Diagnosing Scientific Reasoning Skills in Inquiry Settings: The Role of Pre-Service Teachers' Professional Knowledge

Amadeus J. Pickal, University of Hildesheim, pickal@uni-hildesheim.de
Clark A. Chinn, Department of Educational Psychology, Rutgers University, clark.chinn@gse.rutgers.edu
Raimund Girwidz, Chair of Physics Education, LMU Munich, girwidz@lmu.de
Birgit J. Neuhaus, Biology Education, Faculty of Biology, LMU Munich, birgit.neuhaus@lrz.uni-muenchen.de
Christof Wecker, University of Hildesheim, christof.wecker@uni-hildesheim.de

Abstract: This study extends the research on the influence of teachers' professional knowledge on their diagnostic competences by evaluating the influence of professional knowledge on the diagnosis of cross-domain skills such as scientific reasoning. Biology and physics pre-service teachers completed tests for their scientific reasoning skills, domain knowledge, and knowledge about the structure and diagnosis of scientific reasoning skills, and diagnosed a student's scientific reasoning skills based on a video simulation. The results suggest that the influence of different aspects of professional knowledge may vary between different subjects, and that domain knowledge alone may have an effect on the accuracy of the diagnoses. So far, this could only be established for biology pre-service teachers. Since this study is still in progress, the results have to be considered preliminary.

Keywords: Pre-service teacher education, video simulations, scientific reasoning

Introduction

In order to be able to provide optimally targeted support for their students' learning progress, schoolteachers have to rely on diagnostic competences concerning their students' knowledge and skills (Schrader, 2009). As it is the case for various other competences, diagnostic competences may depend on a teachers' professional knowledge (see Baumert et al., 2010; Förtsch et al., 2018). Research on diagnostic competences has focused mainly on diagnosing subject-specific knowledge and skills (Südkamp, Kaiser, & Möller, 2012). This is the case even though cross-domain skills such as scientific reasoning are also important educational goals. As cross-domain skills can be fostered in different school subjects, teachers in these different subjects may all have to contribute to diagnosing them and fostering their development (Wecker, Hetmanek, & Fischer, 2016).

In this paper, we focus on scientific reasoning skills as an exemplary case of cross-domain skills. Scientific reasoning skills comprise a combination of several sub-skills all concerning mental processes regarding scientific activities such as formulating hypotheses, conducting experiments or drawing conclusions (e.g. de Jong & van Joolingen, 1998; Dunbar & Fugelsang, 2007; Klahr & Dunbar, 1988). Scientific reasoning skills play a particularly important role in natural science subjects such as biology or physics. Here teachers often use inquiry settings for implementing and teaching scientific reasoning skills. Although the appropriateness of such tasks is debated in terms of their authenticity with respect to real life scientific inquiry, they are still a standard approach (Chinn & Malhotra, 2002). Because experimental research on the diagnosis of scientific reasoning skills is hard to implement in regular school lessons, an appropriate approach to investigate factors influencing the accuracy of diagnoses of learners' scientific reasoning skills may be simulations (cf. Südkamp, Möller, & Pohlmann, 2008). Simulations are representations of certain segments of reality that offer the possibility to manipulate specific parameters (Heitzmann et al., 2019). In line with this idea, we developed a video simulation that shows students collaborating during inquiry tasks in school and thereby provides the context for diagnosing the scientific reasoning skills of one of the collaborating students.

Based on the research on diagnostic competences mentioned above, we expect several facets of (pre-service) teachers' professional knowledge (Shulman, 1986) to be important for diagnosing scientific reasoning. One aspect of professional knowledge is teachers' own mastery of the knowledge and skills in question (so-called content knowledge, CK). In the case of scientific reasoning skills as an educational goal, on the one hand this role is played by teachers' own *scientific reasoning skills*. Because, however, scientific reasoning can only be taught in the context of a specific topic (e.g. experimenting with lenses in a physics lesson), (pre-service) teachers' own *domain knowledge* may also influence the accuracy of their diagnoses. Another important aspect of teachers' professional knowledge is pedagogical content knowledge (PCK), which includes knowledge about cognitive structures, typical barriers for understanding, common errors, and instructional ways to overcome them within a particular domain. In the case of scientific reasoning skills as the educational goal in question, this includes *knowledge about the structure and diagnosis of scientific reasoning* (including common student mistakes, etc.).

Against this background, we investigated the following research questions:

1. To what extent do pre-service teachers' own *scientific reasoning skills*, *domain knowledge* and *knowledge about the structure and diagnosis of scientific reasoning skills* influence their *accuracy* in diagnosing learners' scientific reasoning skills?
2. Is the effect of professional knowledge on the *accuracy* similar in two different school subjects (biology and physics education)?

Method

Participants

So far 55 pre-service teachers specializing in biology education (age: $M = 22.36$, $SD = 2.47$; 57 % female; semester: $M = 5.25$, $SD = 1.94$) and 14 pre-service teachers specializing in physics education (age: $M = 24.14$, $SD = 5.52$; 21 % female; semester: $M = 4.07$, $SD = 2.37$) have participated in this study. All of them were enrolled in university-based teacher education programs in two federal states of Germany at the time of participation.

Design

The design was correlational with respect to the predictors *scientific reasoning skills* and *domain knowledge*. Because low values and restricted spontaneous variation were expected for *knowledge about the structure and diagnosis of scientific reasoning skills*, this variable was experimentally manipulated as a between subjects-factor by means of a written text on background information in a control-group design.

Operationalization of independent variables

The participants' own *scientific reasoning skills* were tested with seven items of the "Classroom Test of Scientific Reasoning" (Lawson, 1978). Five of the items consisted of two parts, asking both for an answer and for a reason for the answer. These items are scored as correct only if both parts are answered correctly. Hence, scores could range from zero to seven points. The internal consistency of the scale amounted to Cronbach's $\alpha = 0.59$.
Domain knowledge was tested by two subject-specific tests each comprising four two-part questions. For the biology pre-service teachers, this test was a shortened version of the test on plant growth by Lin (2004). For the physics pre-service teachers, this test was a structurally similar, custom-made test on optical lenses. In both tests, scores could range from zero to four. The internal consistency of the scale amounted to Cronbach's $\alpha = 0.50$ for the biology test. At this point, the physics test was completed by only 14 participants and will be evaluated once the data collection is further progressed.

As we anticipated low values and restricted spontaneous variation for *knowledge about the structure and diagnosis of scientific reasoning skills*, we manipulated this knowledge by providing half of the participants in both subjects with a 2-page text containing information about the sub-skills of scientific reasoning and guidelines on how to diagnose them. The sub-skills were grouped under the headings "formulating hypotheses" (e. g. "precision of hypotheses"), "designing and conducting experiments" (e. g. "control of variables") and "drawing conclusions" (e. g. "detecting confounded experiments"). Within the two subjects, pre-service teachers were randomly assigned to one of the two conditions with or without the text. A treatment check was implemented by means of two open-ended questions that were administered immediately after reading the text and that probed knowledge about the structure and diagnosis of scientific reasoning skills. A statistically significant positive effect of reading the text supported the efficacy of the manipulation in terms of the participants' knowledge about the structured and diagnosis of scientific reasoning skills; $t(67) = 4.99$; $p < .001$; $d = 1.21$.

Simulation environment

To create a scenario for the measurement of pre-service teachers' diagnostic competences as reflected in the *accuracy* of their diagnoses, we developed two video-based simulations that were structured in parallel – one for biology and one for physics. First the pre-service teachers were introduced to the simulation and to the task of collecting information and writing a diagnosis of the scientific reasoning skills of a student who can be observed in the simulation video while collaborating with a peer on a series of experiments in an inquiry setting. In the video simulation, the screen was divided into four parts. In the *video area* that covered the largest part of the screen, a video of the two students was displayed. Because control of the levels of scientific reasoning skills was a concern, the videos were staged, i. e. the students in the video played their roles based on written scripts. The scripts were developed based on skill profiles that contain information about all the respective student's mastery of each of the sub-skills. On this basis, it was also ensured that both simulations had an identical structure. In both subjects, the videos showed the same two female students conducting experiments to find out whether and, if so, how each of four independent variables influences a particular outcome variable. In biology, the outcome variable was the growth of a sprouted seed, and the four independent variables were the amount of water, salt, a fertilizer stick, and an undefined white powder. In physics, the outcome variable was the distance between a lens and a screen at which the image of an object was depicted in focus, and the four independent variables were the distance of the lens and the projected object, the size of the lens, the curvature of the lens, and a so-called polarizing filter. In both videos, the students formulated hypotheses, conducted experiments to test these hypotheses, and drew conclusions from their observations. The screen of the video simulation also displayed an *inquiry table* that incrementally documented the research questions and hypotheses, settings of independent variables, measurements of the outcome variable, and conclusions from the experiments the two students in the video conducted. The *navigation area* displayed links to "interact" with the students in the video by asking them pre-formulated questions. These video links appeared and disappeared on the screen at certain points during the simulation. When participants decided to ask one of the questions (e. g. "What do you want to find out now?"), the answer was presented via a video segment that was inserted in the main video in the video area. After each selected video segment with a question and the corresponding answer, the main video continued. As there was only limited time for asking questions, participants were encouraged only to ask questions that were relevant to their task of collecting information about the target student's scientific reasoning skills. A *notepad* was provided in the fourth part of the screen to help the participants keep track of their diagnostic conjectures and observations.

Measurement of the dependent variable

In order to determine the *accuracy* of the diagnoses, the participants' diagnoses were coded in terms of quantity and quality: All scientific reasoning sub-skills mentioned in the diagnoses (the same as the ones mentioned in the text for manipulating the *knowledge about the structure and diagnosis of scientific reasoning skills*) were coded, and then their number was counted based on this analysis. Additionally, all scientific reasoning sub-skills that were diagnosed in agreement with the competence profile underlying the video case were coded before the number of correctly diagnosed scientific reasoning sub-skills was counted. The *accuracy* score was defined as the achieved percentage of the 23-point-maximum from both counts; including an additional point for the awareness for the differentiation between actual behavior and underlying latent sub-skill.

Results

Two separate regression analyses with the *accuracy* of the diagnoses as the criterion variable and participants' own *scientific reasoning skills*, *domain knowledge,* and *text* (no text vs. text) as predictors were conducted for biology and physics pre-service teachers. The results from the two analyses are displayed in Table 1.

Table 1: Standardized regression coefficients (with standard errors) for the predictors of accuracy in diagnosing scientific reasoning skills

	Biology pre-service teachers	Physics pre-service teachers
Scientific reasoning skills	.17 (*SE* = 0.06)	.28 (*SE* = 0.41)
Domain knowledge	.26* (*SE* = 0.05)	.18 (*SE* = 0.43)
Text	.04 (*SE* = 0.03)	.13 (*SE* = 0.13)
R^2	.13 (*SE* = 0.10)	.13 (*SE* = 0.23)

Note. $^{*}p < .05$ (one-sided)

Although the proportion of variance accounted for was substantial and identical in the two subjects, the findings concerning the individual predictors differed. In biology, domain knowledge was the only statistically significant predictor. In physics, none of the predictors was statistically significant, although the regression coefficient for domain knowledge was of similar size as in biology. One needs to bear in mind, however, that data collection in physics is still in progress, and the sample included in the analysis was rather small.

Discussion

According to the preliminary findings reported, the only aspect of pre-service teachers' professional knowledge that could be established as a predictor of the *accuracy* of their diagnoses was *domain knowledge*, and this finding was restricted to the biology pre-service teacher sub-group. When comparing the results between the biology and physics pre-service teachers, however, the descriptive patterns were fully parallel. In fact, for the physics sub-group, descriptively, the pre-service teachers' own *scientific reasoning skills* had the highest regression coefficient. Due to the still small sample size in physics, it is currently impossible to determine whether this high coefficient reflects a true relation between these variables or whether it is merely due to chance. As we are still in the process of collecting data, the results have to be considered preliminary. With these reservations in mind, these preliminary findings seem to suggest that professional knowledge does not play a large role in the diagnosis of scientific reasoning skills. If this should be the case, there might be other factors with greater predictive value.

However, limitations of the current analysis include the rather low reliability of several instruments, although the internal consistency of tests for rather broad fields of knowledge rarely reaches very high levels. Low reliability may lead to downward-biased estimates of regression coefficients. Hence, amongst other things, future research will have to bare this in mind before dependable consequences about their role for the diagnosis of cross-domain skills can be drawn.

References

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180. https://doi.org/10.3102/0002831209345157

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*(2), 175–218. https://doi.org/10.1002/sce.10001

De Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of educational research, 68*(2), 179. https://doi.org/10.2307/1170753

Dunbar, K., & Fugelsang, J. (2007). Scientific thinking and reasoning. In K. J. Holyoak (Ed.), *The Cambridge handbook of thinking and reasoning* (pp. 705–725). Cambridge: Cambridge Univ. Press.

Heitzmann, N., Seidel, T., Hetmanek, A., Wecker, C., Fischer, M. R., Ufer, S., ... & Obersteiner, A. (2019) Facilitating Diagnostic Competences in Simulations in Higher Education. *Frontline Learning Research, 7*(4), 1-24.

Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M., Girwidz, R., Obersteiner, A., . . . Neuhaus, B. (2018). Systematizing professional knowledge of medical doctors and teachers: Development of an interdisciplinary framework in the context of diagnostic competences. *Education Sciences, 8*(4), 207. https://doi.org/10.3390/educsci8040207

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive science, 12*, 1–48.

Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, *15*(1), 11–24.

Lin, S.-W. (2004). Development and application of a two-tier diagnostic test for high school students' understanding of flowering plant growth and development. *International Journal of Science and Mathematics Education*, *2*(2), 175–199. https://doi.org/10.1007/s10763-004-6484-y

Schrader, F.-W. (2009). Anmerkungen zum Themenschwerpunkt Diagnostische Kompetenz von Lehrkräften [Notes on the issue of diagnostic comeptences of teachers]. *Zeitschrift für Pädagogische Psychologie, 23*(34), 237–245. https://doi.org/10.1024/1010-0652.23.34.237

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, *15*(2), 4-14.

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. https://doi.org/10.1037/a0027627

Südkamp, A., Möller, J., & Pohlmann, B. (2008). Der simulierte Klassenraum: Eine experimentelle Untersuchung zur diagnostischen Kompetenz. [The simulated classroom: An experimental study on diagnostic competence]. *Zeitschrift für Pädagogische Psychologie, 22*(34), 261–276. https://doi.org/10.1024/1010-0652.22.34.261

Wecker, C., Hetmanek, A., & Fischer, F. (2016). Zwei Fliegen mit einer Klappe? Fachwissen und fächerübergreifende Kompetenzen gemeinsam fördern. [Two birds with one stone? Fostering domain-specific knowledge and cross-domain competences simultaneously]. *Unterrichtswissenschaft, 44*(3), 226–238.

Acknowledgments

The research presented in this paper was funded by a grant of the Deutsche Forschungsgemeinschaft (DFG) to Christof Wecker, Birgit Neuhaus and Raimund Girwidz (WE 5426/2-1).