their own abilities (Lohse-Bossenz et al., 2019). For instance, research indicates that future teachers with elevated reflection-related self-efficacy expectations, especially pre-service teachers in their traineeship, demonstrate enhanced resilience in the face of challenges (Lohse-Bossenz et al., 2019). Teachers with a high level of reflection-related self-efficacy examine and question their experiences, which should be related to higher reflection quality.

Attitudes towards reflection. Attitudes are defined as the cognitive and affective responses of an individual towards a particular object, individual, or situation, and are influenced by one's own and others' personal beliefs, values, and experiences (Ajzen & Fishbein, 2005). Thus, the attitudes–reflection relationship can be characterized as a valued person–object relationship. In the paradigm of reflection, reflection-related attitudes pertain to the attitudes and beliefs concerning reflection and its pertinence. In the context of education, both experienced teachers and student teachers exhibit a generally positive attitude towards reflection (Göbel & Neuber, 2020; Wyss, 2013). Research has demonstrated that educators who hold positive attitudes towards reflection are more likely to evaluate their own practices and further develop their teaching (Stender et al., 2021).

Consequently, based on the fundamental principles of self-efficacy theory and attitude theory, positive reflection-related dispositions can be assumed to align with the evolution of reflection processes and, therefore, with higher reflection quality.

Professional experience. Professional experience is frequently identified as a prerequisite for most aspects of professional development and competence (Berliner, 2001; Gruber, 2021). It is assumed that greater professional experience goes hand in hand with greater expertise (Berliner, 2001). Since reflection is thought to contribute to a higher level of

expertise among teachers (Gruber, 2021), it can be assumed that the level of professional experience and reflection quality are positively related.

Need for cognitive closure. The concept of need for cognitive closure describes an individual's desire for the resolution of actions and events, as well as their need for unambiguous responses to questions (Kruglanski, 2004). Consequently, the need for cognitive closure is closely associated with low tolerance for ambiguity (Kruglanski, 2004; Schlink & Walther, 2007). For example, the need for cognitive closure might influence how flexible, open, or deep a teacher reflects on teaching, especially regarding the extent to which they identify alternative actions. Therefore, to avoid ambiguity, teachers with a high need for cognitive closure may be more likely to draw rapid conclusions after immediate evaluation of actions during the reflection process rather than putting effort into reflecting more deeply and making further references to themselves or to their knowledge, skills, and beliefs. In terms of reflection quality, this means that teachers with higher levels of a need for cognitive closure are likely to show lower reflection quality.

1.4. Low-Threshold Promotion of Reflection in Teaching and Teacher Education

The promotion of reflection represents a crucial research objective, both in theoretical and practical terms (Kleinknecht & Gröschner, 2016; Prieto et al., 2020). Implementing the promotion of different elements of reflection usually involves complex interventions (Göbel et al., 2022; Kleinknecht & Gröschner, 2016; Klug et al., 2018). However, it is of particular interest to the professional development of practitioners to employ low-threshold interventions, which can be easily implemented in their daily practice.

In their daily practice, teachers experience many experiences on which they can reflect with the aim of improving their professional competence and teaching practice. Thus, the question arises of whether *repeated reflection* will, in reality, further improve teachers' reflective competence. Research has shown that repetition can improve test results in diverse domains (e.g., testing effect; Roediger & Karpicke, 2006). Hence, implementing repeated exercise by having learners reflect on two (or more) teaching scenarios consecutively might function as a low-threshold intervention for improving reflective competence. In this sense, subsequent reflections serve as an exercise to consolidate and deepen the structure of reflections.

Another low-threshold intervention could involve the use of *prompts* in order to facilitate structured reflection. Many studies have demonstrated that utilizing prompts is an effective method for fostering the development of reflective competence in the context of teaching and teacher education (e.g., Jung et al., 2021). The available evidence suggests that prompts, whether in the form of structured questions (Bradbury et al., 2020; Jung et al., 2021), structured journals (Cengiz, 2020; Korthagen, 1999), or collaborative dialogs (Clarà et al., 2019; Fazio, 2009), can facilitate the development of reflective practices among pre-service and in-service teachers. Structural support, especially in the form of targeted questions (Jung et al., 2021), seems to be particularly effective in facilitating the integration of reflection sub-processes (Weber et al., 2022). Prompts also function as a kind of feedback during the reflection process, as they promote complete reflections regarding the diverse constituent elements of reflection.

1.5. The Present Study

In this exploratory study, a *multidimensional approach to describing the composition of the reflection process* was implemented, in which five dimensions of reflection (describing, evaluating, naming alternatives, justification, and self-reference) are jointly applied. The study followed two main research objectives: First, the multidimensional approach was investigated exploratively by comparing it with different measures of reflection quality

(depth of reflection, breadth of reflection, holistic grading of reflection quality) as well as with reflection-related dispositions (reflection-related dispositions, professional experience, need for cognitive closure). Second, the potential effects of low-threshold interventions (repeated practice and structuring prompts) were examined regarding their functionality for changing the composition of the reflection process and reflection quality in teachers' daily practice. To achieve these objectives, teachers reflected on two teaching situations presented in videos delivered one after the other; each video was reflected on twice—first using open reflection, and then using structured reflection based on prompts. The following research questions were addressed:

Research Question 1. To what extent are the dimensions of the proposed multidimensional approach to describing the composition of the reflection process and different measures of reflection quality related to each other?

Measures on the depth of reflection, the breadth of reflection, and the holistic grading of reflection quality were related to the five multidimensional approach dimensions and each other. It was assumed that the measures would complement one another by showing at least moderate interrelations.

Research Question 2. To what extent are the dimensions of the proposed multidimensional approach to describing the composition of the reflection process and different measures of reflection quality related to reflection-related dispositions, professional experience, and the need for cognitive closure?

It was assumed that attitudes, reflection-related self-efficacy expectations, and level of professional experience would be positively related to higher reflection quality. Conversely, a negative correlation was postulated between the need for cognitive closure and reflection quality, since a high level of the need for cognitive closure can impede deeper and more elaborate reflection. Correlations with the five reflection process dimensions were assumed without a presumed direction of effect regarding the sub-processes describing, evaluating, and naming alternatives. It was assumed that the dimensions of self-reference and justification would show similar correlations with the reflection-related dispositions as with the quality measures.

Research Question 3. To what extent do low-threshold interventions (repeated practice and structuring prompts) foster reflection quality and change the composition of the reflection process regarding the five dimensions of the multidimensional approach?

A training effect through repeated practice was assumed in the form that reflection quality would improve, as measured by three reflection quality measures (i.e., depth of reflection, breadth of reflection, and holistic grading). In addition, a change in the proportions of the five reflection process dimensions exhibited was expected. The training effects were hypothesized for the overall comparison of reflections on the first and second video, for the comparison of the open reflections of the first and second video, and for the comparison of the structured reflections of the first and second video.

Furthermore, it was assumed that structuring prompts would improve reflection quality in all the reflection quality measures and that the proportions of all the multidimensional approach dimensions would shift due to the prompting. Such an effect was postulated for the overall comparison of the structured and open reflections, as well as for the assessments of each video individually.

In addition, a higher reflection quality was assumed for the open reflection on the second video compared to the structured reflection on the first video. A composition-specific change was also expected here for the five reflection process dimensions.

2. Materials and Methods

2.1. Participants, Design, and Procedure

Twenty-nine teachers from German secondary schools (13 female, age: M = 45.03, SD = 12.67) participated voluntarily (without receiving any reward) in this experimental study.

The study was carried out in a within-subjects design, with participants undergoing four conditions. Two videos (v1 and v2) were shown. For each video, the teachers were asked to verbalize their open reflections (or) and structured reflections (sr) with the help of prompts.

The participants were individually tested in an online setting for around 60 to 90 min. After being invited to a video conference, informed consent in written form was obtained from all the participants at the beginning of the study. After being welcomed and introduced to the procedure, the participants first completed an online demographic data questionnaire (see Figure 2). Subsequently, the first video was streamed via the video conferencing platform by displaying it on a split screen. The participants were first asked to reflect openly on this video (v1_or). This was then followed by their structured reflection with the assistance of the prompts (v1_sr). The participants then watched the second teaching video and reflected on it, first, openly (v2_or), and second, structured by prompts (v2_sr). Finally, the participants completed questionnaires on further variables (see Figure 2). The reflections were provided verbally, audio recorded, and later transcribed for analysis.

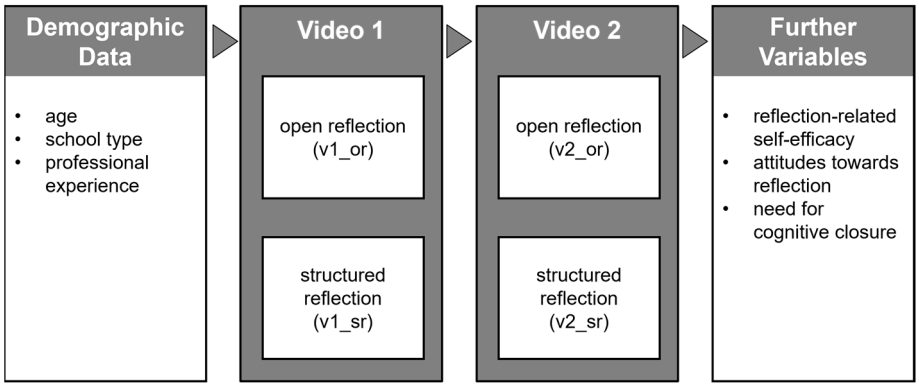

Figure 2. Design and procedure. (or = open reflection; sr = structured reflection; v1 = video 1; v2 = video 2).

2.2. Materials

Videos. Two videos were selected from the ProVision video portal of the University of Münster (Hemmer et al., 2018a, 2018b) for inclusion in this study. They were chosen on the basis of their similarity. In order to facilitate a comparison between the two videos, it was especially essential to ensure that they were aligned in terms of content and the general conditions of the teaching situation. The content of both videos comprised a sequence for the instruction of a learning phase in geography. The classrooms were presented in an overview, which alternated between the perspectives of the teacher and the students. The duration of both videos was approximately the same at between three and four minutes. Although the videos were largely similar in content, one video depicted a fifth-grade class (vf), while the other featured a seventh-grade class (vs) at a German secondary school. The potential effects of these differences in the videos were analyzed to secure internal validity.

Repeated practice. To implement the repeated practice element of the research, the two videos were presented to each participant. The order of the two videos was counterbalanced to eliminate video sequence effects.

Prompts. The participants were initially asked to verbalize an open reflection on both videos ("Reflect on the teaching situation. Express your thoughts loudly."). For the struc-

tured reflection, six structuring prompts were used. The prompts were selected according to theoretical considerations regarding the reflection process and were particularly focused on the sub-processes of describing, evaluating, and naming alternatives (e.g., "Think about how the teacher could have acted differently."). The decision to focus on these specific tasks was made to ensure that the participants were not overwhelmed by an excess of demanding tasks. Furthermore, prompts for justification and self-reference are more individual in nature and more difficult to depict in a standardized setting. For each video, the sequence of requesting an open and then a prompted (structured) reflection was the same.

2.3. Assessment of Reflection Quality

All audio recordings of verbalized reflections were transcribed and subdivided into sense-making units. The whole coding process was conducted by two independent raters with sufficient inter-coder reliability, all $\kappa > 0.75$.

Depth of reflection. In order to assess the depth of reflection, each sense-making unit was coded using the three sub-processes of the reflection process based on the integrated model for reflection (Arendt et al., 2025). The code "describing" was assigned when the situation was reproduced from the video (e.g., "She had an upward differentiation option in it."). In cases where the situation was evaluated, for instance, through the expression of impressions or feelings about the situation, the code "evaluating" was used (e.g., "I thought it was good that she visualized the work assignments on the worksheet [. . .]."). The code "naming alternatives" was assigned to statements on alternatives to the teacher's actions, as depicted in the video (e.g., "I would have proceeded in smaller steps."). A hierarchical interpretation of the code scores (0 = describing, 1 = evaluating, 2 = naming alternatives) was then used to calculate a final evaluation of the highest level reached in each reflection.

Breadth of reflection. In order to assess the breadth of reflection, additional elements were taken into account based on the idea that reflection not only involves analyzing a situation (i.e., by following the three subprocesses of describing, evaluating, and naming alternatives) but also includes justification and self-reference (Arendt et al., 2025). For this purpose, all sense-making units that referred to using personal knowledge, skills, or beliefs as justifications were coded "justification" (e.g., "Because you know that there are different learning channels. Not every child works well on the same learning channel or is better addressed by other work assignments."). Additionally, if a reflection contained self-references to oneself or one's own practice (e.g., "I can see myself in there a bit right now."), it was coded "self-reference". Afterwards, each reflection received a final evaluation score depending on whether it comprised justifications or self-references (0 = neither, 1 = one of the two, 2 = both).

Holistic grading of reflection quality. The holistic grading of reflection quality was applied as an intuitive and practical assessment approach to evaluating reflection quality. The raters evaluated each of the reflections based on their understanding of high-quality reflections. Their evaluations were informed by established theoretical models and prior research on reflection quality (e.g., Arendt et al., 2025; Hatton & Smith, 1995; Lenske & Lohse-Bossenz, 2023). Important criteria for holistic grading were the depth of insight, coherence of argumentation, and the presence of self-referential elements (von Aufschnaiter et al., 2019). In the context of holistic grading, school grades were assigned to each reflection according to the German grading system (1 = very good, 2 = good, 3 = satisfactory, 4 = sufficient, 5 = poor, 6 = insufficient). The weighted agreement of two independent raters, who assessed all the reflections, was satisfactory ($\kappa > 0.85$), and cases of uncertainty were discussed.

2.4. Multidimensional Approach to the Reflection Process

The multidimensional approach to describing the reflection process focuses on five dimensions—describing, evaluating, naming alternatives, justifying, and self-referencing—based on the components of the integrated model for reflection (Arendt et al., 2025). In this study, proceeding from this approach, whole sentences or sense-making phrases in all the reflections were coded according to the five dimensions using qualitative content analysis (Mayring, 2022), as in the cases of the depth of reflection and the breadth of reflection. For every dimension, the number of codes was counted. This number was then relativized in accordance with the total number of codes for the sub-processes identified in each reflection. By constructing this as a percentage value, the length of the verbalized reflections could be ignored. It was imperative to consider the five dimensions in relation to each other, thereby creating a structuring heuristic. By quantifying the reflections, this heuristic could be utilized to analyze the reflection process and make an initial statement about the quality of the reflections.

2.5. Measures

Professional Experience. To obtain data on the professional experience of the teaching staff, in an open questionnaire, the teachers were requested to indicate the number of years they had been employed in their profession since the completion of their traineeship ("How long have you been working in your profession since you finished your traineeship?").

Reflection-related self-efficacy. The validated questionnaire developed by Lohse-Bossenz and colleagues (2019) was utilized to assess reflection-related self-efficacy (*Cronbach's* $\alpha = 0.72$). The scale comprised 13 items (e.g., "It is easy for me to formulate explanations for the actions of individual pupils."), which were rated on a five-point Likert scale (1 = does not apply at all to 5 = fully applies).

Attitudes towards reflection. Attitudes towards reflection on teaching were assessed using a subscale of the ScRIPS (Student Feedback on Teaching and its Contribution to Lesson Reflection in the Practical Semester; Neuber & Göbel, 2018) on the relevance of reflection on teaching (Cronbach's $\alpha = 0.73$). This scale contained seven items (e.g., "For me, reflection on teaching is a central component of the teaching profession and should therefore be promoted as part of teacher training."), which were rated on a four-point Likert scale ranging from 1 = do not agree at all to 4 = fully agree.

Need for cognitive closure. The short scale by Schlink and Walther (2007) was employed to measure the need for cognitive closure (Cronbach's $\alpha = 0.82$). The eight items (e.g., "I don't like it when a person's statements are ambiguous.") were rated on a six-point Likert scale ranging from 1 = does not apply at all to 6 = fully applies.

2.6. Analytic Strategy

For all the global tests of significance, $\alpha = 0.05$ was applied as the conventional level of significance. All the significance tests were interpreted as two-tailed. Due to the limited sample size, correlations (Research Questions 1 and 2) were not analyzed by significance level, but the correlation coefficient effect size was considered. According to Cohen's (1988) convention, $|r| > 0.10$ indicates a small effect, $|r| > 0.30$ indicates a medium effect, and $|r| > 0.50$ indicates a large effect. For Research Question 3, planned contrasts were calculated according to the hypotheses. Cohen's d_z was calculated as the effect size for paired samples. The interpretation was based on the following convention: a $|d_z|$ of 0.2 indicated a small effect, 0.5 indicated a medium effect, and 0.8 indicated a large effect (Cohen, 1988).

3. Results

Table 1 shows the mean and standard deviations for all the variables by condition and total score, as averaged across the four conditions.

Table 1. Means and standard deviations for all the variables by condition and total score (average of the four condition values).

Variables	Condition				Total (N = 29)
	v1_or	v1_sr	v2_or	v2_sr	
Depth of reflection (0 to 2)	1.62 (0.56)	2.00 (0.00)	1.41 (0.50)	1.86 (0.35)	1.72 (0.23)
Breadth of reflection (0 to 2)	1.72 (0.46)	1.76 (0.44)	1.72 (0.46)	1.79 (0.41)	1.75 (0.23)
Holistic grading (school grades, 1 to 6)	3.45 (1.06)	2.79 (1.11)	3.55 (1.10)	2.76 (1.12)	3.14 (0.77)
Multidimensional approach					
Describing (proportion)	0.40 (0.21)	0.30 (0.16)	0.40 (0.16)	0.26 (0.11)	0.34 (0.10)
Evaluating (proportion)	0.48 (0.21)	0.50 (0.14)	0.53 (0.15)	0.53 (0.14)	0.51 (0.09)
Naming alternatives (proportion)	0.12 (0.13)	0.20 (0.08)	0.07 (0.10)	0.21 (0.12)	0.15 (0.06)
Justification (proportion)	0.29 (0.16)	0.28 (0.19)	0.27 (0.19)	0.26 (0.14)	0.27 (0.12)
Self-reference (proportion)	0.19 (0.19)	0.15 (0.12)	0.15 (0.15)	0.16 (0.12)	0.16 (0.10)
Professional experience (years)	-	-	-	-	15.28 (11.41)
Reflection-related self-efficacy (max = 5)	-	-	-	-	3.75 (0.35)
Attitudes towards reflection (max = 4)	-	-	-	-	3.56 (0.37)
Need for cognitive closure (max = 6)	-	-	-	-	3.11 (0.70)

Note. or = open reflection; sr = structured reflection; v1 = video 1; v2 = video 2.

To secure the internal validity of this study, an analysis was conducted to determine whether the reflections differed relative to the videos used: of a fifth-grade class (vf) vs. a seventh-grade class (vs). The results of the paired *t*-tests showed no significant differences for any of the reflection quality measures or reflection dimensions (see Table 2). Therefore, it could be assumed that the counterbalanced order of the videos was effective.

Table 2. Comparison of the reflections on the two videos: a fifth-grade class (vf) vs. a seventh-grade class (vs).

	Open Reflections vf vs. vs		Structured Reflections vf vs. vs	
	t(28)	p	t(28)	p
Depth of reflection	−1.16	0.255	−1.00	0.326
Breadth of reflection	−1.54	0.136	0.33	0.745
Holistic grading	1.56	0.130	0.42	0.676
Multidimensional approach				
Describing	0.83	0.412	0.89	0.381
Evaluating	0.08	0.469	−0.10	0.459
Naming alternatives	−1.36	0.185	−1.03	0.313
Justification	−0.58	0.570	−0.34	0.738
Self-reference	0.56	0.579	−0.88	0.387

3.1. Relationships Between the Different Measures of Reflection Quality and the Multidimensional Approach to Describing the Reflection Process

Table 3 displays the correlations between the mean values of all the reflection quality measures, the five multidimensional approach dimensions, professional experience, and the reflection-related dispositions.

Table 3. Bivariate correlations between the mean values for reflection quality, the reflective process dimensions, professional experience, and reflection-related dispositions.

	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Depth of reflection	0.26	**−0.52**	**−0.31**	−0.16	**0.73**	−0.01	−0.02	−0.09	0.16	0.16	−0.11
2. Breadth of reflection	1	**−0.67**	−0.02	−0.06	0.12	0.27	**0.66**	−0.10	0.06	0.10	0.06
3. Holistic grading		1	0.19	0.10	**−0.46**	**−0.34**	**−0.38**	0.16	0.06	−0.02	−0.02
Multidimensional approach											
4. Describing			1	**−0.80**	**−0.40**	−0.02	−0.20	−0.17	0.10	0.18	0.21
5. Evaluating				1	−0.24	−0.07	0.19	0.26	0.01	−0.01	**−0.35**
6. Naming alternatives					1	0.14	0.03	−0.13	−0.17	−0.27	0.19
7. Justification						1	**0.42**	**−0.36**	**−0.44**	−0.15	−0.03
8. Self-reference							1	0.17	−0.07	0.08	−0.07
9. Professional experience								1	0.12	0.11	−0.23
10. Reflection-related self-efficacy									1	**0.66**	−0.18
11. Attitudes towards reflection										1	**−0.33**
12. Need for cognitive closure											1

Note. All correlations are provided without a significance level due to the small sample size. Correlations above $|r| > 0.30$ are marked in bold.

The depth and breadth of reflection correlated positively with each other, but only with a small effect. Theepth of reflection correlated positively with the naming alternatives dimension, with a strong effect, and negatively with the describing and evaluating dimensions, with a medium and small effect, respectively. The depth of reflection did not correlate with the multidimensional approach dimensions (justification and self-reference).

The breadth of reflection did not correlate with describing and evaluating, but there was a small positive correlation between the breadth of reflection and the naming alternatives and justification. A strong positive correlation was detected between the breadth of reflection and the self-reference dimension.

A strong negative correlation was detected between the depth of reflection, the breadth of reflection, and holistic grading. Holistic grading also correlated with three of the five dimensions—naming alternatives, justification, and self-reference—but with a medium effect. Holistic grading correlated positively with describing and evaluating, but with a small effect.

The five dimensions of the reflection process (multidimensional approach) only partly correlated with one another: a strong negative correlation was observed between describing and evaluating, while describing and naming alternatives correlated negatively, with a medium effect. Describing and justification did not correlate. Conversely, a weak negative correlation was identified between describing and self-reference. Furthermore, evaluating and justification did not correlate, while a weak positive correlation was observed between evaluating and self-reference. A modest positive correlation was observed between the naming of alternatives and justification, whereas no correlation was evident between the naming of alternatives and self-reference.

3.2. Correlations Between Reflection Quality, the Multidimensional Approach Dimensions, the Reflection-Related Dispositions, and Professional Experience

The presumed correlations between the different indicators of reflection quality, the multidimensional approach dimensions, and the other variables did not appear (see Table 3). Professional experience had a weak negative correlation with the dimensions describing, naming alternatives, and self-reference, as well as with the breadth of reflection. A weak

positive correlation was observed between professional experience and evaluating, as well as with holistic grading. For professional experience and justification, a correlation with a medium effect size was detected, although this was negative, contrary to the initial assumption. Professional experience and depth of reflection did not correlate.

Reflection-related self-efficacy correlated weakly and positively with depth of reflection and describing, and negatively with naming alternatives. It did not correlate with the breadth of reflection and holistic grading, or with evaluating and self-reference. A strong negative correlation was found between reflection-related self-efficacy and justification.

Attitudes towards reflection correlated weakly and positively with depth of reflection and breadth of reflection, as well as with describing, and negatively with naming alternatives and justification. Holistic grading, as well as evaluating and self-reference, did not correlate with attitudes towards reflection.

For the need for cognitive closure, positive correlations with a small effect size were observed for the dimensions describing and naming alternatives, and a negative correlation with a small effect size was observed for depth of reflection. The breadth of reflection, holistic grading, and justification, as well as self-reference, did not correlate with the need for cognitive closure. A negative correlation with a medium effect size was found between the need for cognitive closure and evaluating.

3.3. Promotion of Reflection Quality and the Composition of the Reflection Process by Repeated Practice and Structuring Prompts

3.3.1. Depth of Reflection

Figure 3 displays the depth of reflection values for the four conditions. The contrasts between the open and structured reflections were significant in the overall comparison as well as for the first and the second video, respectively (see Table 4). The potential training effect was not significant in the overall comparison between the first and second video, nor in the comparisons of the two open reflections, and the comparison of the two structured reflections. However, there was a significant effect in the expected direction for depth of reflection between the structured reflection on the first video and the open reflection on the second video.

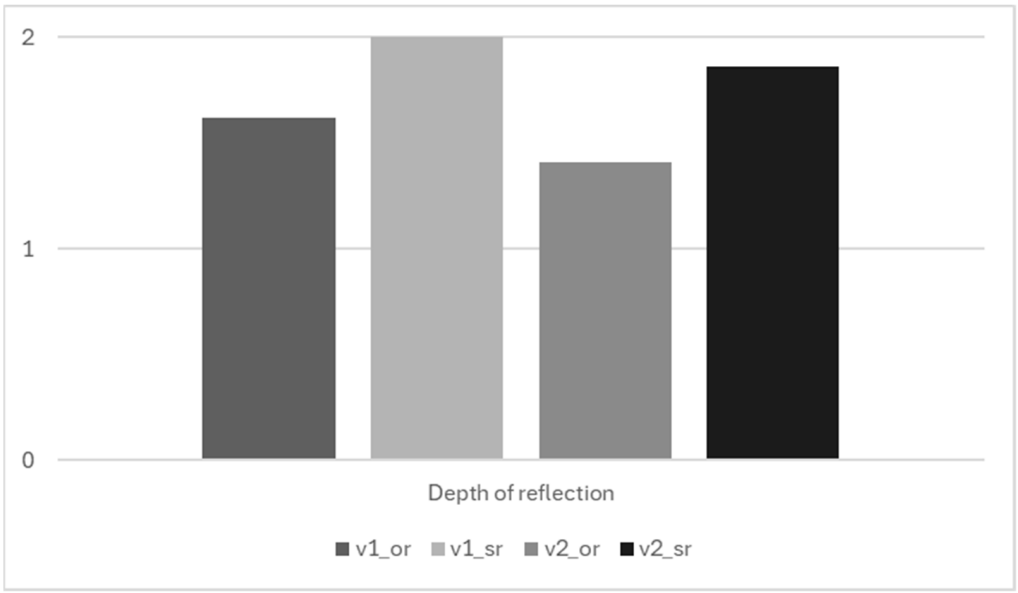

Figure 3. Mean values for the depth of reflection in all four conditions (or = open reflection; sr = structured reflection; v1 = video 1; v2 = video 2).

Table 4. Results of the contrast tests for all the reflection quality measures and the multidimensional approach dimensions.

	Contrasts for Effects of Structuring Prompts									**Contrasts for Training Effect**									**Additional Contrast**		
	or vs. pr			**v1_or vs. v1_pr**			**v2_or vs. v2_pr**			**v1 vs. v2**			**v1_or vs. v2_or**			**v1_pr vs. v2_pr**			**v1_pr vs. v2_or**		
	t	***p***	**| d_z |**	**t**	***p***	**| d_z |**	**t**	***p***	**| d_z |**	**t**	***p***	**| d_z |**	**t**	***p***	**| d_z |**	**t**	***p***	**| d_z |**	**t**	***p***	**| d_z |**
Depth of reflection	**−5.37**	**<0.001**	**0.51**	**−3.45**	**0.001**	**0.33**	**−4.11**	**<0.001**	**0.39**	**2.24**	**0.027**	**0.21**	1.90	0.060	0.18	1.27	0.209	0.12	**5.38**	**0.001**	**0.51**
Breadth of reflection	−0.63	0.528	0.06	−0.30	0.766	0.03	−0.60	0.552	0.06	−0.21	0.833	0.02	0.00	1.00	0.00	−0.30	0.766	0.03	−0.30	0.766	0.03
Holistic grading	**3.56**	**0.001**	**0.34**	**2.28**	**0.025**	**0.21**	**2.78**	**0.007**	**0.26**	−0.17	0.866	0.02	−0.36	0.720	0.03	0.12	0.905	0.01	**−2.64**	**0.010**	**0.25**
Multidimensional Approach																					
Describing	**3.85**	**<0.001**	**0.36**	**2.18**	**0.032**	**0.21**	**3.27**	**0.001**	**0.31**	0.52	0.602	0.05	−0.18	0.861	0.02	0.92	0.362	0.09	**−2.35**	**0.020**	**0.22**
Evaluating	−0.11	0.912	0.01	−0.29	0.775	0.03	0.13	0.896	0.01	−1.23	0.220	0.12	−1.08	0.282	0.10	−0.66	0.509	0.06	−0.79	0.429	0.08
Naming alternatives	**−5.78**	**<0.001**	**0.54**	**−2.92**	**0.004**	**0.28**	**−5.25**	**<0.001**	**0.49**	1.11	0.269	0.10	1.95	0.053	0.18	−0.38	0.703	0.04	**4.87**	**<0.001**	**0.46**
Justification	0.11	0.916	0.01	0.14	0.888	0.01	0.01	0.993	0.00	0.56	0.575	0.05	0.46	0.644	0.04	0.33	0.741	0.03	0.32	0.748	0.03
Self-reference	0.56	0.578	0.05	0.86	0.390	0.08	−0.07	0.941	0.09	0.62	0.537	0.06	0.91	0.367	0.09	−0.03	0.975	0.00	0.04	0.966	0.00

Note. Significant results are printed in bold; or = open reflection; sr = structured reflection; v1 = video 1; v2 = video 2.

3.3.2. Breadth of Reflection

Figure 4 shows the mean values for the breadth of reflection for all four conditions. In contrast to the hypotheses, there were no significant differences in the breadth of reflection (Figure 4, Table 4). Neither were there any training effects for the overall analysis, open reflections, or structured reflections. Nor were there any applied prompts effects (overall, video 1, video 2), or any significant differences between the structured reflection on video 1 and the open reflection on video 2.

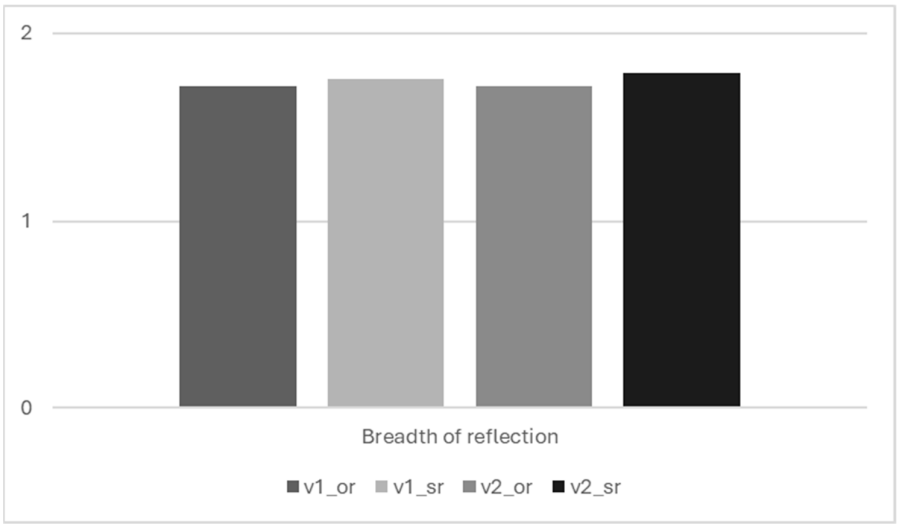

Figure 4. Mean values for the breadth of reflection in all four conditions (or = open reflection; sr = structured reflection; v1 = video 1; v2 = video 2).

3.3.3. Holistic Grading of Reflection Quality

The holistic gradings of reflection quality for all four conditions are displayed in Figure 5. As assumed, contrast tests showed that the open reflections had a significantly worse holistic rating compared to the structured reflections, overall and for both videos (see Table 4). However, the assumed training effect was not detectable for holistic grading (for the overall analysis, open reflections, structured reflections). Contrary to the initial hypothesis, the open reflection on the second video achieved a significantly worse grade than the structured reflection on the first video.

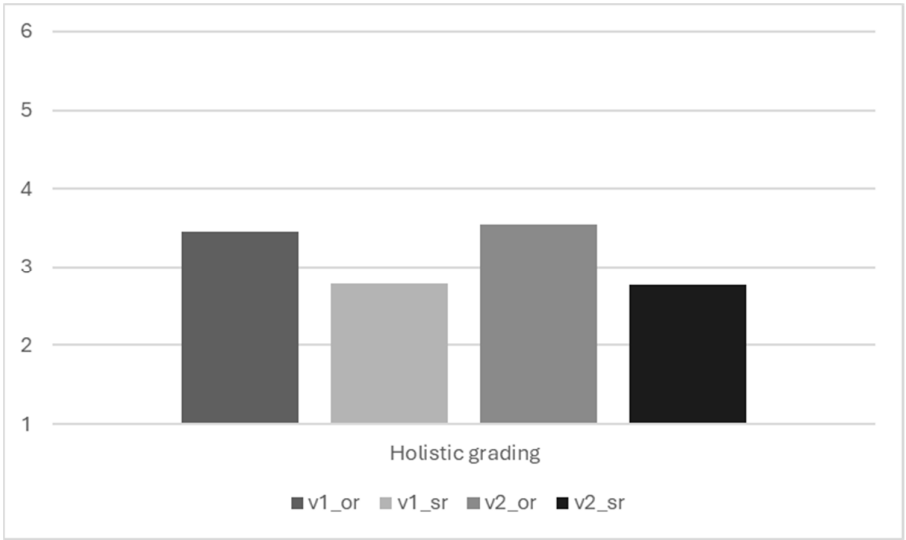

Figure 5. Mean values for holistic gradings in all four conditions (or = open reflection; sr = structured reflection; v1 = video 1; v2 = video 2).

3.3.4. Multidimensional Approach to Describing the Reflection Process

Figure 6 displays the mean values for the five reflection process dimensions for the four conditions. The evaluating dimension obtained the highest scores across all the conditions, while the naming alternatives dimension showed relatively lower scores. The dimensions of justification and self-reference in particular had remarkably low proportions.

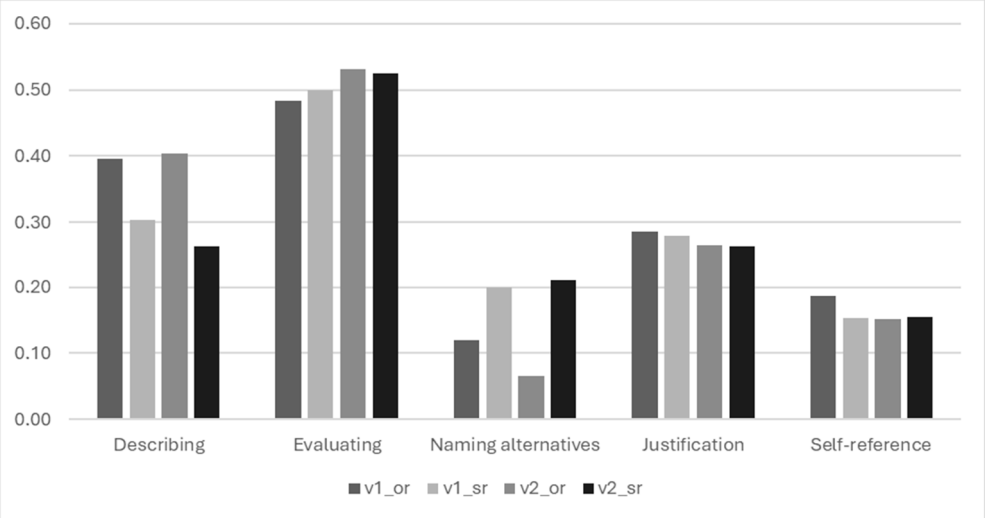

Figure 6. Proportions of the five reflection process dimensions in all four conditions (or = open reflection; sr = structured reflection; v1 = video 1; v2 = video 2).

The results demonstrated that significant differences were exclusively observed for the describing and naming alternatives dimensions (Figure 6, Table 4). All the contrast tests for the other reflection process dimensions—evaluating, self-reference, and justification—were not significant.

In the describing dimension, a significant and small overall effect for the difference between open and structured reflections was detected (see Table 4), with a larger proportion of describing elements in the open reflections. This effect was also evident when comparing the open and structured reflections on each of the two videos with each other, respectively. Regarding the potential training effect, there were no significant differences between the proportions of describing elements of the first and the second video settings overall, or when comparing both open and structured reflections with each other, respectively. However, a significant difference with a small effect size was evident between the structured reflection on the first video and the open reflection on the second video, indicating a higher proportion of describing statements in the open reflections on the second video.

For the naming alternatives dimension, a significant medium effect for the difference between open and structured reflections was detected. This significant difference could also be shown when comparing open and structured reflections of each of the two videos, respectively, with each other. The naming alternatives dimension was higher in structured reflections compared to open reflections. Regarding the potential training effect, there were no significant differences between the naming alternatives proportions for the first and the second video overall, as well as when comparing both open and structured reflections with each other, respectively. Nevertheless, a significant difference with a small effect size was evident between the structured reflection on the first video and the open reflection on the second video, reflecting a lower proportion of named alternatives in the open reflections on the second video.

4. Discussion

This study presented a multidimensional approach to describing the reflection process. In an exploratory manner, a comparison was made between this multidimensional approach and already existing measures of reflection quality. Furthermore, the relationships between reflection-related dispositions (attitudes towards reflection, reflection-related self-efficacy) as well as personal variables (professional experience, need for cognitive closure), the five reflection process dimensions (multidimensional approach), and three reflection quality measures (depth of reflection, and breadth of reflection, as well as a holistic grading) were exploratively examined to determine what teachers need to facilitate effective reflection. Last but not least, this study also examined the potential effects of low-threshold interventions (repeated practice and structuring prompts) to foster teachers' reflection process and reflection quality.

4.1. Operationalization of the Reflection Process as a Multidimensional Construct

In accordance with the integrated model of reflection on teaching (Arendt et al., 2025), five dimensions were identified as components of the reflection process: describing, evaluating, naming alternatives, justification, and self-reference.

The results indicate that the operationalization of the reflection process as a multidimensional construct seems plausible: The results demonstrate that the five dimensions exhibited low to medium intercorrelations, which indicates there is potential for several dimensions of reflection. This approach allows the reflection process to be interpreted in terms of five different, interrelated dimensions. As a result, the present study supports the multidimensional structure of the reflection process, as derived from Arendt et al.'s (2025) integrated model of reflection. However, an empirical test of the model is still required.

Noticeably, evaluating achieved the highest level of reflection across all the dimensions, which can be explained by the fact that teachers tend to adopt an evaluative attitude when reflecting on other people's teaching videos (Kleinknecht & Schneider, 2013). The participants' judgmental attitude may in part be attributed to the fact that the teachers in the videos looked rather young (which implies they are relatively new teachers). Some participants referred to this aspect in their reflections and attributed less professional experience to the teachers, justifying their criticism of some of the teachers' behavior. In a subsequent study, it would be beneficial to introduce a range of expertise among the video models in order to ascertain the extent to which the experience attributed to the teachers in the video influences the proportion of the evaluating dimension. In addition, the models could be varied to show either optimal or dysfunctional teacher behavior. The extent to which reflections differ when reflecting on functional or dysfunctional teacher behavior could be relevant for further research. This is particularly important for promoting reflection processes and should be considered in intervention studies.

As a result of the high levels of evaluating, other aspects of reflection, such as justifications or constructing a self-reference, tended to be neglected, which may explain the low scores in these two dimensions.

These results suggest that the multidimensional approach can be useful, especially for future intervention studies. Specifically, the construction of self-reference and the use of justifications in the reflection process should be a focus of future research, as these were weakly developed in this study. The examination of self-reference in reflections is a current focus of research; for instance, Merkert et al. (2023) examined various types of self-references created in reflections. Nonetheless, research in this domain remains in its nascent stages. Therefore, further clarification of the content of these dimensions and how to promote self-references is required. Studies on justifications have already been conducted. The focus was primarily on the integration of information texts into

reflections (Hartmann et al., 2021). In addition, Schellenbach-Zell et al. (2023) demonstrated that prompts support the integration of theoretical knowledge into reflections and that expert feedback on the reflections generates self-references. It would be valuable for future research to investigate to what extent the integration of information tends to occur more spontaneously, i.e., without prior input through information texts, in order to depict a more practical situation.

4.2. Comparison of the Multidimensional Approach with Different Measures of Reflection Quality

The first Research Question aimed to analyze the relationship between all measures of reflection quality and the multidimensional approach dimensions. Moderate interrelations between all measures were assumed.

The correlation between the depth of reflection and the breadth of reflection, although positive, was only weak. Thus, the depth of reflection and the breadth of reflection seem to primarily relate to different reflective elements. Other studies have also reported only a moderate correlation between reflection depth and breadth (Klempin & Rehfeldt, 2023), underlining this finding.

A negative correlation was evident between holistic grading and both the depth of reflection and the breadth of reflection. Furthermore, a negative correlation with a medium effect size was observed with some multidimensional approach dimensions, specifically with the dimensions of naming alternatives, justification, and self-reference. Consequently, a deeper and broader reflection was associated with a better holistic grading. With regard to the multidimensional approach, higher proportions of named alternatives, justifications, and self-references indicated better holistic grading.

Such a relationship was also recognizable in the connection between the depth of reflection and the multidimensional approach dimensions. The depth of reflection was primarily associated with a lower level of description and a higher level of named alternatives in a reflection. Based on these findings, it could be assumed that such describing and naming alternative proportions might be associated with higher reflection quality. Indeed, a comparable finding was reported in the study by Wulff et al. (2022). However, it must be noted that the correlations must be interpreted cautiously, as they can be explained by the definition and operationalization of the depth of reflection, as the presence of naming alternatives is a criterion for achieving the highest value in the depth of reflection.

A particularly strong correlation was evident between the breadth of reflection and a high level of self-reference. Therefore, it can be assumed that self-reference plays an important role in the assessment of reflection quality. This assumption is predicated on the premise that self-reference facilitates the differentiation of reflection from other constructs, such as the analysis of teaching situations (Arendt et al., 2025; Lenske & Lohse-Bossenz, 2023). Despite this conclusion, it must also be noted that this correlation, although strong, needs to be interpreted with caution due to the operationalization of both measures; the very definitions of the measures mean that the correlation was to be expected.

It is noteworthy that the reflections in this study received relatively high ratings considering the depth and breadth of the reflection measures. The values of these variables were elevated and exceeded the mean value. However, the multidimensional analyses revealed that the reflections exhibit weaknesses. When comparing the breadth of reflection and the dimensions of justification and self-reference, it could be assumed that, when evaluating the breadth of reflection, the quality of the reflections was overestimated. A comparable phenomenon can be observed concerning the depth of reflection. In this measure, reflections already receive the best rating when only one alternative is mentioned in the entire reflection, which represents the already mentioned critique of such a rating (see Section 1.2.1). However, it is important to note that this reflection does not address all

the elements relevant to the subject matter. In comparison, the multidimensional approach demonstrates the reflection process in a more differentiated way, by examining individual dimensions in relation to the overall reflection.

In conclusion, it can be stated that the various measures for assessing reflection quality and the multidimensional approach dimensions were interrelated. However, the findings of this study demonstrated that the multidimensional approach provides a more comprehensive understanding of the structure and nature of the reflections.

4.3. Relationships Between the Five Reflection Process Dimensions, Reflection Quality, Reflection-Related Dispositions, and Teachers' Personal Characteristics

Research Question 2 was aimed at identifying the relationships between the reflection-related dispositions, teachers' personal characteristics, dimensions of the reflection process, and reflection quality. It was assumed that positive attitudes, higher self-efficacy expectations, and more professional experience would correlate positively with reflection quality and the dimensions of justification and self-reference. Further, it was expected that a high need for cognitive closure would be related negatively to reflection quality. For the dimensions describing, evaluating, and naming alternatives, correlations were assumed but without a presumed direction of effect.

The finding that reflection-related self-efficacy correlated negatively with the five reflection process dimensions (multidimensional approach), as well as with depth of reflection, contradicts previous research findings (Lohse-Bossenz et al., 2019). Issues with the operationalization of the construct can be disregarded in this context, as reflection-related self-efficacy demonstrated a relatively strong correlation with attitudes towards reflection, which is consistent with the anticipated findings. This outcome may be ascribed to the small, relatively homogenous sample with regard to the reflection-related dispositions. Therefore, it is evident that further large-scale studies are required on the relationships between the reflection process dimensions, reflection quality, and the reflection-related dispositions. Conversely, the outcomes of the reflection-related self-efficacy evaluation may have been impacted by the fact that this variable was only assessed following completion of the reflection tasks. It is conceivable that the participants' self-efficacy expectations were downwardly revised in response to their self-assessed performance in the reflection task.

Concerning attitudes towards reflection, we also did not find the expected positive correlations. Despite the teachers' reporting extremely positive attitudes towards reflection, this is not shown in correlations between this variable and the multidimensional approach dimensions, or the depth of reflection, breadth of reflection, and the holistic grading measures. This incongruity remains unexplained by existing research findings (Göbel & Neuber, 2022; Göbel et al., 2022) but suggests the presence of a gap between teachers' self-perceptions and their actual reflection practices. This phenomenon may also be explained by the attitude–behavior gap. Alternatively, social desirability may have influenced the classification of the items for this variable. The elevated mean value of the scale, in conjunction with the minimal level of rigor, lends credence to this hypothesis. It is recommended that further research be conducted to investigate the correlations between the quality of reflection and attitudes towards reflection, utilizing a more extensive and heterogeneous sample.

The absence of a recognizable positive correlation between the measures of reflection quality and professional experience, and especially the negative correlation between the justification dimension and professional experience, is particularly noteworthy in the context of the expertise paradigm. It was hypothesized that a higher level of professional experience would be associated with a higher quality of reflection, as well as higher values for the dimensions of justification and self-reference. However, teaching experience alone was not sufficient to ensure the initiation of higher-quality reflection processes or a more

elaborate reflection due to justifications and self-references. Instead, numerous years of reflective practice, underpinned by a restructured knowledge base, did not constitute the assumed requisite for attaining higher-quality reflection processes (Gruber, 2021). Consequently, the hypothesis that more experienced teachers are superior reflectors was not supported by our data. Instead, it is proposed that experienced teachers require recurrent opportunities to further develop their reflective competence and facilitate continuous in-service reflection. Therefore, reflection research can contribute to the professionalization of teachers because reflective competence does not appear to evolve independently; rather, it must be cultivated as part of the professionalization process.

With regard to the need for cognitive closure, the anticipated average negative correlation was observed to be confined to the measures of reflection quality (depth of reflection, breadth of reflection, and holistic grading). The need for cognitive closure was seemingly related to only one aspect of the reflection process, the evaluating dimension. Hence, this correlation was only revealed through the multidimensional approach and not by the quality measures, substantiating the value of this approach. Analysis of the correlations between the reflection process dimensions and other personal teacher variables that are related to reflection and reflection quality would appear to be a logical progression in this field of enquiry. Despite the exclusion of these variables from the reflection models (e.g., Hatton & Smith, 1995) and the integrated model of reflection employed in this study (Arendt et al., 2025), their consideration within the framework of the professionalization of teachers remains fruitful to the field of reflection research.

Taken together, the relationships between the reflection quality measures, the reflection process dimensions, and the reflection-related dispositions, as well as personal variables, were rather inconsistent and unexpected. However, other studies have also shown that reflection-related dispositions only explain a small proportion of the variance (Stender et al., 2021). Presumably, other dispositions exist that are more closely related to reflection quality, for example, motivation or task value, or knowledge about reflection. Moreover, these findings show that the multidimensional approach was particularly efficacious and more differentiated in enhancing understanding of the relationships between reflection processes and teachers' personal characteristics, as is illustrated, for example, by the findings regarding reflection-related self-efficacy. Despite the observed correlation between reflection-related self-efficacy and the depth of reflection, no correlation was found with the other reflection quality measures—breadth of reflection or holistic grading.

4.4. Low-Threshold Promotion of Reflection Quality

Research Question 3 addressed the extent to which structuring prompts and repeated practice fostered reflection quality and altered the composition of the reflection process from a multidimensional perspective. In all the quality measures and reflection process dimensions, the hypothesized training effect was not demonstrated in either the open or structured reflections comparing video 1 and video 2, or in the overall comparison. Furthermore, the hypothesized immediate effect of the structuring prompts at the first measurement point on the subsequent open reflection at the second measurement point could not be substantiated. For the depth of reflection, holistic grading, and two multidimensional approach dimensions, an opposing effect was observed: the dimensions describing and naming alternatives revealed that the descriptions in the reflections increased in the open reflections on the second video compared to the structured reflections on the first video, while the proportion of named alternatives decreased. In the comparison of open and structured reflections, an enhancement in the quality of reflection in the structured reflections was hypothesized. However, significant differences were only observed for the describing and naming alternatives dimensions, with the proportion of descriptions in

the structured reflections decreasing and the proportion of alternatives named increasing. These differences were also evident in the depth of reflection and holistic grading.

Here again, it becomes evident that the multidimensional approach enables a more differentiated view of the reflection process, as it provides a more substantial amount of information regarding the structure of the reflections.

As the results demonstrate, the values for the justification and self-reference dimensions, which were not explicitly prompted, were not influenced by the prompts. The prompts exhibited the greatest changes in the describing and naming alternatives dimensions, while they had no effect on evaluating. Notably, the prompt for evaluating the teaching situation may require further refinement and revision for future research, as it may not have aligned with the expectations for the participants. Conversely, the reflections exhibited a substantial degree of evaluation, which suggests that supplementary support for this dimension might not have been essential.

Overall, the training using the two teaching videos appears to have been ineffective. This may be attributed to the complexity and time-consuming nature of the reflection tasks. There may have been a lack of motivation to engage with the same work task again, particularly for the second video. Seemingly, the prompts for the first video were not sufficient to generate a learning effect for the open reflection on the second video. This suggests that the low-threshold approach used was not effective. It would, therefore, be beneficial for future studies to present a more realistic setting that is more practical and motivating for the participants, thereby making the reflection task appear more meaningful for them.

Another aspect that can be addressed is the extent to which these low-threshold intervention measures constituted an effective intervention for the participants. The results also hint at a potential discrepancy between self-perceived competence and actual performance. It seems that the teachers were unable to demonstrate complete reflective performance when they were asked to reflect openly. It can be assumed that, in this context, they lacked the ability to utilize their existing reflective skills in a targeted manner. However, the structured reflections indicated that the participants possessed the ability to effectively manage the demands associated with reflection. Guided by the structuring prompts, the majority of the participants demonstrated a high level of engagement with the sub-processes of describing, evaluating, and naming alternatives, although this engagement was not necessarily apparent in their open reflections. The discrepancy between competence and performance represents a pivotal point for the promotion of reflective competence. However, the efficacy of the low-threshold approach in this study has been mixed, as evidenced by the findings. To enhance the effectiveness of low-threshold approaches in this context, it is essential to address all five dimensions of reflection. The two dimensions, justification and self-reference, should also be more explicitly and directly promoted.

Adopting an alternative approach would also be conducive to the promotion of reflective competence. As demonstrated in this study, a discrepancy between performance and competence may exist, which suggests that support programs should be designed with the initial aim of analyzing the reflections provided by others rather than generating one's own reflections. In order to continue working on this discrepancy, reflective competence must be considered in a fine-grained manner, like with the multidimensional approach. In addition, reflection-related dispositions must also be considered, as proposed by reflection models (Arendt et al., 2025; von Aufschnaiter et al., 2019). While reflection-related self-efficacy or attitudes towards reflection are frequently investigated (e.g., Göbel et al., 2022; Lohse-Bossenz et al., 2019), knowledge about reflection has rarely been included in prior research. However, Wulff et al. (2022) were able to show that the promotion of knowledge about reflection had a positive influence on the reflection process and, thus, on reflection

quality. Furthermore, training courses on reflective competence should incorporate more repetition to facilitate deeper and more sustained support. The utilization of prompts as a support measure has demonstrated efficacy in this and other studies (e.g., Schellenbach-Zell et al., 2023), so following this approach in subsequent training programs is recommended. It also appears rational to employ text vignettes in a targeted training program as opposed to videos, which engender a multitude of reflection opportunities, thereby reducing the variability in content and enhancing the comparability of reflections with respect to different quality dimensions.

4.5. Transfer of the Multidimensional Approach to the Reflection Process on Reflection Quality

The present study demonstrated that the multidimensional approach has the potential to provide a more structured description of reflection processes. The study also demonstrated the limitations of existing reflection quality measures. It can therefore be posited that the multidimensional approach has the potential to function not only as a means of describing reflection processes but also as a foundation for enhancing existing reflection quality measures and re-evaluating the assessment of reflection quality. Thus, the question arises whether the multidimensional approach could also constitute a reflection quality measure.

There is a major point of criticism as to why the multidimensional approach cannot already be regarded as an extended quality of reflection measure. The multidimensional approach is primarily based on the presence of constituent parts of the reflection process. A precise categorization of the correctness and quality of verbal statements in the reflections based on the various dimensions is missing in this approach. Therefore, it would be beneficial to further differentiate the single dimensions and define specific quality characteristics or levels for each one. In addition, mixed-methods approaches appear to be a promising strategy for achieving a more profound understanding of the complex, multidimensional structure of reflection and the reflection quality.

Nevertheless, it could also be argued that existing quality measures were unable to make a comprehensive and elaborate statement on the quality of reflections. As stated above (see Section 1.2), measures on the depth or breadth of reflection mostly refer to reflection quality as a unidimensional concept. Mapping reflection quality to one overall value eases its quantitative analysis, but it always faces the critique of a loss of information.

The multidimensional approach must not be considered a genuinely new measure of reflective quality. Rather, it should be regarded as an integrative measure that combines existing measures and expands them using further elements (e.g., a further differentiation of the dimensions). Consequently, the existing quality measures for reflection should be integrated into the multidimensional approach. The concept of the depth of reflection is addressed in the dimensions of describing, evaluating, and naming alternatives. The relations between describing and naming alternatives and the depth of reflection in this study highlight how the multidimensional approach could serve as a quality measure. But due to the inherent nature of multidimensionality, it is possible to focus on all three sub-processes (and their proportions) instead of focusing only on the highest level achieved. For the breadth of reflection, different elements have been grouped under this category. In the multidimensional approach, all these elements could be explained as justifications or self-references. Consequently, a more differentiated assessment of this category is possible. The relationships between the breadth of reflection and self-reference in this study underpin this assumption. Nonetheless, it is recommended that subsequent studies focus on a more detailed examination of the reflection process, with a particular emphasis on differentiating between the single dimensions across varying quality levels.

Further discussion is required on the manner in which quality can be interpreted on the basis of the multidimensional approach. Due to the results of the low-threshold intervention, several first assumptions can be made: First, if prompts contribute to the promotion of reflection processes (e.g., Weber et al., 2022), then the consideration of the multidimensional approach in this context allows for initial conclusions regarding reflection quality. The prompts resulted in a significant increase in the proportion of named alternatives and a decrease in the proportion of descriptions. The number of evaluations, however, remained largely unchanged. From a normative perspective, it could therefore be hypothesized that high-quality reflections are associated with a reduced number of descriptions and an increased number of alternatives. Second, a further assumption regarding the multidimensional approach could be that the dimensions of the three sub-processes are present in a balanced ratio. This would result in the dimensions being equalized in their relationship to each other. In order to test this hypothesis, it would be necessary to explore the reflection process in greater depth, to ascertain how the sub-processes in the reflections build on each other in terms of the content of reflections. Thirdly, the presence of justifications and self-references in this study was found to be minimal, and they did not appear to be influenced by the interventions. However, for high-quality reflections, it can be assumed that these should be highly prominent, as they represent more sophisticated and diversified reflection processes. In accordance with the prevailing normative assumptions about reflection, it could be imperative that all sub-processes are accompanied by substantiated justifications and self-references, as part of a comprehensive, high-quality reflection. If necessary, these only make sense in a reflection if they are truly relevant for the consequences that arise for one's own practice. In this sense, the consideration of a "filter" for reflections might be reasonable, in which, for a high-quality reflection, the reflection content that is important for one's learning process is first selected before the reflection process has to be completed in full. Consequently, further studies should be conducted to examine such relationships, with a view to refining the concept of quality for reflection. As demonstrated, the multidimensional approach has the capacity to function as a foundational framework for the advancement of the prevailing measures of reflection quality.

4.6. Limitations and Implications

The present study was conducted using a within-subjects design. Despite the implementation of counterbalancing in the video sequence, the sequence of open and structured reflection did not vary. From a pedagogical perspective, this approach is not optimal, as the prompts could have influenced the open reflection. Consequently, the available data are nested, which limits the analysis and interpretability of the results. It is also important to note that the resulting correlations must be interpreted with caution due to the limited sample size. In a different study design, with a larger sample, it would have been possible to consider the variables as between-subjects factors, something that was not possible in the present design. However, with larger samples, a fine-grained analysis of the different reflections is extremely time-consuming. Notwithstanding the design limitations, the exploratory character of this study helps us to understand reflection processes in more detail and evaluate reflection quality in a more comprehensive way. Even though this study was conducted with only 29 participants and consequently does not permit the formulation of generalizable statements, it nevertheless provides valuable initial insights into the advantages of taking a multidimensional perspective on reflection processes and reflection quality.

A further limitation of the study design is that justifications and self-reference in the structured conditions were implicitly rather than explicitly prompted. This may also explain why the low-threshold intervention did not work well for these dimensions. Future

works should develop and test prompts that explicitly target these two dimensions, as they are crucial for the promotion of reflection and reflection quality.

As previously mentioned, the small sample size limits the generalizability of the results. Additionally, the composition of the sample may contribute to this limitation. In the present sample, the teachers participated voluntarily and without receiving any remuneration. It is possible that mainly teachers who consider reflection to be important self-selected to participate in this study. Furthermore, the extent to which the observed effects can be considered verifiable in the long term remains uncertain. The intervention measures were of a short-term nature, and open and structured reflections followed each other directly, and were repeated immediately afterwards with the second video. Consequently, the results cannot be interpreted in terms of their sustainability.

Another limitation of this study is the use of two teaching videos. Reflecting on a video of someone else's teaching is not the same as reflecting on one's own teaching (Kleinknecht & Schneider, 2013). Nevertheless, videos of someone else's teaching are a good starting point for experimental studies because it is difficult to produce individual videos of one's own teaching for each participant and to establish standardized conditions for the associated reflections. Furthermore, individual video effects can be dismissed, as the videos were selected according to similarity and presented in a counterbalanced order. Despite this, further studies would benefit from a more thorough examination of the video material. For instance, the evocation of different reflections by different types of video material could be investigated.

5. Conclusions

This study demonstrated the crucial importance of ongoing efforts to further develop the construct of reflection in the context of teacher education. The assessment of the reflection process can be supported by employing the integrated model of reflection (Arendt et al., 2025), as presented here. Thus, a substantial foundation for the advancement of reflection in the field of teacher education can be established. Assessing the quality of reflection is also important for practitioners. The question is often asked whether a reflective practitioner is genuinely a better practitioner. This disregards the fact that the quality of reflection is also a crucial element in the development of one's own practice. Furthermore, based on this assessment of the quality of reflection, in the context of the multidimensional approach presented here, it is possible to develop targeted support to help (future) teachers enhance their reflection skills. This, in turn, will enable them to utilize the positive effects of reflection in their practice. Here, future research should start by further clarifying the process of reflection, in particular, the individual dimensions of reflection and their interplay. Additionally, support programs should be individually adapted to the needs of teachers and specifically tailored to their previous reflection quality.

Author Contributions: Conceptualization, K.A., L.S., A.F., R.B., and R.S.; methodology, K.A., L.S., A.F., R.B., and R.S.; formal analysis, K.A. and L.S.; investigation, K.A.; data curation, K.A.; writing—original draft preparation, K.A.; writing—review and editing, K.A., L.S., A.F., R.B., and R.S.; visualization, K.A.; supervision, L.S., A.F., R.B., and R.S.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Graduate Program of the Center for Teacher Education (GRA-PRO SAAR) at Saarland University.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Faculty of Empirical Human Sciences and Economics, Saarland University (24-35, 18 November 2024).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request due to privacy reasons.

Acknowledgments: We would like to thank Annika Fendel and Marieke Reimers for their assistance with the qualitative data analysis.

Conflicts of Interest: The authors declare no conflicts of interest.

References

Aeppli, J., & Lötscher, H. (2016). EDAMA—Ein rahmenmodell für reflexion [EDAMA—A framework model for reflection]. *Beiträge zur Lehrerinnen- und Lehrerbildung*, *34*(1), 78–97. [CrossRef]

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Erlbaum.

Arendt, K., Stark, L., Friedrich, A., Brünken, R., & Stark, R. (2025). Reflexion von unterricht in der lehrkräftebildung—Ein scoping-review [Reflection on teaching in teacher education—A scoping review]. *Unterrichtswissenschaft*. [CrossRef]

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, *35*(5), 463–482. [CrossRef]

Bradbury, O., Fitzgerald, A., & O'Connor, J. (2020). Supporting pre-service teachers in becoming reflective practitioners using conversation and professional standards. *Australian Journal of Teacher Education*, *45*, 18–34. [CrossRef]

Cengiz, C. (2020). The effect of structured journals on reflection levels: With or without question prompts? *Australian Journal of Teacher Education*, *45*(2), 22–43. [CrossRef]

Clarà, M. (2015). What is reflection? Looking for clarity in an ambiguous notion. *Journal of Teacher Education*, *66*(13), 261–271. [CrossRef]

Clarà, M., Mauri, T., Colomina, R., & Onrubia, J. (2019). Supporting collaborative reflection in teacher education: A case study. *European Journal of Teacher Education*, *42*, 175–191. [CrossRef]

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge. [CrossRef]

Fazio, X. (2009). Teacher development using group discussion and reflection. *Reflective Practice*, *10*, 529–541. [CrossRef]

Fränkel, S., Ferencik-Lehmkuhl, D., & Schroeder, R. (2022). Wie reflektieren studienanfänger*innen inklusiven unterricht? Ergebnisse einer qualitativen studie zur reflexionskompetenz von lehramtsstudierenden [How do first-year students reflect on inclusive teaching? Results of a qualitative study on the reflective competence of student teachers]. *Lehrerbildung auf dem Prüfstand*, *15*(1), 195–215.

Göbel, K., Bönte, J., Gösch, A., & Neuber, K. (2022). The relevance of collegial video-based reflection on teaching for the development of reflection-related attitudes. *Teaching and Teacher Education*, *120*, 103878. [CrossRef]

Göbel, K., & Neuber, K. (2020). Einstellungen zur reflexion von angehenden und praktikzierenden lehrkräften [Attitudes towards reflection of prospective and practising teachers]. *Empirische Pädagogik*, *34*(1), 64–78.

Göbel, K., & Neuber, K. (2022). Verändern sich reflexionsbezogene einstellungen von studierenden nach der nutzung von schüler-rückmeldungen im praxissemester? Befunde einer interventionsstudie [Do students' attitudes towards reflection change after using student feedback in the practical semester? Findings of an intervention study]. *Zeitschrift für Erziehungswissenschaft*, *25*(3), 721–744. [CrossRef]

Gruber, H. (2021). Reflexion. Der königsweg zur expertise-entwicklung [The royal road to expertise development]. *Journal für LehrerInnenbildung*, *21*(1), 108–117. [CrossRef]

Hartmann, U., Kindlinger, M., & Trempler, K. (2021). Integrating information from multiple texts relates to pre-service teachers' epistemic products for reflective teaching practice. *Teaching and Teacher Education*, *97*, 103205. [CrossRef]

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33–49. [CrossRef]

Hemmer, M., Holodynski, M., & Möller, K. (2018a). *Geographie—Klasse 5—Instruktionen der arbeitsphase—Einzelstunde 3—Clip 4* [*Geography—Class 5—Instructions for the work phase—Individual lesson 3—Clip 4*] [Video]. ProVision Videoportal. Available online: https://vsso.uni-muenster.de/ProVision/video/#GEO_K5_LK6_3ES_URL_Clip4 (accessed on 5 January 2023).

Hemmer, M., Holodynski, M., & Möller, K. (2018b). *Geographie—7. Klasse—In den savannen—Doppelstunde 1—Clip 5* [*Geography—7th grade—In the savannahs—Double lesson 1—Clip 5*] [Video]. ProVision Videoportal. Available online: https://vsso.uni-muenster.de/ProVision/video/#GEO_K7_LK2_1DS_SAV_Clip5 (accessed on 5 January 2023).

Jung, J., Lu, Y., & Ding, A. (2021). How do prompts shape preservice teachers' reflections? A case study in an online technology integration class. *Journal of Teacher Education*, *73*, 301–313. [CrossRef]

Kleinknecht, M., & Gröschner, A. (2016). Fostering preservice teachers' noticing with structured video feedback: Results of an online- and video-based intervention study. *Teaching and Teacher Education*, *59*, 45–56. [CrossRef]

Kleinknecht, M., & Schneider, J. (2013). What do teachers think and feel when analyzing videos of themselves teaching and other teachers teaching? *Teaching and Teacher Education*, *33*, 13–23. [CrossRef]

Klempin, C., & Rehfeldt, D. (2023). How to promote and measure reflective skills in depth and breadth of English and physics teacher trainees. In L. Mientus, C. Klempin, & A. Nowak (Eds.), *Reflexion in der lehrkräftebildung. Empirisch—Phasenübergreifend—Interdisziplinär* (pp. 115–124). Universitätsverlag Potsdam.

Klug, J., Schultes, M.-T., & Spiel, C. (2018). Assessment at school—Teachers' diary-supported implementation of a training program. *Teaching and Teacher Education*, *76*, 298–308. [CrossRef]

Korthagen, F. A. J. (1999). Linking reflection and technical competence: The logbook as an instrument in teacher education. *European Journal of Teacher Education*, *22*(2–3), 181–207. [CrossRef]

Kruglanski, A. W. (2004). *The psychology of closed mindedness*. Psychology Press.

Kulgemeyer, C., Kempin, M., Weißbach, A., Borowski, A., Buschhüter, D., Enkrott, P., Reinhold, P., Riese, J., Schecker, H., Schröder, J., & Vogelsang, C. (2021). Exploring the impact of pre-service science teachers' reflection skills on the development of professional knowledge during a field experience. *International Journal of Science Education*, *43*(18), 3035–3057. [CrossRef]

Leijen, Ä., Valtna, K., Leijen, D. A. J., & Pedaste, M. (2012). How to determine the quality of students' reflections? *Studies in Higher Education*, *37*(2), 203–218. [CrossRef]

Lenske, G., & Lohse-Bossenz, H. (2023). Stichwort: Reflexion im pädagogischen kontext [Keyword: Reflection in a pedagogical context]. *Zeitschrift für Erziehungswissenschaft*, *26*, 1133–1164. [CrossRef]

Leonhard, T., Nagel, N., Rihm, T., Strittmatter-Haubold, V., & Wengert-Richter, P. (2010). Zur entwicklung von reflexionskompetenz bei lehramtsstudierenden [On the development of reflection skills in student teachers]. In A. Gehrmann, U. Hericks, & M. Lüders (Eds.), *Bildungsstandards und kompetenzmodelle. Beiträge zu einer aktuellen diskussion über schule, lehrerbildung und unterricht* (pp. 111–127). Klinkhardt.

Leonhard, T., & Rihm, T. (2011). Erhöhung der reflexionskompetenz durch begleitveranstaltungen zum schulpraktikum? Konzeption und ergebnisse eines pilotprojekts mit lehramtsstudierenden [Increasing reflective competence through accompanying events to the school internship? Concept and results of a pilot project with student teachers]. *Lehrerbildung auf dem Prüfstand*, *4*(2), 240–270.

Lohse-Bossenz, H., Schönknecht, L., & Brandtner, M. (2019). Entwicklung und validierung eines fragebogens zur erfassung reflexionsbezogener selbstwirksamkeit von lehrkräften im vorbereitungsdienst [Development and validation of a questionnaire to assess reflection-related self-efficacy of pre-service teachers]. *Empirische Pädagogik*, *33*(2), 164–179.

Mayring, P. (2022). *Qualitative inhaltsanalyse: Grundlagen und techniken* [*Qualitative content analysis: Basics and techniques*]. Beltz.

Merkert, A., Lohse-Bossenz, H., Neuber, K., & Lenske, G. (2023). Selbstbezug in videobasierten unterrichtsreflexionen von lehramtsstudierenden im Bachelor [Self-reference in video-based lesson reflections of student teachers in the Bachelor's programme]. *Zeitschrift für Erziehungswissenschaft*, *26*, 1259–1280. [CrossRef]

Neuber, K., & Göbel, K. (2018). *Schülerrückmeldungen zum unterricht und unterrichtsreflexion: Dokumentation der entwickelten erhebungsinstrumente im projekt "Schülerrückmeldungen zum unterricht und ihr beitrag zur unterrichtsreflexion im praxissemester (ScRiPS)"* [Student feedback on teaching and lesson reflection: Documentation of the survey instruments developed in the project "Student feedback on teaching and its contribution to lesson reflection in the practical semester (ScRiPS)"]. *Aktualisierte Skalenanalysen*. [CrossRef]

Nowak, A., Kempin, M., Kulgemeyer, C., & Borowski, A. (2019). Reflexion von physikunterricht [Reflection on physics lessons]. In C. Maurer (Ed.), *Naturwissenschaftliche bildung als grundlage für berufliche und gesellschaftliche teilhabe. Jahrestagung in kiel 2018* (p. 838). Gesellschaft für Didaktik der Chemie und Physik.

Petko, D., Schmid, R., Müller, L., & Hielscher, M. (2019). Metapholio: A mobile app for supporting collaborative note taking and reflection in teacher education. *Technology Knowledge and Learning*, *24*(4), 699–710. [CrossRef]

Prieto, L. P., Magnuson, P., Dillenbourg, P., & Saar, M. (2020). Reflection for action: Designing tools to support teacher reflection on everyday evidence. *Technology Pedagogy and Education*, *29*(3), 279–295. [CrossRef]

Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Psychological Science*, *17*(3), 249–255. [CrossRef]

Schellenbach-Zell, J., Molitor, A.-L., Kindlinger, M., Trempler, K., & Hartmann, U. (2023). Wie gelingt die anregung von reflexion über pädagogische situationen unter nutzung bildungswissenschaftlicher wissensbestände? Die bedeutung von prompts und feedback [How can reflection on pedagogical situations be stimulated using educational science knowledge? The importance of prompts and feedback]. *Zeitschrift für Erziehungswissenschaft*, *26*, 1189–1211. [CrossRef]

Schlink, S., & Walther, E. (2007). Kurz und gut: Eine deutsche Kurzskala zur erfassung des bedürfnisses nach kognitiver geschlossenheit [Short and good: A German short scale to assess the need for cognitive coherence]. *Zeitschrift für Sozialpsychologie*, *38*, 153–161. [CrossRef]

Seidel, T., Blomberg, G., & Renkl, A. (2013). Instructional strategies for using video in teacher education. *Teaching and Teacher Education*, *34*(1), 56–65. [CrossRef]

Stender, J., Watson, C., Vogelsang, C., & Schaper, N. (2021). Wie hängen bildungswissenschaftliches professionswissen, einstellungen zu reflexion und die reflexionsperformanz angehender lehrpersonen zusammen? [What is the relationship between professional knowledge in educational science, attitudes towards reflection and the reflective performance of prospective teachers?]. *HLZ*, *4*(1), 229–248. [CrossRef]

Ullmann, T. D. (2019). Automated analysis of reflection in writing: Validating machine learning approaches. *International Journal of Artificial Intelligence in Education*, *29*, 217–257. [CrossRef]

von Aufschnaiter, C., Fraij, A., & Kost, D. (2019). Reflexion und reflexivität in der lehrerbildung lehrerbildung [Reflection and reflexivity in teacher training]. *HLZ*, *2*, 144–159. [CrossRef]

Weber, K. E., Elstrodt-Wefing, N., & Hoge, K. (2022). Expertise- und stimulusbedingte unterschiede in schriftlichen videobasierten reflexionen [Expertise- and stimulus-related differences in written video-based reflections]. In E. Gläser (Ed.), *Reflexion und reflexivität im kontext grundschule, perspektiven für forschung, lehrer:innenbildung und praxis* (pp. 77–89). Klinkhardt.

Wulff, P., Buschhüter, D., Westphal, A., Nowak, A., Becker, L., Robalino, H., Stede, M., & Borowski, A. (2021). Computer-based classification of preservice physics teachers' written reflections. *Journal of Science Education and Technology*, *30*, 1–15. [CrossRef]

Wulff, P., Mientus, L., Nowak, A., & Borowski, A. (2022). Utilizing a pretrained language model (BERT) to classify preservice physics teachers' written reflections. *International Journal of Artificial Intelligence in Education*, *33*, 439–466. [CrossRef]

Wulff, P., Westphal, A., Mientus, L., Nowak, A., & Borowski, A. (2023). Enhancing writing analytics in science education research with machine learning and natural language processing—Formative assessment of science and non-science preservice teachers' written reflections. *Frontiers in Education*, *7*, 1061461. [CrossRef]

Wyss, C. (2013). *Unterricht und reflexion: Eine mehrperspektivische untersuchung der unterrichts- und reflexionskompetenz von lehrkräften* [*Teaching and reflection: A multi-perspective study of teachers' teaching and reflection skills*]. Waxmann.

Wyss, C., & Mahler, S. (2021). Mythos reflexion: Theoretische und praxisbezogene erkenntnisse in der lehrer*innenbildung [The myth of reflection: Theoretical and practical findings in teacher training]. *Journal für LehrerInnenbildung*, *21*(1), 16–25. [CrossRef]