

Beyond the surface: a qualitative study on professional vision of deep structures within geography education

Sebastian Streitberger, Ulrike Ohl

Angaben zur Veröffentlichung / Publication details:

Streitberger, Sebastian, and Ulrike Ohl. 2025. "Beyond the surface: a qualitative study on professional vision of deep structures within geography education." In Teacher professional vision: empirical perspectives; volume 2, edited by Rebekka Stahnke and Andreas Gegenfurtner, 150–67. London: Routledge.
<https://doi.org/10.4324/9781003370604-12>.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:

Deutsches Urheberrecht

Weitere Informationen finden Sie unter: / For more information see:

<https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/>



BEYOND THE SURFACE

A qualitative study on professional vision of deep structures within geography education

Sebastian Streitberger and Ulrike Ohl

University of Augsburg, Germany

Deep structures within geography education and professional vision

In the educational sciences, the discussion about significant characteristics of teaching quality and successful learning processes has established a distinction between so-called surface and deep structures of learning. Surface structures describe the immediately perceptible features of teaching, i.e. they can be observed directly (Lipowsky & Lotz, 2015, p. 195; Mehren & Mehren, 2020, p. 4). Surface structures include organisational forms, methods, and the type of cooperation (e.g. partner or group work). They are of great importance in the planning and teaching of lessons on an organisational level, although decisions about them do not directly lead to successful learning processes (Bohl, 2017, p. 262; Trautwein et al., 2018, p. 6).

Deep structures, on the other hand, largely determine the quality of learning and teaching (Kunter & Trautwein, 2013, p. 76; Mehren & Mehren, 2020, p. 4). They refer to the quality of the learners' interaction with the learning content and to the quality of the interaction between the participants of a given lesson (Trautwein et al., 2018, p. 8). They are considered more effective than the surface structures, especially with regard to the acquisition of domain-specific competencies (Klieme & Rakoczy, 2008, p. 228). Deep structures refer to features of lessons that are usually not immediately perceptible, however, and require some form of interpretation of what is happening in the classroom (Mehren & Mehren, 2020, p. 4; Trautwein et al.,

2018, p. 8). The key question is now, towards which characteristics do we then direct our interpretative view?

Here, professional vision (PV) comes into play. PV is understood as “the ability to notice and interpret relevant features of classroom situations” (Seidel & Stürmer, 2014, p. 741) and thus the ability to respond spontaneously to these situations. The ability to notice “is the ability to make connections between specific events and the broader ideas they present” (van Es & Sherin, 2002, p. 574), namely principles of teaching and learning. Van Es and Sherin (2002, pp. 574–575) as well as Seidel, Blomberg, and Stürmer (2010, pp. 297–298) add the component of knowledge-based reasoning (see also Sherin et al., Chapter 3 of Volume 1). By this, they mean the teachers’ ability to use their general but also, and particularly, their domain-specific knowledge in order to analyse what they notice. They also suggest a multilevel mode of knowledge-based reasoning, depending on descriptive, explanatory, and predicting elements. Since teachers’ knowledge bases and their PV abilities may differ, specific and often highly complex classroom situations may be processed very differently, too (Ansorge & Leder, 2017, pp. 8–9). With professional vision in mind, we now must rephrase our key question from before: Which characteristics of the surface structures do we notice in order to draw knowledge-based inferences about deep structures?

According to Lipowsky and Lotz (2015, p. 188), it is particularly crucial to focus on so-called basic dimensions of high-quality teaching. These resemble strategies that promote the pupils’ processing and understanding of content. The first basic dimension is classroom management (Kunter & Trautwein, 2013, pp. 65, 76–77). This dimension predominantly centres around means of control, e.g. measures to prevent class interruptions or to make more effective use of time (Mehren & Mehren, 2020, p. 5). According to the model of teachers’ professional competence (Kunter et al., 2011), teachers have to draw entirely on their pedagogical knowledge (PK) to manage the classroom. Due to the lack of domain specificity, classroom management will not be included in this study as a result. However, for the other two dimensions, i.e. cognitive activation and constructive support of learners (Kunter & Trautwein, 2013, pp. 65, 76–77), teachers rely on comprehensive domain-specific content knowledge (CK) and pedagogical content knowledge (PCK). This is because cognitive activation, e.g. in the form of selecting and sequencing cognitively challenging tasks, cognitively stimulating classroom conversations, or cognitively challenging opportunities to practice (Mehren & Mehren, 2020, p. 5), can be realised only on the basis of domain-specific competencies. The same applies to measures of constructive support, such as a good structuring of lessons or differentiated feedback related to the pupils’ subject-specific performance (Mehren & Mehren, 2020, p. 5).

Consequently, Klieme and Rakoczy (2008, p. 229) emphasise the importance of every discipline to specify what cognitive activation or structuring really mean in their specific contexts. Hence, the presented study focuses on these two basic dimensions.

Cognitive activation basically describes the process of stimulating pupils to engage more profoundly with a specific task or content and, thus, it helps to deepen their knowledge (Kunter & Voss, 2011, p. 88). The importance of domain-specific CK and PCK becomes even clearer since cognitive activation can only be successful when connected to specific learning goals, the individual prior knowledge of each pupil and the attention to ongoing cognitive processes. In the context of a PV of deep structures, it is now important to notice specific situations in the surface structures that provide valuable information beyond the surface and on the more relevant deep structures. For cognitive activation, these noticeable surface structures could be situations in which teachers encourage their pupils to justify their statements or when being provocative in their own statements, for example. Pupil-centred examples would be when pupils ask content-related questions or compare and evaluate different solution paths (Lipowsky & Lotz, 2015, p. 192).

The other basic dimension on a deep structure level is the constructive support of learners. For the purpose of this study, this dimension has been split into three constructs: teachers dealing adequately with pupils' misconceptions, appropriate scaffolding, and fading strategies. For all three constructs, the importance of domain-specific CK and PCK applies as well. In order to diagnose misconceptions of their pupils, teachers need CK to be able to assess the correctness of statements, on the one hand, and PCK about the acquisition of misconceptions, on the other. PCK is also relevant when it comes to finding solutions for appropriately supporting conceptual change. On a surface structure level, teachers could notice statements of their pupils, for example, in which they find clues about the misconceptions which are on a deep structure level.

Scaffolding represents another strategy to ensure the quality of learning processes at the level of deep structures. The term refers to instructional actions taken by the teacher to help close the gap between pupils' current competencies and the learning goal (or zone of proximal development) according to Rosenshine (1996, p. 202). Scaffolding measures include assisting pupils with their tasks, e.g. by simplifying the complexity of tasks (Kunter & Voss, 2011, p. 90). Likewise, scaffoldings can be interventions in case of comprehension deficits in which teachers help restructure the learning process (Kunter & Voss, 2011, p. 90). Scaffolds can either be situational or anticipated before the actual instructional situation as part of lesson planning (Arnold et al., 2016, pp. 25–26). With regard to PV, situational scaffolding is more in focus. A central component of scaffolding is fading (Arnold et al., 2016, pp. 25–26). Fading describes the absence of the teacher as a conscious decision in terms of constructive learning support (Lajoie, 2005, p. 548).

Deciding which scaffolding strategies are appropriate for which pupil in which situation can be done only based on an adequate assessment of the pupils' learning processes and subject-related performance in the specific situation in the classroom. Thus, again, domain-specific expertise is necessary. Surface structure indicators relevant to the deep structures of learning might be once more the teachers' and pupils' statements. After noticing them, these surface structures could be interpreted with regard to whether the teachers' support is constructive or misses its mark. Regarding fading, the lack of a statement by the teacher could be a surface structure that might be indicative of the deep structures and reasoned about further.

Although domain-specific knowledge seems to be essential, the differentiation and investigation of high-quality characteristics of domain-specific instruction has not yet been addressed systematically, nor is it evidence-based in geography education research (Mehren & Mehren, 2020, p. 9). Nonetheless, it seems obvious that strategies taken by the teacher to cognitively activate and constructively support the pupils should be directed towards the goal that pupils acquire subject-relevant competencies.

Therefore, we argue for educational deep structures as a particularly relevant focus for PV research since such structures are, as shown earlier, immensely important for successful learning processes yet challenging to observe in teaching (Mehren & Mehren, 2020, p. 4). To date, there have been few empirical studies on professional vision in geography education. In a qualitative study, Scholten et al. (2019) explored geography teachers' ability to notice classroom incidents that are critical for student learning. Scholten and Sprenger (2020) compared the noticing skills of geography teachers at three different stages of their professional development. A qualitative content analysis of teachers' think-aloud responses to video vignettes of geography teaching showed that the three groups differed in terms of the number of incidents noticed, their interpretation of the situations, and their decisions regarding adequate educational actions. To what extent geography teachers or future geography teachers possess skills of professional vision related to deep structures or how these skills can be specifically trained, has not yet been investigated in geography education research. Due to the great relevance of these skills, the study presented here begins precisely there, focusing on the skills of future geography teachers.

Research questions

The importance and challenge for teachers to notice and interpret educational interactions in a knowledge-based manner as described earlier is the key fundament of our research interest as well as fostering these abilities before the teacher trainees leave university for the classroom. In this study, we will focus on noticing and interpreting the highly relevant deep structures. Therefore, we ask the following question:

RQ 1: To what degree are future geography teachers capable of noticing and reasoning about geographical lessons before and after an intervention with regard to deep structures?

We want to answer this research question by answering the following sub-questions, each related to the situation before and after the intervention:

RQ 1.1: How many classroom situations are future geography teachers able to notice in a given video vignette?

RQ 1.2: To what extent do they notice surface structures and make inferences about deep structures?

RQ 1.3: To what extent do they notice and – in a knowledge-based manner – reason about constructs at the level of domain-specific deep structures which are elaborated in the intervention (i.e. cognitive activation, misconceptions, scaffolding, fading)?

Research design

The design of this research project includes two major components: an intervention and two measurements. The study follows a qualitative pre-post design (Perels & Otto, 2010, p. 253) and includes 45 future geography teachers who were sampled via a criteria-based purposeful approach, with criteria such as different types of school (e.g. primary, secondary) or different school subjects besides geography (Flick, 2019, pp. 155–156). The longitudinal design, which was selected for this study, is particularly suitable for documenting and analysing changes (Flick, 2019, pp. 183–184) as is required by the research questions. Measurements take place at two points in time (t_0 and t_1) and use an instrument which was developed specifically for this study: a non-standardised projective questionnaire (Döring & Bortz, 2016, pp. 401–402) in which the future teachers respond to a 10-minute video of a geography lesson and this task: “Analyse the video clip of a geography lesson with regard to situations you deem significant for successful learning.” The video vignette shows a segment of a geography lesson which addresses desert oases, their specific sources of water supply, and economic division in touristic and/or agricultural importance. In this segment, a group of four pupils works on several tasks. The pupils interact both with each other and with a teacher, who stands next to the group table. Both t_0 and t_1 use the same video.

Unlike with item-based questionnaires, a very open response format is chosen during the measurements in order to elicit the active analytical vocabulary of the participants, their ability to notice and reason about events freely and, finally, to allow for an individually diverse analysis, which may be expected

due to the intentionally high complexity of the videotaped classroom situation. The participants were given 40 minutes for their analysis of the vignette and could stop the video at any time during this time frame.

The data from the projective questionnaire are qualitatively analysed according to Kuckartz (2018), with the unit of analysis consisting of every single situation of the educational setting which is rated relevant to learning processes by the future teachers, i.e. every time the participants *notice* something which is relevant to them. With the given task, everything participants write down can be considered relevant to them.

The deductive categories of the content analysis are based on theoretical assumptions and empirical findings about successful domain-specific and non-domain-specific learning processes (as described in Table 9.1). Table 9.1 shows the operationalisation of the two central codes of “surface structure” and “deep structure” from the coding manual. If the distinction between the two is unclear, a third, “unclear” code is used.

TABLE 9.1 Example of the category system from the coding manual

<i>Code</i>	<i>Code Definition</i>	
	<i>Code Description</i>	<i>Application</i>
Surface Structure	<p>This code includes comments that address features of teaching that can be assigned to the surface structure.</p> <p>Surface structure includes forms of organisation (e.g. class teaching, remedial teaching), methods (e.g. project work, cooperative learning, direct instruction) and social form like single or group work (Trautwein et al., 2018, pp. 6–8). Features of the surface structure can be observed comparatively well – visually or auditorily (Bohl, 2017, pp. 262).</p> <p>Analyses of surface structures include utterances by the teacher or the pupils as well as descriptions of actions without a reference to cognitive (learning) processes being established.</p>	<p>Very descriptive reproductions of observable (and obvious) actions or statements are usually coded here: Indicators are descriptions of actions or quotes from the videotaped lesson (e.g. if a teacher asks a question, this is directly observable). Mental (teaching-learning) processes do not occur or are not focused on.</p>

(Continued)

TABLE 9.1 (Continued)

Code	Code Definition	
	Code Description	Application
Deep Structure	<p>This includes comments that address features of teaching that can be assigned to the deep structure.</p> <p>Deep structures include features of teaching-learning processes such as the quality of classroom management, the cognitive activation of pupils, or the constructive support of learning processes by the teacher (Trautwein et al., 2018, pp. 8–9). Features of deep structures cannot usually be observed directly (Trautwein et al., 2018, p. 8).</p> <p>Teaching-learning processes include situations both at the beginning of the teaching-learning process (e.g. activation), during it (e.g. pupils reflect on a question) and at its end.</p>	<p>As soon as there are explicit references to teaching-learning processes, the processing and understanding of learning content by the pupils and/or the way of dealing with learning content, this code is applied. Clues can be ideas or other mental products of pupils, cognitive processes, indicators for knowledge, causes for thought, misconceptions, comprehension processes, learning progress and achievement, scaffolding measures, cognitive activation (e.g. through provocative questions), etc.</p> <p>The analysis goes beyond the mere reproduction of (obvious) actions or statements.</p>

Important deductive categories are the aforementioned constructs related to the basic dimensions of high-quality teaching: Cognitive activation corresponds to the second basic dimension, i.e. also cognitive activation, as well as scaffolding, fading, and dealing with misconceptions corresponding to the third basic dimension, i.e. constructive support of learners. If participants noticed more than one of these four constructs in a single situation, the unit of analysis is accordingly coded repeatedly. At the same time, every code (e.g. misconceptions) is divided into two disjoint subcodes, of which no more than one can be used for a single unit of analysis: one code for a more explicit and somewhat technical form of reasoning about the construct (referred to as *explicit*, i.e. *misconceptions–explicit*); one for a vaguer, more descriptive and less technical way of reasoning (referred to as *implicit*, i.e. *misconceptions–implicit*).

The study follows the two-part procedure suggested by Flick (1995, p. 164), according to which the data are first evaluated at the individual case level and then compared and contrasted with other cases. These exclusively

qualitative steps are then complemented by quantifications for further insights (Smith, 2000, p. 314). With the use of frequency analyses of both the noticings of the participants (i.e. unit-of-analysis based) and the participants (i.e. case based) we aim to structure the data to help us answer the research questions (for similar approaches, also see Harder, 2018, p. 111; Maxwell, 2010, p. 478; Sherin & van Es, 2009).

The intervention implemented between t_0 and t_1 consists of an elaborate training in geography education which spans a four-month period during a regular German university semester. During the intervention, in a first phase, students (i.e. future teachers) work on topics of differentiation, task culture, the relevance of and approach to deep structures of teaching and learning, and hence the already-mentioned constructs of cognitive activation, misconceptions, scaffolding, and fading as well as PV. After this rather theoretical phase (four sessions at 90 minutes), students begin to analyse lessons by watching videos of geography lessons and applying the acquired theoretical knowledge (one session at 90 minutes). They then plan their own lesson (two sessions at 90 minutes) which is held in an authentic school environment. This lesson is filmed and subsequently analysed in the last phase of the intervention – both individually and collaboratively (one 8-hour session).

Results

Results focus on the aforementioned research questions and include quotes from the participants' analyses and aggregated findings based on frequency analyses as described earlier.

First of all, regarding RQ 1.1, an overall increased number of classroom situations is noticed in t_1 compared to t_0 (see Table 9.2). Individually, participants show tendencies in different directions, i.e. increasing, decreasing, stagnating numbers (see Figure 9.1). In detail, this means the following: All 45 participants together provide a combined 504 noticings for t_0 and another 599 noticings for t_1 . This leads to an arithmetic average of 11.2 (t_0) and 13.3 (t_1) noticings per participant with a standard deviation of 5.5 for t_0 and 7.3 for t_1 and a median of 10 for t_0 and 11 for t_1 (see Table 9.2).

As becomes apparent in Figure 9.1, in total, 23 of the 45 participants notice more situations at t_1 compared to t_0 , 15 participants notice fewer situations, and 7 participants notice equally as many. Sixty-seven per cent of the participants with noticings below the median (i.e. 14 of 21) notice more events at t_1 compared to t_0 , whereas only 37% of participants (i.e. 7 of 19) with noticings above the median do (see Table 9.2 and Figure 9.1). Also visible in Figure 9.1 are participants who are able to notice considerably more situations than others (both at t_0 and t_1). Of these proficient noticers, however, few can maintain their level over the course of both measurements.

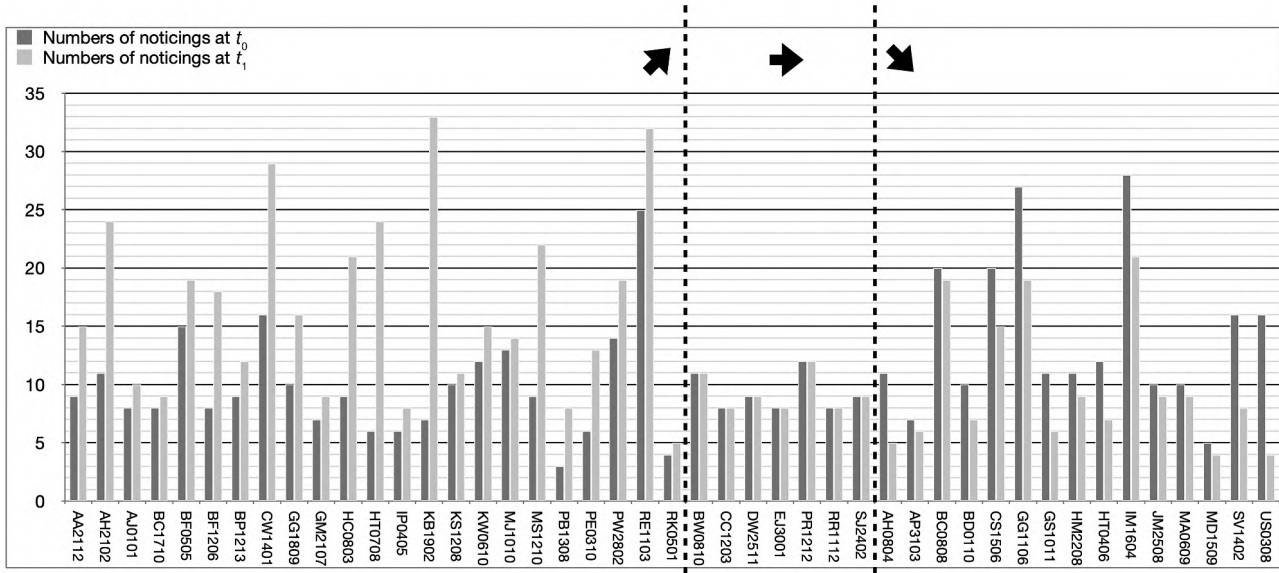


FIGURE 9.1 Summary of quantitative results of individual developments for research question 1.1 (left side: increasing from t_0 to t_1 ; centre: stagnating; right side: decreasing).

TABLE 9.2 Summary of quantitative results for research question 1.1

<i>Feature</i>	t_0	t_1
Noticings (total)	504	599
Arithmetic average of noticings per participant	11.2	13.3
Standard deviation	5.5	7.3
Median	10	11

When looking at RQ 1.2 and hence the distribution of noticings of surface structures that stay on the surface level, and noticings of surface structures that lead to inferences about deep structures, we see an increase of the latter (see Table 9.3 and Figure 9.2). Of the overall 504 units of analysis at t_0 , 51% (i.e. 256 units) provide noticings on the surface level and make no inferences about deep structures. This number decreases to 29% of the units (i.e. 176 of 599) at t_1 . Deep structure analyses behave reversely: At t_0 , 43% of units (i.e. 215 of 504) are indicative of the deep structure level, whereas 60% (i.e. 357 of 599) are at t_1 .

Examples of noticings of surface structures are “Very small groups, one teacher per group?” (t_0 MA0609: 2) or “Insecure posture, resting her hands on her hips, feet crossed – again from minute [2:02]” (t_0 KW0610: 4). Noticings which refer to deep structures relate to cognitive aspects: The example “Teacher makes a sketch to illustrate the facts to the pupils => intervenes exactly when pupils get stuck” (t_1 CW0610: 14) refers to the comprehension process. “Summary of the teacher to guide pupils and focus them more on questions and pictures, as pupils mention lots of things they also know” (t_0 CW1401:11) illustrates constructive support. An example of an unclear code, whose distribution is also illustrated in Figure 9.2, is “[Teacher] summarises: Hotel owners, farmers and asks about other groups of people, summary could also have been done by a pupil“ (t_1 AJ0101: 3). Here, the first part seems to be a description of the event on a surface level; the second part might indicate some form of potential for cognitive activation (i.e. on the level of deep structures) but does not provide a definitive argument.

At t_0 every participant is already able to make inferences about deep structures in at least one of his or her noticings. At t_1 , 34 of the 45 participants show more inferences about deep structures in their analyses compared to t_0 (see Table 9.3). There are also 7 of the 45 participants who make inferences about deep structures in all of their noticings at t_1 . Nothing of the sort can be seen at t_0 .

In the noticings which make inferences about deep structures, participants regularly apply geographical CK and PCK, as becomes apparent with the following examples:

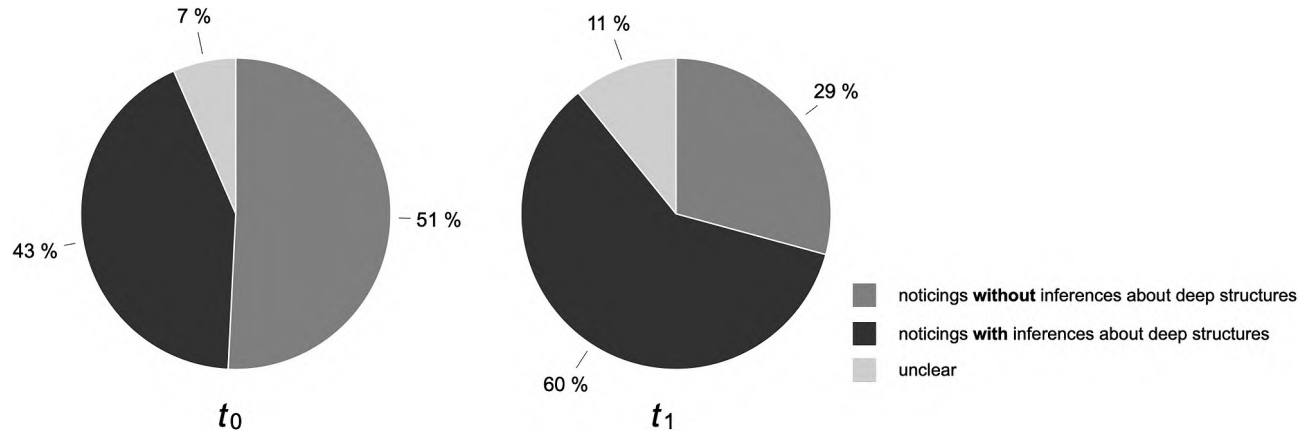


FIGURE 9.2 Comparison of t_0 and t_1 with regard to surface structures with and without inferences about deep structures.

TABLE 9.3 Summary of quantitative results for research question 1.2

<i>Feature</i>	t_0	t_1	<i>No. (n = 45)</i>
Numbers of noticings (total)	504	599	
Noticings of surface structures without inferences about deep structures (total)	256	176	
Noticings of surface structures with inferences about deep structures (total)	215	357	
Participants who notice surface structures without inferences about deep structures at t_0			45
Participants who notice surface structures with inferences about deep structures at t_0			45
Participants who make more inferences about deep structures at t_1 compared to t_0			34
Participants who make inferences about deep structures with every noticing at t_0			0
Participants who make inferences about deep structures with every noticing at t_1			7

- “Again summary by the teacher, at the same time assistance --> through this the students have new ideas and build on them more and more (from the employees in the hotel to supplier companies)” (t_0 GG1106: 7) → The analysis focuses on the pupils’ construction of new knowledge as well as on cognitive activation.
- “Explanation becomes more imprecise. Stone above groundwater; like a lake; lake does not grow back > Wording too imprecise and misleading!!!” (t_1 CW1401: 70) → The analysis focuses on the correctness of the teacher’s explanations and its impact on the pupils’ learning. The participant sees the danger of fostering misconceptions.

Results to RQ 1.3 include both the look at noticings of constructs relevant to the deep structure level that played a central role in the intervention (i.e. cognitive activation, misconceptions, scaffolding, and fading) as well as how participants used their knowledge base while reasoning about them. Since these constructs are part of the intervention, an increase of associated noticings can be expected from t_0 to t_1 – and the results are in line with this expectation. Also, participants reason more in t_1 than t_0 in an explicit manner about these constructs (see Tables 9.4 and 9.5).

In detail, 28% of the units of analysis at t_0 (143 of 504) were coded for at least one of the constructs, and 42% at t_1 (251 of 599); implicit reasonings more or less stagnate in relation to the total number of noticings, while explicit ones increase (see Table 9.4). As can be seen in Table 9.5, 43 of 45 participants already notice surface structures at t_0 which allows them to draw inferences to at least one of the four constructs. At t_1 , every participant does. In addition, more participants reason in an explicit manner after the

TABLE 9.4 Summary of quantitative results for research question 1.3

<i>Feature</i>	t_0	t_1
Number of noticings (total)	504	599
Number of noticings in which participants reason about at least one of the intervention constructs	143	251
Relevant to deep structures (i.e. cognitive activation, misconceptions, scaffolding, fading)		
Explicit reasoning		
with regard to cognitive activation	8	27
with regard to misconceptions	9	45
with regard to scaffolding	1	21
with regard to fading	0	6
Implicit reasoning		
with regard to cognitive activation	91	116
with regard to misconceptions	2	10
with regard to scaffolding	67	75
with regard to fading	27	29

TABLE 9.5 Summary of quantitative results of individual developments for research question 1.3

<i>Feature</i>	t_0	t_1
Number of participants (total)	45	45
Participants who notice surface structures with inferences about at least one of the intervention constructs relevant to deep structures	43	45
Participants who reason explicitly about		
cognitive activation	2	6
misconceptions	5	24
scaffolding	1	12
fading	0	5

intervention (also see Table 9.5). The length of the statements is generally very short for both implicit and explicit reasonings at both t_0 (18.3 words per reasoning) and t_1 (21.7 words per reasoning).

Discussion

Summary and interpretations of the results

The aim of this study was to explore the ability of future geography teachers to notice and reason about classroom situations before and after an intervention and, more specifically, to what extent they do so with regard to surface and deep structures. The results not only give insight into the PV abilities of

future geography teachers but also indicate the possibility for some positive developments of their abilities. RQ 1.1 can be answered twofold: On a momentary perspective, some few future geography teachers are able to notice more than twice as many classroom situations as most of the other participants can (both at t_0 and t_1). On a longitudinal perspective, there are developments in every direction; while most show an increase in noticings, a third of the participants are able to notice fewer situations at t_1 . According to RQ 1.2, participants are already able to draw inferences from surface structures about deep structures before the intervention. However, their ability to do so increases for three-quarters of the participants from t_0 to t_1 . Regarding RQ 1.3, participants are able to notice and reason about deep structure-level constructs that are highlighted in the intervention – even before the intervention. The manner in which they reason about them becomes more professional after the intervention, though. Results also indicate that the number of noticings is not necessarily connected to a high or poor quality of reasoning.

Additionally, while Sunder, Todorova, and Möller (2015, pp. 9–10) conclude that PV can be trained for primary school teachers and their specific topics, and Sherin and van Es (2009, pp. 32–33) see similar results for mathematics education, the results of this study suggest that PV can also be improved for a geography-centred educational setting. This becomes apparent since the intervention is able to stimulate more inferences on domain-specific deep structures for future geography teachers. Again, this seems to be consistent with Sherin and van Es (2009, pp. 29–30), who saw an increasing focus on thinking processes by teachers after their interventions. Positive influences of interventions are already seen by van Es and Sherin (2002, p. 591); however, geography-specific forms of development have not been looked at before. That is why more inherently domain-specific deep structures were focused. The fact that participants are able to analyse deep structures even without being urged or prompted to do so may indicate that future geography teachers see the importance of analysing on a deep structure level, i.e. focusing on teaching and learning, already at a very early stage of their training; yet, still, they are not capable of explicitly reaching this level of PV on a constant level for every noticing they make. After this intervention, however, they gain expertise in that regard.

Furthermore, only about half of the participants are able to notice more situations at t_1 . One crucial factor might lie in the participants' intent to notice beyond the surface which they may have gained from the intervention. Since we can assume that noticings of deep structures are more challenging (Trautwein et al., 2018, p. 8), it is not a stretch to also assume that they require certain abilities not every participant can display consistently, as mentioned earlier. This is supported by the fact that participants who rank highest at the number of noticings are also predominantly making inferences about deep structures at t_1 . So, they do not trade quality for quantity. This might also indicate that developments of PV are very much individual ones. Some are able to increase their

overall skills and to notice more and do so more often with inferences about deep structures, while others are limited by the more challenging inference and must choose between quality and quantity. For each group of teacher trainees, specific training concepts might be necessary.

Although the main focus of this study is on noticing, the results on knowledge-based reasoning seem encouraging as well. In accordance with Sherin and van Es (2009, p. 30) we also see an increasing amount of explicitness in the results. The participants' terminology became more elaborate and specific and included explicit technical vocabulary more often. Hence, the manner in which participants reasoned about the events they noticed became more knowledge based. This seems particularly important because of the power of a shared professional language (Egnal, 2015, pp. 19–28). The training effect becomes especially visible since it is the four theoretical constructs from the intervention that show a more elaborate form of reasoning. The fact that participants use a fairly concise way of reasoning is also found by Scholten, Höttecke, and Sprenger (2019, p. 174) for more experienced geography teachers. Here, this can be confirmed for less experienced teachers at university level as well, independent from the intervention. This suggests that the level of experience is not an essential factor for conciseness or brevity in the context of PV. Accordingly, teacher trainers at university should not equate brevity with a lack of quality.

Limitations

Nevertheless, this study also presents some limitations: First, it does not juxtapose the correctness and plausibility of the participants' analyses to an external standard yet. Such a standard could be achieved via an expert rating in which experts are asked to analyse the same video vignette in a consensus-orientated manner (Bortz & Döring, 2003, pp. 335–336; see also Gold et al., Chapter 7 of Volume 1). Comparable expert ratings are suggested for similar studies by Gold, Hellermann, and Holodynski (2016, pp. 116–117). In this study, the standard value is instead an internal one and established by comparing pre- and post-measurements of the 45 future geography teachers only.

Furthermore, the coding system of this study is exclusively deductive and limits the ability of the researcher to interpret the data more openly. Rather, it is focused on specific deep structure-level constructs that play a major role during the intervention. A more inductive approach may lead to a more open integration of the participants' analyses and include common noticings that the researchers have not been aware of beforehand. Deductive categories could also be added (e.g. via codes that categorise the units of analysis into *descriptive*, *explanatory*, or *evaluative* reasoning). They may provide further insight in the way participants reason about their noticings according to Seidel and Stürmer (2014). In addition to the expert rating, this could be another external standard.

When using the same video vignette as we did for t_0 and t_1 , the significance of the results may also be limited because of memory effects (Dreier et al., 2018, p. 157; Schwarz et al., 2020, p. 326). However, the two measurements are separated by a fourth-month intervention, which also includes several highly complex tasks that probably interfere with the participants' recollection of the original measurement (Schwarz et al., 2020, pp. 327–329).

For a qualitative study, a sample of 45 is quite substantial; however, quantifications from this sample do not allow for generalisations but may only show trends. Nonetheless, this explorative study may hint at some connections follow-up studies should apply additional focus on and explore further.

Conclusion

The study presented here was able to show that PV with regard to the highly learning-relevant deep structures can develop in a fairly short period of time at university level. With noticing and knowledge-based reasoning, two very important competencies of teachers were brought into view, which are also of high practical relevance for everyday classroom interactions. PV is a field of research in geography education that has hardly been explored so far. Hence, this research opens up excellent possibilities for future research, e.g. developing group-specific trainings for teacher trainees with a higher or lower focus on deep structures, investigating most commonly noticed constructs, or filtering out deep structures which are most difficult to notice and reason about via observations on the surface. At the same time, efforts to integrate the findings into the training of both future and experienced teachers in the sense of a theory-practice transfer between university and school seem immensely valuable.

References

- Ansorge, U., & Leder, H. (2017). *Wahrnehmung und Aufmerksamkeit* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-658-12912-5>
- Arnold, J., Kremer, K., & Mayer, J. (2016). Scaffolding beim Forschenden Lernen. *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 21–37. <https://doi.org/10.1007/s40573-016-0053-0>
- Bohl, T. (2017). Umgang mit Heterogenität im Unterricht. Forschungsbefunde und didaktische Implikationen. In T. Bohl, J. Budde, & M. Rieger-Ladich (Eds.), *Umgang mit Heterogenität in Schule und Unterricht. Grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen* (pp. 257–273). Klinkhardt.
- Bortz, J., & Döring, N. (2003). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3rd ed.). Springer.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5th ed.). Springer. <https://doi.org/10.1007/978-3-642-41089-5>

- Dreier, L., Leuthold-Wergin, A., & Lüdemann, J. (2018). Qualitativer Längsschnitt in der Bildungsforschung. In M. S. Maier (Ed.), *Qualitative Bildungsforschung* (pp. 153–168). Springer VS. https://doi.org/10.1007/978-3-658-18597-8_9
- Egnal, B. (2015). *Leading through language. Choosing words that influence and inspire*. Wiley. <https://doi.org/10.1002/9781119176602.ch2>
- Flick, U. (1995). Stationen des qualitativen Forschungsprozesses. In U. Flick, E. von Kardoff, H. Keupp, L. von Rosenstiel, & S. Wolff (Eds.), *Handbuch Qualitative Sozialforschung. Grundlagen, Konzepte, Methoden und Anwendungen* (2nd ed., pp. 147–173). Beltz.
- Flick, U. (2019). *Qualitative Sozialforschung. Eine Einführung* (9th ed.). Rowohlt.
- Gold, B., Hellermann, C., & Holodynski, M. (2016). Professionelle Wahrnehmung von Klassenführung. Vergleich von zwei videobasierten Erfassungsmethoden. In K. Schwippert & D. Prinz (Eds.), *Der Forschung – Der Lehre – Der Bildung. Aktuelle Entwicklungen der empirischen Bildungsforschung* (pp. 103–118). Waxmann.
- Harder, M. (2018). “Entweder sind die durch Zufall entstanden oder Gott hatte seine Finger im Spiel.” *Eine Interviewstudie zu Kindervorstellungen über das Naturphänomen Fluss im Kontext des Conceptual Change*. Klinkhardt.
- Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237. <https://doi.org/10.25656/01:4348>
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Beltz.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Waxmann.
- Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. Schöningh. <https://doi.org/10.36198/9783838538952>
- Kunter, M., & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV. Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (pp. 85–113). Waxmann.
- Lajoie, S. P. (2005). Extending the scaffolding metaphor. *Instructional Science*, 33(5/6), 541–557. <https://doi.org/10.1007/s11251-005-1279-2>
- Lipowsky, F., & Lotz, M. (2015). Ist Individualisierung der Königsweg zum erfolgreichen Lernen? Eine Auseinandersetzung mit Theorien, Konzepten und empirischen Befunden. In G. Mehlhorn, K. Schöppe, & F. Schulz (Eds.), *Begabungen entwickeln Kreativität fördern* (pp. 155–219). Kopaed.
- Maxwell, J. A. (2010). Using numbers in qualitative research. *Qualitative Inquiry*, 16(6), 475–482. <https://doi.org/10.1177/1077800410364740>
- Mehren, M., & Mehren, R. (2020). Über die Tiefenstrukturen des (Geographie-) Unterrichts. *Praxis Geographie*, 50(4), 4–9.
- Perels, F., & Otto, B. (2010). Evaluation von Interventionen. Erfassung der Veränderung versus Erfassung des Prozesses? In T. Hascher & B. Schmitz (Eds.), *Pädagogische Interventionsforschung. Theoretische Grundlagen und empirisches Handlungswissen* (pp. 252–259). Juventa.
- Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions. A review of intervention studies. *Review of Educational Research*, 66(2), 181–221. <https://doi.org/10.3102/00346543066002181>

- Scholten, N., Höttecke, D., & Sprenger, S. (2019). How do geography teachers notice critical incidents during instruction? *International Research in Geographical and Environmental Education*, 29(2), 163–177. <https://doi.org/10.1080/10382046.2019.1670915>
- Scholten, N., & Sprenger, S. (2020). Teachers' noticing skills during geography instruction: An expert-novice comparison. *Journal of Geography*, 119(6), 206–214. <https://doi.org/10.1080/00221341.2020.1809695>
- Schwarz, H., Revilla, M., & Weber, W. (2020). Memory effects in repeated survey questions reviving the empirical investigation of the independent measurements. *Survey Research Methods*, 14(3), 325–344. <https://doi.org/10.18148/srm/2020.v14i3.7579>
- Seidel, T., Blomberg, G., & Stürmer, K. (2010). “Observer” – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. Projekt OBSERVE. *Zeitschrift für Pädagogik* (suppl. 56), 296–306. <https://doi.org/10.25656/01:3438>
- Seidel, T., & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *American Educational Research Journal*, 51(4), 739–771. <https://doi.org/10.3102/0002831214531321>
- Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60(1), 20–37. <https://doi.org/10.1177/00224871083281>
- Smith, C. P. (2000). Content analysis and narrative analysis. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 313–335). Cambridge University Press.
- Sunder, C., Todorova, M., & Möller, K. (2015). Kann die professionelle Unterrichtswahrnehmung von Sachunterrichtsstudierenden trainiert werden? Konzeption und Erprobung einer Intervention mit Videos aus dem naturwissenschaftlichen Grundschulunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 22(1), 1–12. <https://doi.org/10.1007/s40573-015-0037-5>
- Trautwein, U., Sliwka, A., & Dehmel, A. (2018). *Grundlagen für einen wirksamen Unterricht*. Landesinstitut für Schulentwicklung.
- van Es, E. A., & Sherin, M. G. (2002). Learning to notice. Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10(4), 571–596.