Mentoring & Tutoring: Partnership in Learning

ISSN: 1361-1267 (Print) 1469-9745 (Online) Journal homepage: https://www.tandfonline.com/loi/cmet20

Teachers' diagnosis of students' research skills during the mentoring of the undergraduate thesis

Bas T. Agricola, Frans J. Prins, Marieke F. van der Schaaf & Jan van Tartwijk

To cite this article: Bas T. Agricola, Frans J. Prins, Marieke F. van der Schaaf & Jan van Tartwijk (2018) Teachers' diagnosis of students' research skills during the mentoring of the undergraduate thesis, Mentoring & Tutoring: Partnership in Learning, 26:5, 542-562, DOI: 10.1080/13611267.2018.1561015

To link to this article: https://doi.org/10.1080/13611267.2018.1561015

MENTORING & TUTORING: PARTNERSHIP IN LEARNING
2018, VOL. 26, NO. 5, 542–562
https://doi.org/10.1080/13611267.2018.1561015

Routledge
Taylor & Francis Group

ARTICLE

Teachers' diagnosis of students' research skills during the mentoring of the undergraduate thesis

Bas T. Agricola[a,b], Frans J. Prins[a], Marieke F. van der Schaaf[a,c] and Jan van Tartwijk[a]

[a]Faculty of Social and Behavioural Sciences, Utrecht University, Utrecht, The Netherlands; [b]Faculty of Health and Behavioural Studies, HAN University of Applied Sciences, Nijmegen, The Netherlands; [c]Center for Research and Development of Education, Utrecht Medical Center, Utrecht, The Netherlands

ABSTRACT

In higher education, students often write an undergraduate thesis and receive one-to-one or small group support. During mentoring, 10 teachers ideally diagnose students' research skills, to be able to adapt their support to students' needs. In this study, we aimed to answer the question of how mentors apply the diagnostic phases of a diagnostic question, a diagnosis, a diagnostic check and an intervention, during mentoring meetings about 15 students' research skills. Four mentors participated in this multiple case study. Qualitative data were gathered and sixteen videotaped mentoring meetings were coded on the four diagnostic phases. The results were compared within and between mentors, showing that mentors asked several diagnostic questions, seldom articulated and shared their diagnoses explicitly with students, and mainly used interventions. We concluded that more support is needed for mentors who do not automatically use their diagnostic questions to formulate explicit diagnoses about students' research skills.

ARTICLE HISTORY
Received 11 August 2017
Accepted 29 August 2018

KEYWORDS
Diagnosing; research skills; undergraduate thesis; mentoring; mentors; thesis; observation

Introduction

In higher education (HE), it is common for students to carry out a research project and to write a thesis about it in the final part of their undergraduate programme. Undergraduate research projects often consist of several phases in which students have to carry out a literature review, prepare a research plan, collect and analyze their data, and, finally, present the findings in a thesis. Often, students are awarded between 15 and 30 European Credit Transfer and Accumulation Systems (ECTS) credits for successfully completed research projects. In the Netherlands, one year corresponds to 60 credit points and one credit point equals to 28 working hours. For most students, this is the first research project they have engaged in during their educational career and, as

CONTACT Bas T. Agricola b.t.agricola@uu.nl Faculty of Social and Behavioural Sciences, Utrecht University, Heidelberglaan 1, Utrecht, The Netherlands

a result, they have to develop their research skills during the project (de Kleijn, R. A. M., Mainhard, Meijer, Brekelmans, & Pilot, 2012)

During the research process, a student normally receives several forms of mentoring support. Written comments on theses in progress are widely used to improve students' academic writing (Basturkmen, East, & Bitchener, 2014). Frequent one-on-one mentoring meetings with verbal feedback lead to theses being completed and to students being satisfied with the supervision (Heath, 2002; Shanahan, Ackley-Holbrook, Hall, Stewart, & Walkington, 2015). During these meetings, mentors use dialogue and communication with their students to obtain better understanding of students' experiences and perceptions (Greenbank & Penketh, 2009; Jaldemark & Lindberg, 2013). These dialogues are normally a two-way process, with teacher-student interaction and active student engagement (Graesser, Person, & Magliano, 1995; Nicol, 2010), as students can check their interpretation of the feedback or ask for further explanation (Nicol & Macfarlane-Dick, 2006). For mentors, this interaction is important because they can check whether their students understand their feedback and explanations.

The mentoring process of undergraduate research is complex as it includes communication, collaboration, and conflict in addition to understanding the mentor-student interactions (Palmer, Hunt, Neal, & Wuetherick, 2015). During mentoring meetings many mentors struggle with the balance between intervening and providing support, on the one hand, and allowing students to find their own ways, on the other (Todd, Bannister, & Clegg, 2004; Vehviläinen & Löfström, 2014). For example, students struggle with the production of a specific research question, which is one of the most challenging aspects during the research process for undergraduate students and mentors (Todd et al., 2004). Although mentors feel that responsibility for the thesis belongs to the students, they find it hard to determine a balance between direct instruction and the student's own decision making (Todd, Smith, & Bannister, 2006). A dependent relationship is created when students rely heavily on their mentor to provide feedback and when the mentor is willing to give it (Sambrook, Stewart, & Roberts, 2008).

Mentors need to be sensitive to all the differences between the students and, as such, the level and amount of support needs to be adapted to students' needs and the support will differ from student to student (Engebretson et al., 2008; Shanahan et al., 2015; Todd et al., 2006). Mentors have emphasized the importance of tailoring their guidance to the individual learning needs of the student; depending on the individual needs of the student they provide varying levels of guidance (Manathunga, 2005). However, mentors are unsure about how extensive and detailed this support should be and, as a result, they have troubles tailoring their support to students' abilities and needs (Vehviläinen & Löfström, 2014).

When teachers diagnose their students' understanding accurately, they can develop and apply more effective and efficient supervising or mentoring strategies, with conscious consideration of students' needs (Hedin & Gaffney, 2013; Südkamp, Kaiser, & Möller, 2012). Teachers' diagnosing is crucial for the quality of research mentoring (de Kleijn, Meijer, Brekelmans, & Pilot, 2015). Therefore, we examined how mentors diagnose students' research skills by describing different characteristics of mentors' diagnostic behaviour.

Teachers' diagnostics

Teachers' diagnosing can be operationalized as teachers' ability to judge students' achievement or task difficulties (Klug, Bruder, Kelava, Spiel, & Schmitz, 2013). In this study, we defined teachers' diagnostics as their ability to judge their students' research skills. Teachers' diagnosing has been a research topic for some time in contexts other than HE (Hoth et al., 2016; Klug et al., 2013). Researchers investigating teachers' diagnostic skills in classroom settings in primary education (PE) and secondary education (SE) have included findings that indicated diagnosing is complex (Ruiz-Primo & Furtak, 2007; Van de Pol, Volman, & Beishuizen, 2011). Diagnosing students' level of understanding in classrooms is rather difficult for teachers and hardly occurs during teacher-student interactions (Graesser et al., 1995; Putnam, 1987; Van de Pol, Volman, & Beishuizen, 2010). When teachers do diagnose, their diagnoses are often far from perfect and there is plenty of room for improvement (Südkamp et al., 2012; Van de Pol & Elbers, 2013). Instead of diagnosing, teachers either focused on objectives of their own (Nathan & Kim, 2009) and on beliefs about what they thought was difficult for the students (Van de Pol et al., 2011), or they intervened immediately (Ruiz-Primo & Furtak, 2007).

De Kleijn, Bronkhorst, Meijer, Pilot, and Brekelmans (2014) examined 12 mentor-student dyads and found that thesis mentors provided support that was adapted to their own goals and to students' goals. These mentors collected information about the students and sometimes the mentors explicitly formulated a student's characteristic that they observed. In the study of de Kleijn et al. (2015) interviews and group discussion meetings with five expert master's thesis mentors were conducted. They concluded that these mentors carefully diagnosed students' characteristics, such as competence level and determination. If we want to understand more about the diagnostic process of HE research mentors, we need to know what they are doing and observe their naturalistic behaviour with their students.

We believe that diagnosing is a complex skill and that teachers should be diagnosing during interactions with their students. This paper builds on results from primary/secondary education and applies this research literature to the HE context. Given the fact that the quality of the diagnosing process is crucial for adaptive mentoring and student learning (de Kleijn et al., 2014, 2015; Hedin & Gaffney, 2013), and the knowledge that, in

general, teachers' diagnostic processes need to be improved (Südkamp et al., 2012), it is worthwhile to unravel the diagnostic process involved in thesis mentoring.

Identifying diagnostic phases in research mentoring

In order to examine the diagnostic process of HE teachers, we used a framework that summarizes three popular models from primary/secondary education (Klug et al., 2013; Ruiz-Primo & Furtak, 2007; Van de Pol et al., 2011). Klug et al. (2013) described the diagnostic process as teachers who diagnose their student's level of understanding. Teachers make a prediction about a student's performance and possible underlying learning difficulties. Initially, the teachers interpret the information that is gathered, come to a concluding diagnosis, and finally give feedback (Klug et al., 2013). Ruiz-Primo and Furtak (2007) described the diagnosing process as an assessment conversation. Initially, the teachers elicit a question, then recognize the student's response, and finally use the collected information to support student learning. Van de Pol et al. (2011) used a model of contingent teaching in which the diagnosing process is described. Initially, the teachers apply diagnostic strategies, then check the diagnosis, and finally intervene. Although the diagnostic process in these three models sometimes lacks a specific diagnosis (Van de Pol et al., 2011)and sometimes a diagnostic check (Klug et al., 2013; Ruiz-Primo & Furtak, 2007), the three models are described quite similarly, with three cyclical phases within the teacher-student interaction. In our study, we combined the phases from these studies (Klug et al., 2013; Ruiz-Primo & Furtak, 2007; Van de Pol et al., 2011) and distinguished four diagnostic phases consisting of a diagnostic question, a diagnosis, a diagnostic check and an intervention. This model can be used as a lens to observe and determine mentors' diagnostic behaviour as seen in Figure 1.

Diagnostic question

In the first phase of *diagnostic questions*, mentors ask the students questions to gather information on their research skills, e.g. 'How would you make this research question more specific?' The students respond and show (or do not show) the skills. This gives the mentors a basis on which to decide whether they know enough about the student or more diagnostic questions can be asked before a diagnosis is reached. These questions provoke an interactive mentor-student mentoring meeting (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; Chin, 2006) and elicit further information about students' research skills. Diagnostic questions can lead to a prediction about a student's development and possible underlying difficulties (Klug et al., 2013) and to a more accurate diagnosis.

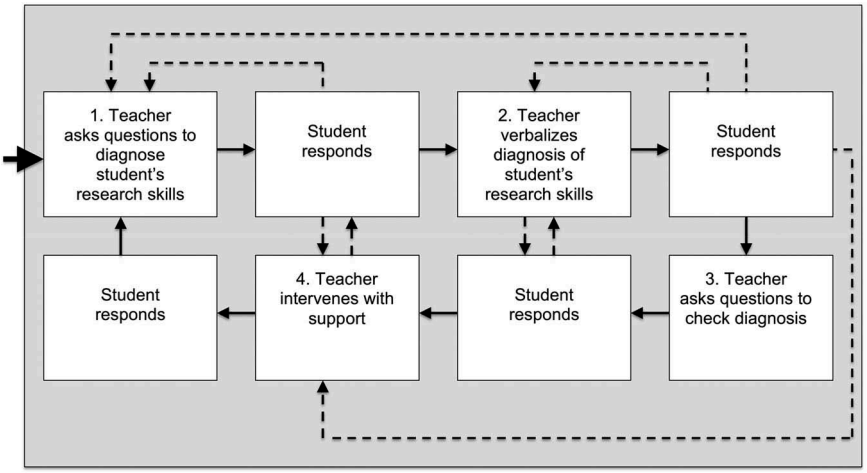

Figure 1. Four phases of diagnostic process during research mentoring (adapted from Klug et al., 2013; Ruiz-Primo & Furtak, 2007; Van de Pol et al., 2011).

Diagnosis

In the second phase of *diagnosis*, mentors can interpret the gathered information and come to a concluding diagnosis in which the mentors articulate a diagnosis about student's research skills, e.g. 'You don't know yet how to compose specific research questions, because you're doing this for the first time.' By explicitly articulating a diagnosis, the mentors are judging students' achievement, giving the student the opportunity to respond with a confirmation or rejection of the diagnosis. A diagnosis should be made explicit (Klug et al., 2013), as then it allows a shared understanding between the mentor and the student (Van de Pol et al., 2011) and they can work on common goals (Halse & Malfroy, 2010).

Diagnostic check

In the third phase of *diagnostic check*, mentors check the diagnosis of a student's research skills, e.g. 'If I understand you correctly, you don't know because you haven't read the book chapter about this topic yet?' Mentors verify whether they captured the research skills of the student correctly (Van de Pol et al., 2011). Although diagnostic checks can also provide mentors with additional information about students' research skills, the goal of a diagnostic check is verification of the mentor's own diagnosis.

Intervention

In the fourth phase of *interventions*, mentors support the student, preferably in a way that is adapted to the diagnosis (Klug et al., 2013; Van de Pol et al., 2011), e.g. 'Because you have one group of children and one experiment with

no comparison, the study you are conducting will probably be a prospective cohort study.' Mentors intervene; they give feedback and explanations, and want to help the student with understanding the subject matter or with the task of approaching the thesis. In this phase, the mentors are engaged in moment-by-moment decision-making in order to support the student in understanding the relevant issues (Hedin & Gaffney, 2013). Framing the intervention phase as an outcome of the diagnostic process is important as mentors can only provide adaptive support, when they have asked enough diagnostic questions and gathered enough information to make a diagnosis.

Present study

In our study, we focused on mentors who have mentor-student dialogues, in a setting in which teacher mentors can focus on diagnosing students' research skills without the classroom context. One might expect these teachers to be better able to diagnose their individual students' learning than PE and SE teachers. Teachers in higher education often supervise students during one-to-one and/or small group mentoring meetings when students are writing a thesis about a research project (McCallin & Nayar, 2012). The mentors and students in our study are used to having these mentoring meetings and, thus, give opportunities to observe the diagnostic behavior in a naturalistic setting.

We focused our study on mentoring meetings regarding students' writing of a research plan. The writing of a research plan is a crucial phase, as it involves making decisions about the direction of the project. When students write a research plan, they review the literature to develop a conceptual framework, determine the aim and focus of the study, compose the research questions, specify a research design, and choose their research instruments (Wisker, 2009). Mentors can support students in their development of these research skills. Given the importance of arriving at an accurate diagnosis for mentors, it is interesting to investigate what kind of diagnostic phases they apply during the mentoring of an undergraduate thesis. The research question for our study was as follows: How do mentors apply the diagnostic phases of a diagnostic question, diagnosis, diagnostic check and intervention during mentoring meetings about students' research skills?

Method

Design and participants

A multiple case study method was chosen, with a case being defined as one mentor teacher. Ten mentor teachers from a four-year bachelor of health undergraduate programme at a Dutch university agreed to participate. The supervised students were in their final year and worked on their research

project for 20 weeks (30 ECTS; 840 hours). Mentors had several (approximately 8–10) mentoring meetings with their students.

Case sampling

As we were interested in the diagnostic behaviour of mentors, three or four cases with quite similar characteristics were sufficient to predict similar results (Yin, 2013). We selected four mentors with a most similar method, where cases are similar on specified variables (Seawright & Gerring, 2008). The four cases were all female mentors; this was representative for the complete teaching staff. The characteristics that were used for case selection were age, highest degree earned, number of supervising years, and number of supervised theses. The four mentors were between 28–34 years of age (M = 30.50; SD = 2.65) and all had a master's degree. Two of the selected mentors were in their first year supervising and had a low number of supervised theses (Amy: 5 theses; Brooke: 8 theses). The other two selected were slightly more experienced mentors (Claire: 3 years; Debby: 4 years), and off course had supervised more theses (Claire: 30 theses; Debby: 35 theses). All the names used are pseudonyms.

Procedure

The four teacher mentors were supervising students who were writing their thesis in pairs. A mentor supervised a group of no more than five or six pairs. The mentor offered support to the students during mentoring meetings, consisting of a dialogue between one teacher and two students. The research plan was discussed in the third week of the research project. A few weeks before the start of the project, teacher mentors and students were informed that the study focused on the interaction between a mentor and their students. All the participants gave informed consent before the data collection started. For each mentor, four mentoring meetings were recorded on video, without the researcher being present. The students were used to cameras, as they often videotaped their own conversations for assessment and self-reflection. The mentors informed us that they were aware of the camera for the first few minutes but that they forgot that they were being recorded as the discussion with the students continued. The 16 videotaped mentoring meetings were transcribed verbatim; the videos were transcribed literally, with punctuation, pauses, and continuers (e.g. hm, yeah) but without intonation and voice volume.

Data analyses

The coding of the 16 transcribed mentoring meetings was conducted in four steps: (a) coding on topic, (b) segmentation in units of meaning, (c) coding on the diagnostic phases, and a (d) sequential analysis. To get insight on the

variety of diagnosing phases among mentors, we examined how mentors differed in how they conducted their mentoring. First, we identified which codes were used most. A within-case analysis was conducted between the four mentoring meetings of each mentor and a between-case analysis was conducted between all four mentors. The relative scores of the diagnostic phases and the significant results from the sequential analysis were used to determine similar and/or different patterns (Miles & Huberman, 1994). Several excerpts from the mentoring meetings were selected to show the application or absence of diagnostic phases and to illustrate sequences of diagnostic phases. In the next subsections, more details are provided about the four coding steps.

Coding on topic

As three different actors attended every mentoring meeting, the transcripts contained three different turns: teacher turns, student1 turns and student2 turns. Speech turn taking was used as the first segmentation criterion because it fits with the natural course of the conversation (Chi, 1997). The analysis focused on the teacher turns only. Since we were interested in mentor behaviour that concerned the mentoring of an undergraduate research project, all mentors' turns were first coded for being 'on topic' or 'off topic.' An on topic turn is a turn in which elements of the undergraduate thesis were discussed, such as 'the definition of the research problem,' 'the research questions,' and 'the theoretical framework.' The unit of analysis consisted of each mentor's turn. The first author coded all mentors' turns (n = 2989) to determine if they were on topic. The interrater reliability was determined by coding 10% of all mentors' turns (2/16 mentoring meetings; 304/2989 mentor turns) independently with the second (1/16 meeting; 215/2989 turns) and third author (1/16 meeting; 89/2989 turns). Interrater reliability between the first and the second author (Krippendorff's $\alpha = .79$) and between the first and the third author (Krippendorff's $\alpha = .76$) was satisfactory (Krippendorff, 2004).

Segmentation in units of meaning

In the second step, we segmented all on topic mentor turns in units of meaning, based on semantic features and, specifically, on ideas or topics of discussion (Chi, 1997). Every on topic mentor's turn was characterized as either (a) a single unit of meaning or (b) segmented into several units of meaning. For example, when a mentor was addressing two different topics within one turn (e.g. 'research questions' and 'data gathering'), this turn was segmented into two units of meaning. The first author segmented all mentors' on topic turns (n = 1881) into units of meaning. The reliability of the segmentation was computed as the proportion agreement because there was only one category involved with two values (agree = 1, disagree = 0) (Strijbos, Martens, Prins, & Jochems, 2006). The proportion agreement was determined by segmenting 10% of all mentors' on topic turns (2/16 mentoring meetings; 188/1881 on topic turns) independently with the second (1/16 meeting; 117/1881 turns)

and third author (1/16 meeting; 71/1881 turns). With an agreement percentage of 94% (with the second author) and 92% (with the third author), the reliability of the segmentation appeared to be good (Riffe, Lacy, & Fico, 2005).

Coding diagnostic phases

The third step of coding was conducted on all units of meaning. The first author coded all units of meaning (n = 1963) on one of the four diagnostic phases: (a) diagnostic question, (b) diagnosis, (c) diagnostic check, and (d) intervention. The interrater reliability was determined by coding 10% of all mentors' turns (2/16 mentoring meetings; 204/1963 mentor turns) independently with the second (1/16 meeting; 132/1963 turns) and third author (1/16 meeting; 72/1963 turns). The interrater reliability between the first and second author (Krippendorff's α = .73) and between the first and third author (Krippendorff's α = .75) was satisfactory (Krippendorff, 2004).

Sequential analysis

Finally, a sequential analysis was conducted on all coded units of meaning to analyse the sequential links between the four phases (diagnostic question (DQ), diagnosis (D), diagnostic check (DC), and intervention (I)) and to identify patterns within the sequences of these phases (Jeong, Clark, Sampson, & Menekse, 2011). A simple evaluation of the independence of these phases was done in order to identify whether a particular sequence of behaviour occurred more or less often than might be expected by chance alone (Wampold & Margolin, 1982). The program Multi Episode Protocol Analysis (MEPA) was used and three different scores of possible sequences were computed: a frequency score, an expected score, and a Z-score (Erkens, 2005). Each sequence consisted of exactly two diagnostic phases, starting with one phase (e.g. diagnostic question) and followed by another (e.g. intervention). Thus, the four different diagnostic phases led to 16 different sequences, which were tested. Chi square tests can be performed when the expected value is more than five (Gravetter & Wallnau, 2013). Therefore, a threshold with a minimum expected frequency of five was used for further analysis and reporting of the results. All sequences that involved a diagnostic check occurred very infrequently because only one diagnostic check was coded and, thus, they were not used for further analysis.

Credibility of the study

Several basic key elements to the study design were integrated to enhance the overall study quality and credibility. Clearly written research questions were posed that fitted the case study design (Baxter & Jack, 2008; Yin, 2013). A purposeful case sampling procedure was conducted based on the most similar method described by Seawright and Gerring (2008). The videotaped data were collected and transcribed systematically. The mentoring meetings were

videotaped and coded from the beginning to the end. The goal of selecting excerpts from the different mentoring meetings was to present thick and rich descriptions of the diagnostic phases, and to bring the mentor-student interactions alive (Creswell & Miller, 2000). Triangulation of researchers was applied to explore the diagnosing process of mentors from different perspectives (Baxter & Jack, 2008; Creswell & Miller, 2000). The first author coded all transcripts; the second and third author made coding checks (Guba, 1981; Miles & Huberman, 1994). Interrater reliability and agreement were determined for the first three coding steps (Krippendorff, 2004; Riffe et al., 2005).

Results

First, the descriptive results from coding the four diagnostic phases will be presented. Then, an excerpt is presented from each mentor and each diagnostic phase to provide better insight into the application of the four diagnostic phases, and finally we present the sequential analysis results.

Mentors' diagnostic phases

Regarding our research question about how mentors apply the diagnostic phases of a diagnostic question, diagnosis, diagnostic check, and intervention, the descriptive results show that mentors mainly used *interventions* and applied *diagnostic questions* less frequently (see Table 1). Amy, Claire, and Debby showed similar diagnostic behaviour and applied many interventions. Brooke showed different diagnostic behaviour than the other three mentors, she asked substantially more diagnostic questions and by that Brooke had better opportunities to gather enough information to diagnose students' research skills. All four mentors barely formulated a *diagnosis* in which they articulated the level of the students' research skills, and, consequently, a *diagnostic check* was scarcely observed. Even Brooke who asked a lot of diagnostic questions did not formulate her diagnoses explicitly with her students. As a consequence, these mentors did not share a lot of understanding about the diagnosed research skills with their students.

Sequences of diagnostic phases

The sequential analysis results show that mentors applied the sequences of a diagnostic question and a diagnostic question (DQ-DQ) and the sequence of an intervention and an intervention (I-I) significantly more frequently than might be expected by chance. These sequences were quite similar for all four mentors. It seems these mentors used a lot turns with diagnostic questions to diagnose students' research skills and also used a lot of turns with interventions to support students' skills (see Table 2).

Table 1. Percentage of diagnostic phases per mentor for each meeting and total frequency.

	Amy						Brooke						Claire						Debby					
	A1	A2	A3	A4	M	SD	B1	B2	B3	B4	M	SD	C1	C2	C3	C4	M	SD	D1	D2	D3	D4	M	SD
Question	21	26	37	27	27.77	6.65	57	41	42	43	45.84	7.19	28	19	20	29	23.93	5.08	17	28	21	25	22.98	4.58
Diagnosis	2	4	5	3	3.39	1.29	4	5	4	3	3.97	1.04	7	7	1	0	3.76	3.75	5	0	0	0	1.34	2.68
Check	0	0	0	0	0.00	0.00	0	1	0	0	0.22	0.44	0	0	0	0	0.00	0.00	1	0	0	0	0.17	0.34
Intervention	77	70	58	70	68.84	7.77	39	53	54	54	49.97	7.11	65	74	78	71	72.31	5.55	77	72	79	75	75.51	2.80
Total (n)	114	147	149	165	143.75	21.41	122	114	132	183	137.75	31.05	72	128	93	73	91.50	26.19	149	111	108	103	117.75	21.09

Note. Four mentoring meetings (1–4) were coded for each mentor, A = Amy, B = Brooke, C = Claire, D = Debby

Table 2. Summary of sequential analysis for the diagnostic phases of diagnostic question, diagnosis and intervention.

		Diagnostic question				Diagnosis				Intervention			
		A	B	C	D	A	B	C	D	A	B	C	D
Diagnostic question	Fr	81.00	129.00	26.00	37.00	9.00	5.00	5.00	1.00	70.00	69.00	33.00	68.00
	E	44.24	69.85	10.48	23.91	5.57	5.16	2.26#	1.58#	110.19	127.99	51.26	80.29
	Z	7.59*	10.80*	5.64*	3.47*	-	-	1.99	-	−8.02	−10.53	−6.48	−3.11
Diagnosis	Fr	9.00	2.00	2.00	.00	3.00	5.00	4.00	.00	8.00	8.00	8.00	6.00
	E	5.53	5.16	2.29#	1.58#	.70#	.38#	.49#	.10#	13.77	9.46	11.21	5.30
	Z	-	-	-	-	2.86	7.67	5.17	-	−2.82*	-	−2.19*	-
Intervention	Fr	69.00	72.00	37.00	69.00	8.00	5.00	5.00	6.00	318.00	295.00	277.00	281.00
	E	109.23	127.99	52.23	80.29	13.74	9.46	11.25	5.30	272.03	234.55	255.52	356.00
	Z	−8.22*	−10.00*	−5.13*	−2.85*	−2.82*	−2.40*	−4.24*	-	8.87*	10.72*	6.71*	2.97*

Note 1: Mentor A = Amy; B = Brooke; C = Claire; D = Debby.
Note 2: Fr = Frequency; E = Expected score; Z = Z-score; * $p < .05$; # $E < 5.00$.

Diagnostic question

The excerpt of Brooke shows the repetition of diagnostic questions (DQ-DQ) without a diagnosis. This excerpt was typical for Brooke as a mentor, as she applied many diagnostic questions throughout her four mentoring meetings. She applied the largest quantity of diagnostic questions of the four selected mentors. In Table 3 we present a description of this excerpt.

In this excerpt, Brooke asked several diagnostic questions (turns 1, 3 and 5), in order to diagnose the students' research skills concerning both the formulation of a research question and how to make this question more specific. By asking several diagnostic questions, Brooke seemed to imply that the formulation of the students' research questions was not specific enough. After three diagnostic questions, when the students still could not give the right answer, Brooke intervened by saying 'the word "nutrition" has to be more specific.' This excerpt shows a mentor who skipped the explicit articulation of a diagnosis after her diagnostic questions and provided an intervention immediately.

Diagnosis

The excerpt of Amy shows one of the few instances where the actual diagnosis was articulated. This excerpt illustrates the diagnostic behaviour of Amy, with a few diagnostic questions and many interventions. In Table 4 we present a description of this excerpt.

Amy tried to figure out what kind of research design the students were actually dealing with and what kind of data the students were collecting. Amy started with a diagnostic question about the design, offering students an opportunity to answer it. Instead of asking more diagnostic questions and waiting for an answer of the students, she answered the question herself and even explained why she thought it was a prospective cohort study. She continued with this intervention approach, by stating that this design offered descriptive data. Then, Amy asked another diagnostic question about statistical testing. However, this time, the students offered the correct answers and Amy ended this section with

Table 3. Transcript excerpt from Brooke (dialogue1 lines 292–310).

Turn	Speaker	Utterance	Code
1	Brooke	Uhum, how would you make this research question more specific? Have you thought about that?	Diagnostic question
2	Student2	No.	
3	Brooke	And what is the point of this…well to make this more specific?	Diagnostic question
4	Student1	Yes well, I do not know if this is specific enough or that you think it should be even more specific?	
5	Brooke	Hmhm. What do you think? Do you think that, if this is your question, you can develop a brochure, you can do some literature research and find out what is possible and what is not?	Diagnostic question
6	Student1	Uhm … yeah, I think it is specific enough.	
7	Student2	Yes.	
8	Student1	I do not really know how you can make it more specific.	
9	Brooke	No uhm, well, you want to make the word 'nutrition' more specific.	Intervention

Table 4. Transcript excerpt from Amy (dialogue3 lines 636–665).

Turn	Speaker	Utterance	Code
1	Amy	And will this group of children be compared to another group?	Diagnostic question
2	Student2	No	
3	Amy	Nope. That gives a lot of information. It will probably be a prospective cohort study	Intervention
4	Student2	Ooh…okay	
5	Amy	Because you have one group of children and one experiment will be applied, there is no comparison, thus it is not controlled for and it is not ehm back in time so it is forward in time.	
6	Student2	Hmhm	Intervention
7	Amy	And finally here you got your outcomes, and what is very important is that you get very descriptive data and that you canprobably already think of okay what statistical tests can we do.	Intervention
8	Amy	Well what kind of test can you do?	Diagnostic question
9	Student1	Yes I thought eh descriptive statistics, so one sample t tests And I have been describing the t test here, but I also had a question about it haha because i did not quite get it…	
10	Amy	No?	
11	Student1	This is what we have got so far.	
12	Amy	Well quite good though, you are already very far. So you guys have really thought this through, and now you see how far youcan get, even when you haven't started with the analysis	Diagnosis

a diagnosis about how well the students did. This diagnosis was not followed by new interventions and was more of a conclusive diagnosis.

Diagnostic check

The excerpt of Debby shows another of the few instances with an explicit diagnosis (turn 3), but without a diagnostic check, followed by the repetition of interventions (turns 4, 6, 8, 10 and 12). This excerpt is typical for

Table 5. Transcript excerpt from Debby (dialogue1 lines 224–244).

Turn	Speaker	Utterance	Code
1	Debby	Do you mean an action plan?	Diagnostic question
2	Student2	Yes	
3	Debby	Uhm …yeah yeah I know…but it is quite logical you don't know though, because you're doing this for the first time…	Diagnosis
4	Debby	And uhum, I can well imagine that you do not know exactly what part belongs in and what part of it comes later, but the more you already know how to approach your research right now the better…because you actually…	Intervention
5	Student2	Yes	
6	Debby	…uhm you just want to limit your search uh, and not be searching too long too broadly, because it takes too much time actually.	Intervention
7	Student2	Yes	
8	Debby	So if the two of you go and brainstorm about how we are going to tackle the report of health promotion	Intervention
9	Student1	Uhum (nods yes)	
10	Debby	Uhm it's okay if you then write down all of it in great detail, okay… we will recruit patients that way. We will approach them in that way or we are going to do interviews or a survey, that you have To actually decide yet, how are you going to conduct the survey or Interview. uhm uhm where do you start your literature search, what Kind of literature you could possibly need.	Intervention
11	Student1	Yes	
12	Debby	And where do you search for existing interventions	Intervention

Debby as a mentor, as she applied many interventions throughout her four mentoring meetings, with little diagnostic questioning. In Table 5 we present a description of this excerpt.

Debby first stated that the students do not know the answer (*diagnosis*), obviously, because they are doing research for the first time (*intervention*). After the diagnosis has been articulated, Debby provided support through several interventions, aimed at showing how the students could proceed. This excerpt shows that Debby provided support without a *diagnostic check*. A check would have given Debby the opportunity to verify with her students whether she was right about her diagnosis. In this excerpt, Debby's support seems to be taking over the thought processes of the students, as they are only affirming what the mentor is saying.

Intervention

The excerpt of Claire shows quite a similar pattern to those of Amy and Brooke. Claire asked one or two diagnostic questions (turns 1, 3 and 9) and applied interventions (turns 6, 11 and 13) with explanation and instruction.

First, Claire asked about the literature research the students had done (*diagnostic question*) and, after they gave their responses, proposed an explanation for their failure to find many studies (*intervention*). Subsequently, she asked a more specific diagnostic question about the relationship. Again, this question was immediately followed by two interventions that explained why the students should rephrase their research question. In Table 6 we present a description.

Table 6. Transcript excerpt from Claire (dialogue1 lines 162–207).

Turn	Speaker	Dialogue	Code
1	Claire	Yes. Did you just do a literature research?	Diagnostic question
2	Student1	Yes we did a bit	
3	Claire	What was the result?	Diagnostic question
4	Student2	We particularly looked at how much dental caries exists and if there Were studies done in that area	
5	Student1	We found one study of 1992 … we can use it, but we are not sure if It is up to date	
6	Claire	Heh no, but it could also mean that there is no other research done	Intervention
7	Student1	Yes, we found two …	
8	Student2	I found one of 2007	
9	Claire	Did you do some literature research into the relationship between Nutrition and dental caries?	Diagnostic question
10	Student1	We have particularly searched for that relationship, on PubMed, You can find a lot of hits	
11	Claire	Okay, because it is also important to determine what exactly the Question is that you want to address. Maybe you do not want to Research the correlation, but there may be another question …	Intervention
12	Student1	Yes	
13	Claire	Because you already know that there is a relationship, so you do Not need to re-examine that …	Intervention

Discussion

The aim of our study was to examine how mentors diagnose students' research skills, by describing different characteristics of mentors' diagnostic behaviour. In order to answer the research question, we analysed mentor utterances that could be coded in one of the four diagnostic phases of the model. The excerpts and sequential analysis of the four mentors showed that mentors followed several phases of the cyclical model with the four diagnostic phases.

Diagnostic question

It seems very natural to ask several diagnostic questions about one topic in sequence to gather enough information about the students, just as our mentors did. Amy, Claire and Debby did not ask as many diagnostic questions as Brooke did and, as a result, would find it more difficult to tailor their support to the student's needs (Nicol & Macfarlane-Dick, 2006; Vehviläinen & Löfström, 2014). Overall, our mentors asked an adequate amount of diagnostic questions, which could potentially present them with enough information to formulate a diagnosis about their students' research skills.

Diagnosis

However, our mentors formulated only a few diagnoses in their mentoring meetings. According to the model, the phases of a diagnosis (and a diagnostic check) are ideally placed between a phase of diagnostic question(s) and a phase of intervention(s). The absence of an explicit diagnosis can cause a lack of shared understanding between the mentor and the student about the students' research skills and, consequently, they cannot work on common goals, students do not understand mentor feedback properly, and mentors do not understand why their feedback is not used. These findings within the area of higher education and thesis mentoring fit the findings of Van de Pol et al. (2010) and Graesser et al. (1995), who argued that teachers find it hard to diagnose their students' skills.

Diagnostic check

As our mentors were seldom articulating a diagnosis explicitly to their students, diagnostic checks were scarcely uttered (only twice). We propose three explanations of why mentors do ask enough diagnostic questions, but do not explicitly share their diagnoses with their students and check their diagnoses. Firstly, mentors do not *have* the appropriate diagnostic skills available. Secondly, mentors fail to *enact* the appropriate available diagnostic skills. It seems mentors do have the appropriate skills available, as mentors showed they applied diagnostic questions and even formulated some diagnoses explicitly, but they are not always

capable of enacting these skills. A final explanation could be that the many diagnoses, including the actual process of determining the students' understanding (i.e. arriving at a diagnosis), were made implicitly.

Intervention

Our mentors, especially Amy, Claire and Debby, seemed to struggle with the balance between providing support (interventions) and allowing students to find their own way (asking questions). This fits the results of Todd et al. (2006), who found their mentors to be struggling with this balance as well. These three mentors mainly applied interventions and applied several interventions in sequence. The result of mentors, who provided support immediately without an explicit diagnosis, is similar to the findings of Ruiz-Primo and Furtak (2007). We propose two explanations as to why mentors apply so many interventions. Firstly, it could be that mentors are following their own objectives and own agenda (Nathan & Kim, 2009). Secondly, it could be that mentors are focusing on the difficulties they believe students are dealing with.

Limitations and future research

The four cases have provided insight into the diagnostic phases that are applied by mentors. However, there are some limitations to our multiple case study. This small-scale exploratory multiple case study was performed in one undergraduate program, which might limit the transferability of the findings. Another limitation arises from gathering data in the third week of the research project. These starting weeks of the research project might have led the study to focus on mentors who were very involved in the process and, as a result, might have shown many more interventions than a mentor who is less directly involved and who might provide less instruction and explanation in the finalizing phase of the thesis. We acknowledged the need for more empirical studies to investigate the diagnostic skills of mentors. We focused on how teachers were diagnosing their students. Future researchers could be focused on the implicit diagnosis. Furthermore, studies on mentors who have more mentoring experience could be a topic of research. For example, one could investigate whether they would apply more explicit diagnoses. Further interesting avenues for research include investigating the possible differences between research mentoring of the development of a research plan and the writing of a final research report.

Practical implications

It is necessary that mentors get a sense of what good diagnostic skills are, in order to be better at judging their students' research skills (Südkamp et al., 2012). Mentors should be offered formal mentor training to be kept updated on the

demands of the research agenda (McCallin & Nayar, 2012). When mentors do not have the appropriate diagnostic skills available, a good strategy could be instruction through video (Van Es & Sherin, 2010) and guided reflection (McCullagh, 2012). When mentors fail to *enact* the appropriate available diagnostic skills, mentors could be provided with a list of prompts, as examples of the kind of questions mentors can ask (Chi et al., 2001), and this may improve diagnostic questioning, diagnosis articulating, and diagnostic checking. Mentoring meetings can become more interactive when mentors are trained to suppress their explanations and feedback and, instead, are trained to prompt students with questions.

Conclusion

The context of the mentoring meeting with a one-to-one (or one-to-two) interaction did not enable our mentors to use diagnostic questions automatically to articulate their diagnosis of their students' research skills explicitly. Instead, they mainly use interventions. Since diagnostic questions and the articulation of the diagnosis are conducive to mentoring that is adapted to students' needs, mentors need to be skilled diagnosticians. From this point of view, a goal should be to make teacher mentors more aware of their lack of diagnostic skills and to stimulate the development of strategies for effectively supporting their students' research skills.

Acknowledgement

The authors would like to thank the participating supervisors for observing their experiences, Pien Hermans, Sophie Schouten, and Gijsbert Erkens for their valuable help with data preparation and analyses.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the Dutch Organization for Scientific Research NWO with grant number [023.002.122].

Notes on contributors

Bas T. Agricola is a doctoral candidate at Utrecht University, Department of Education of the Faculty of Social Sciences and a lecturer at HAN University of Applied Sciences. His research aims to understand student-teacher interactions during feedback conversations in higher education. His educational focus is on mentoring students during their undergraduate thesis projects.

Frans J. Prins is an associate professor at Utrecht University, Department of Education, and Educational Consultancy & Professional Development of the Faculty of Social and Behavioural Sciences. His research interests are assessment, feedback, and motivation. His educational focus concerns practice-oriented research in the pre-service teacher education.

Marieke F. van der Schaaf is a professor of Research and Development of Health Professions Education at University Medical Center Utrecht and a researcher at the department of Education at Utrecht University. Her research focuses on quality improvement of formative and summative assessments, and how assessment can stimulate professionals' learning.

Jan van Tartwijk is a professor of applied educational sciences at Utrecht University, Department of Education of the Faculty of Social Sciences. His research, focuses on teacher-student communication processes in the classroom, on learning and assessment at the workplace, on assessment, motivation, and creativity, and the development of teacher expertise.

References

Basturkmen, H., East, M., & Bitchener, J. (2014). Supervisors' on-script feedback comments on drafts of dissertations: Socialising students into the academic discourse community. *Teaching in Higher Education*, *19*(4), 432–445.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, *13*(4), 544–559. https://nsuworks.nova.edu/tqr/vol13/iss4/2

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, *6*(3), 271–315.

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471–533.

Chin, C. (2006). Classroom interaction in science: Teacher questioning and feedback to students' responses. *International Journal of Science Education*, *28*(11), 1315–1346.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice*, *39*(3), 124–130.

de Kleijn, R. A. M., Bronkhorst, L. H., Meijer, P. C., Pilot, A., & Brekelmans, M. (2014). Understanding the up, back, and forward-component in master's thesis supervision with adaptivity. *Studies in Higher Education*, 1–17. doi:10.1080/03075079.2014.980399

de Kleijn, R. A. M., Mainhard, M. T., Meijer, P. C., Brekelmans, M., & Pilot, A. (2012). Master's thesis supervision: Relations between perceptions of the supervisor–Student relationship, final grade, perceived supervisor contribution to learning and student satisfaction. *Studies in Higher Education*, *37*(8), 925–939.

de Kleijn, R. A. M., Meijer, P. C., Brekelmans, M., & Pilot, A. (2015). Adaptive research supervision: Exploring expert thesis supervisors' practical knowledge. *Higher Education Research and Development*, *34*(1), 117.

Engebretson, K., Smith, K., McLaughlin, D., Seibold, C., Terrett, G., & Ryan, E. (2008). The changing reality of research education in australia and implications for supervision: A review of the literature. *Teaching in Higher Education*, *13*(1), 1–15.

Erkens, G. (2005). *Multi episode protocol analysis (MEPA)* (Version 4.10 ed.). Computer Software. The Netherlands: Utrecht University.. http://edugate.fss.uu.nl/mepa

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, *9*(6), 495–522.

Gravetter, F. J., & Wallnau, L. B. (2013). *Statistics for the behavioral sciences* (9th ed. ed.). New York, NY: Cengage Learning.

Greenbank, P., & Penketh, C. (2009). Student autonomy and reflections on researching and writing the undergraduate dissertation. *Journal of Further and Higher Education*, *33*(4), 463–472.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology*, *29*(2), 75.

Halse, C., & Malfroy, J. (2010). Retheorizing doctoral supervision as professional work. *Studies in Higher Education*, *35*(1), 79–92.

Heath, T. (2002). A quantitative analysis of PhD students' views of supervision. *Higher Education Research & Development*, *21*(1), 41–53.

Hedin, L. R., & Gaffney, J. S. (2013). Tutoring sixth graders who struggle with reading: Illustrations of wood's contingent interventions. *Reading Psychology*, *34*(3), 207–256.

Hoth, J., Döhrmann, M., Kaiser, G., Busse, A., König, J., & Blömeke, S. (2016). Diagnostic competence of primary school mathematics teachers during classroom situations. *ZDM Mathematics Education*, *48*(1), 41–53.

Jaldemark, J., & Lindberg, J. O. (2013). Technology-mediated supervision of undergraduate students' dissertations. *Studies in Higher Education*, *38*(9), 1382–1392.

Jeong, A., Clark, D. B., Sampson, V. D., & Menekse, M. (2011). Sequential analysis of scientific argumentation in asynchronous online discussion environments. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing interactions in CSCL. methodology, approaches, and issues* (pp. 207–233). New York, NY: Springer.

Klug, J., Bruder, S., Kelava, A., Spiel, C., & Schmitz, B. (2013). Diagnostic competence of teachers: A process model that accounts for diagnosing learning behavior tested by means of a case scenario. *Teaching and Teacher Education*, *30*, 38–46.

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, *30*(3), 411–433.

Manathunga, C. (2005). Early warning signs in postgraduate research education: A different approach to ensuring timely completions. *Teaching in Higher Education*, *10*(2), 219–233.

McCallin, A., & Nayar, S. (2012). Postgraduate research supervision: A critical review of current practice. *Teaching in Higher Education*, *17*(1), 63–74.

McCullagh, J. F. (2012). How can video supported reflection enhance teachers' professional development? *Cultural Studies of Science Education*, *7*(1), 137–152.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. New York, NY: Sage.

Nathan, M. J., & Kim, S. (2009). Regulation of teacher elicitations in the mathematics classroom. *Cognition and Instruction*, *27*(2), 91–120.

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education*, *35*(5), 501–517.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218.

Palmer, R. J., Hunt, A. N., Neal, M., & Wuetherick, B. (2015). Mentoring, undergraduate research, and identity development: A conceptual review and research agenda. *Mentoring & Tutoring: Partnership in Learning*, *23*(5), 411–426.

Putnam, R. T. (1987). Structuring and adjusting content for students: A study of live and simulated tutoring of addition. *American Educational Research Journal*, *24*(1), 13–48.

Riffe, D., Lacy, S., & Fico, F. G. (2005). Reliability. In *Analyzing media messages* (2nd., pp. 123–159). London, England: Lawrence Erlbaum Associates.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research of Science Teaching*, *44*(1), 57–84.

Sambrook, S., Stewart, J., & Roberts, C. (2008). Doctoral supervision… a view from above, below and the middle! *Journal of Further and Higher Education*, *32*(1), 71–84.

Seawright, J., & Gerring, J. (2008). Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political Research Quarterly*, *61*(2), 294–308.

Shanahan, J. O., Ackley-Holbrook, E., Hall, E., Stewart, K., & Walkington, H. (2015). Ten salient practices of undergraduate research mentors: A review of the literature. *Mentoring & Tutoring: Partnership in Learning*, *23*(5), 359–376.

Strijbos, J., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education*, *46*(1), 29–48.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, *104*(3), 743–762.

Todd, M., Bannister, P., & Clegg, S. (2004). Independent inquiry and the undergraduate dissertation: Perceptions and experiences of final-year social science students. *Assessment & Evaluation in Higher Education*, *29*(3), 335–355.

Todd, M., Smith, K., & Bannister, P. (2006). Supervising a social science undergraduate dissertation: Staff experiences and perceptions. *Teaching in Higher Education*, *11*(2), 161–173.

Van de Pol, J., & Elbers, E. (2013). Scaffolding student learning: A micro-analysis of teacher–Student interaction. *Learning, Culture and Social Interaction*, *2*(1), 32–41.

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, *22*(3), 271–297.

Van de Pol, J., Volman, M., & Beishuizen, J. (2011). Patterns of contingent teaching in teacher-student interaction. *Learning and Instruction*, *21*(1), 46–57.

Van Es, E. A., & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education*, *13*(2), 155–176.

Vehviläinen, S., & Löfström, E. (2014). 'I wish I had a crystal ball': Discourses and potentials for developing academic supervising. *Studies in Higher Education*, 1–17. doi:10.1080/03075079.2014.942272

Wampold, B. E., & Margolin, G. (1982). Nonparametric strategies to test the independence of behavioral states in sequential data. *Psychological Bulletin*, *92*(3), 755–765.

Wisker, G. (2009). *The undergraduate research handbook*. Basingstoke, England: Palgrave Macmillan.

Yin, R. K. (2013). *Case study research. design and applied. social research methods series, vol. 5* (5th ed. ed.). Thousand Oaks, CA: SAGE Publications Inc.