


Article

Patterns of Scientific Reasoning Skills among Pre-Service Science Teachers: A Latent Class Analysis

Samia Khan ^{1,*} and Moritz Krell ² ¹ Department of Curriculum and Pedagogy, University of British Columbia, Vancouver, BC V5K 0A1, Canada² Leibniz Institute for Science and Mathematics Education, University of Kiel, 24103 Kiel, Germany; krell@leibniz-ipn.de

* Correspondence: samia.khan@ubc.ca

Abstract: We investigated the scientific reasoning competencies of pre-service science teachers (PSTs) using a multiple-choice assessment. This assessment targeted seven reasoning skills commonly associated with scientific investigation and scientific modeling. The sample consisted of 112 PSTs enrolled in a secondary teacher education program. A latent class (LC) analysis was conducted to evaluate if there are subgroups with distinct patterns of reasoning skills. The analysis revealed two subgroups, where LC1 (73% of the PSTs) had a statistically higher probability of solving reasoning tasks than LC2. Specific patterns of reasoning emerged within each subgroup. Within LC1, tasks involving analyzing data and drawing conclusions were answered correctly more often than tasks involving formulating research questions and generating hypotheses. Related to modeling, tasks on testing models were solved more often than those requiring judgment on the purpose of models. This study illustrates the benefits of applying person-centered statistical analyses, such as LC analysis, to identify subgroups with distinct patterns of scientific reasoning skills in a larger sample. The findings also suggest that highlighting specific skills in teacher education, such as: formulating research questions, generating hypotheses, and judging the purposes of models, would better enhance the full complement of PSTs' scientific reasoning competencies.

Keywords: scientific reasoning; science teacher education; pre-service teachers; person-centered statistical analyses; latent class analysis



Citation: Khan, S.; Krell, M. Patterns of Scientific Reasoning Skills among Pre-Service Science Teachers: A Latent Class Analysis. *Educ. Sci.* **2021**, *11*, 647. <https://doi.org/10.3390/educsci11100647>

Academic Editor: Ian Hay

Received: 17 August 2021

Accepted: 13 October 2021

Published: 15 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Scientific reasoning has been a subject of study in the field of science education for some time [1]. Assessing this reasoning, however, remains a 21st century challenge for science educators today [2]. The present study is on the scientific reasoning of future science teachers themselves. We have assessed reasoning amongst this group because they will need to teach and demonstrate reasoning to their future students in science, and we can design activities in science teacher education that can enhance their competency in this field.

Scientific reasoning is a competency that encompasses the abilities needed for scientific problem-solving, as well as the capacity to reflect on problem-solving [3,4]. In the sciences, reasoning has been previously distinguished from other constructs such as problem-solving and critical thinking or scientific thinking alone. Descriptions of thinking, problem-solving, and reasoning are often conflated. For example, scientific reasoning has been suggested as being a kind of problem-solving; however, it has also been suggested that reasoning can be distinguished from problem-solving alone in that direct retrieval of a solution from memory is not possible with reasoning [5]. Ford [6] further reinforces that reasoning does not mean following a series of rules either but rather encompasses permanent evaluation and critique, as suggested by the reflective component of the above definition. Reasoning in the sciences requires cognitive processes that can contribute to, or allow for, inquiring and answering questions about the world and the nature of phenomena. These cognitive processes include

formulating and evaluating hypotheses, two of several processes regularly invoked in scientific domains [7,8].

The multiple cognitive processes that have been investigated in research on reasoning in science and science education have been variously described as formal logic, non-formal reasoning, creativity, model-based reasoning, abductive reasoning, analogical reasoning, and probabilistic reasoning [9–12]. These processes may or may not be used in the wider category of critical thinking [13]. Scholars have provided evidence that the ability to use these processes for reasoning is transferable across domains [14], while others such as Kind and Osborne [15] suggest that reasoning is highly variable by the content and the procedural and epistemic knowledge of the reasoner. Scholars have also shown that the ability to reason in science does not necessarily improve with age [16] but that it can be taught and enhanced in both the early years and at university levels [17–19].

Our focus in the present study is on the reasoning competencies of pre-service science teachers (PSTs) enrolled in a university teacher education program. Most studies on pre-service science teachers' scientific reasoning competencies adopt variable-centered approaches and report, for example, average scores for sample groups or populations. For example, one study [20] reported on a group of 66 Australian pre-service science teachers that they performed significantly better on tasks that required skills of 'planning investigations' compared to tasks related to skills of 'formulating research questions' and 'generating hypotheses'. Such insights are valuable but sometimes might be too rough-grained depending on the research questions, as different subgroups with distinct patterns of scientific reasoning skills exist within a sample. In order to identify such subgroups, person-centered analyses are necessary, that, statistically speaking, aim to "[R]educe the 'noise' in the data by splitting the total variability into 'between-group' variability and 'within-group' variability" [21] (p. 2). Hence, person-centered analyses, like latent class analyses (LCA), are finer-grained analyses in the sense that they are case-based and identify individuals with similar patterns of scientific reasoning skills (e.g., [22]). Person-centered analyses are also referred to as 'typological' approaches [23]. Such approaches can be specifically valuable for educators as they move beyond the 'average' and follow, methodologically, "[M]odern developmental theory, in which individuals are regarded as the organising unit of human development" [23] (p. 502). In the present study, we seek to establish whether subgroups of reasoners can be ascertained among PSTs using an LCA. The seven reasoning skills examined are: *formulating research questions, generating hypotheses, planning investigations, analyzing data and drawing conclusions, judging the purpose of models, testing models, and changing models*. While historical examination of scientific work has revealed that practices such as thought experiments, analogies, and imagistic simulation are important to scientists' development of new concepts [24], these seven skills under investigation were identified as key empirical areas of inquiry in science education [25–29] and likely having been taught in undergraduate science programs [3].

2. Materials and Methods

2.1. Sample

A full cohort of 56 PSTs from a university in North America participated in this study. Their mean age was 27 years ($SD = 6.34$; mode = 23). Data collection was done in their science teacher education secondary methods course within a Bachelor of Education after-degree program. To enroll in the secondary program, all students had at least one prior degree (usually 4 years of Science or more). The instrument described below (Section 2.2) was administered to the PSTs in their methods course at the beginning and at the end of the semester (pre–post-assessment). For the purpose of identifying groups with distinct patterns of scientific reasoning, we analyzed pre- and post-assessment data taken together of 56 PSTs. The total response sample for each item was thus $n = n_{pre} + n_{post}$ or $n = 112$. Only PSTs without missing responses have been included in the analysis, resulting in a sample of $n = 101$ for the statistical analysis. The number of PSTs by primary major were: Biology ($n = 30$), Chemistry ($n = 11$), Physics ($n = 8$), Biomedicine ($n = 1$), Earth Sciences

($n = 1$), Mathematics ($n = 1$), n/a ($n = 4$). Most of the PSTs' prior degrees were within the field of Biology ($n = 60$; e.g., general Biology, Applied Biology, or Evolutionary Biology), followed by Chemistry ($n = 25$) and Physics ($n = 6$).

2.2. Data Collection

An established multiple-choice instrument was administered to assess the PSTs' scientific reasoning competencies. The instrument was originally developed in the German language [27] and was later adapted into English, with thorough evaluations [30]. The instrument includes 21 multiple-choice items that were developed to assess seven reasoning skills of *formulating research questions, generating hypotheses, planning investigations, analyzing data and drawing conclusions, judging the purpose of models, testing models, and changing models*. Authentic scientific contexts were included in the items, which are mostly related to general science and Biology as well. As suggested in the organizing device that has been used for test development (see Table 1), these seven skills are related to two sub-competencies: conducting scientific investigations and using scientific models [31]. To correctly solve the multiple-choice items, PSTs have to apply their procedural and epistemic knowledge related to the respective skills [32–34]. Table 1 lists the two sub-competencies, their associated skills, and the specific knowledge necessary to correctly answer the items.

Table 1. Sub-competencies of scientific reasoning and associated skills with necessary procedural and epistemic knowledge, as described by Mathesius et al. [34].

Sub-Competencies	Skills	Necessary Knowledge PSTs Have to Know That . . .
Conducting scientific investigations	formulating questions	... scientific questions are related to phenomena, empirically testable, intersubjectively comprehensible, unambiguous, basically answerable and are internally and externally consistent.
	generating hypotheses	... hypotheses are empirically testable, intersubjectively comprehensible, clear, logically consistent and compatible with an underlying theory.
	planning investigations	... causal relationships between independent and dependent variables based on a previous hypothesis can be examined, whereby the independent variable is manipulated during experiments and control variables are considered. ... correlative relationships between independent and dependent variables based on a previous hypothesis can be examined with scientific observations.
	analyzing data and drawing conclusions	... data analysis allows an evidence-based interpretation and evaluation of the research question and hypothesis.
Using scientific models	judging the purpose of models	... models can be used for hypotheses generation.
	testing models	... models can be evaluated by testing model-based hypotheses.
	changing models	... models are changed if model-based hypotheses are falsified.

2.3. Data Analysis: Latent Class Analysis

A latent class analysis (LCA) was utilized to identify patterns of scientific reasoning skills among PSTs. The R package *poLCA* was employed [35]. All further (classical) statistical analyses, such as *t*-tests and descriptive analyses, were carried out with IBM

SPSS statistics, version 26. In an LCA, PSTs' responses are analyzed on the latent level, all variables are assumed to be (at least) on a nominal level, and there are no restrictions on the kind of relation between the (manifest) variables [33,36,37]. LCA was selected for data analysis because it permits the identification and computation of different groups (i.e., latent classes) of PSTs, with each group consisting of individuals with a response pattern that is as homogenous as possible (low within-group variability) but different from the response patterns of the other groups (high between-group variability). Therefore, LCA would be considered as belonging to the person-centered approaches of data analyses [21,23].

A core question of LCA is to decide on the appropriate number of latent classes [36]. To compare different LCA models, indices such as the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the sample size adjusted Bayesian information criterion (ssaBIC) are typically employed. These indices factor in the parsimony, the sample size, and the likelihood of the LCA models—each of the indices in a different manner [38]. When comparing different LCA models with these information indices, the smallest value of each index points out the comparatively best LCA model; however, the BIC and the ssaBIC were identified as superior indicators compared to the AIC [39] (p. 557), which is why these indicators are used in the present study. On the other hand, the BIC and the ssaBIC often do not identify the same LCA model as optimal [38]. Therefore, one has to use a combination of different insights to decide how many latent classes represent the data set best [38].

It is an important characteristic of LCA that the subjects are not assigned to the different latent classes in a deterministic manner but more so in a probabilistic sense. For diagnostic purposes, it is common to classify each subject to the latent class with the highest probability of assignment. Therefore, an “Additional indicator [of model-goodness] is the average membership probability within each [latent] class” [40] (p. 52); the higher this probability, the better the LCA model. Furthermore, one should analyze the item parameters for extreme values that indicate an estimated probability of 0% or 100% to solve a task; the fewer extreme values, the better the LCA model [40].

3. Results

Table 2 provides the fit-indices for the LCA models compared in this study. Because the BIC (2 latent classes) and ssaBIC (4 latent classes) suggest selecting different LCA models, the number of extreme values and the probability of assignment have been used as additional indicators. Based on these indicators, it can be assumed that the response pattern of the PSTs is best represented using two latent classes. These two latent classes consist of about 73% or 74 PSTs (latent class 1) and 27% or 27 PSTs (latent class 2) of the sample, respectively.

Table 2. Fit-indices of the different LCA models compared. Note that models with more than four latent classes did not fit the data.

LCA Model	BIC	ssaBIC	Extreme Values	Probability of Assignment
2 latent classes	2685	2549	0	0.93 to 0.98
3 latent classes	2722	2517	9	0.92 to 0.98
4 latent classes	2779	2504	11	0.91 to 0.97

Figure 1 illustrates the response profiles for the two latent classes across the seven skills of scientific reasoning covered in the multiple-choice instrument. Generally, PSTs in latent class 1 show a higher mean probability of correct answers within all seven skills. Comparing the mean probability of correct answers between the two latent classes with independent *t*-tests resulted in significant differences for the skills *planning investigations* ($p = 0.04$; $d = 0.48$, small to medium effect size measure), *analyzing data and drawing conclusions* ($p < 0.001$; $d = 1.25$, large effect size measure) as well as *judging the purpose of models*

($p < 0.001$; $d = 1.25$, large effect size measure), *testing models* ($p < 0.001$; $d = 1.49$, large effect size measure), and *changing models* ($p < 0.001$; $d = 0.88$, large effect size measure).

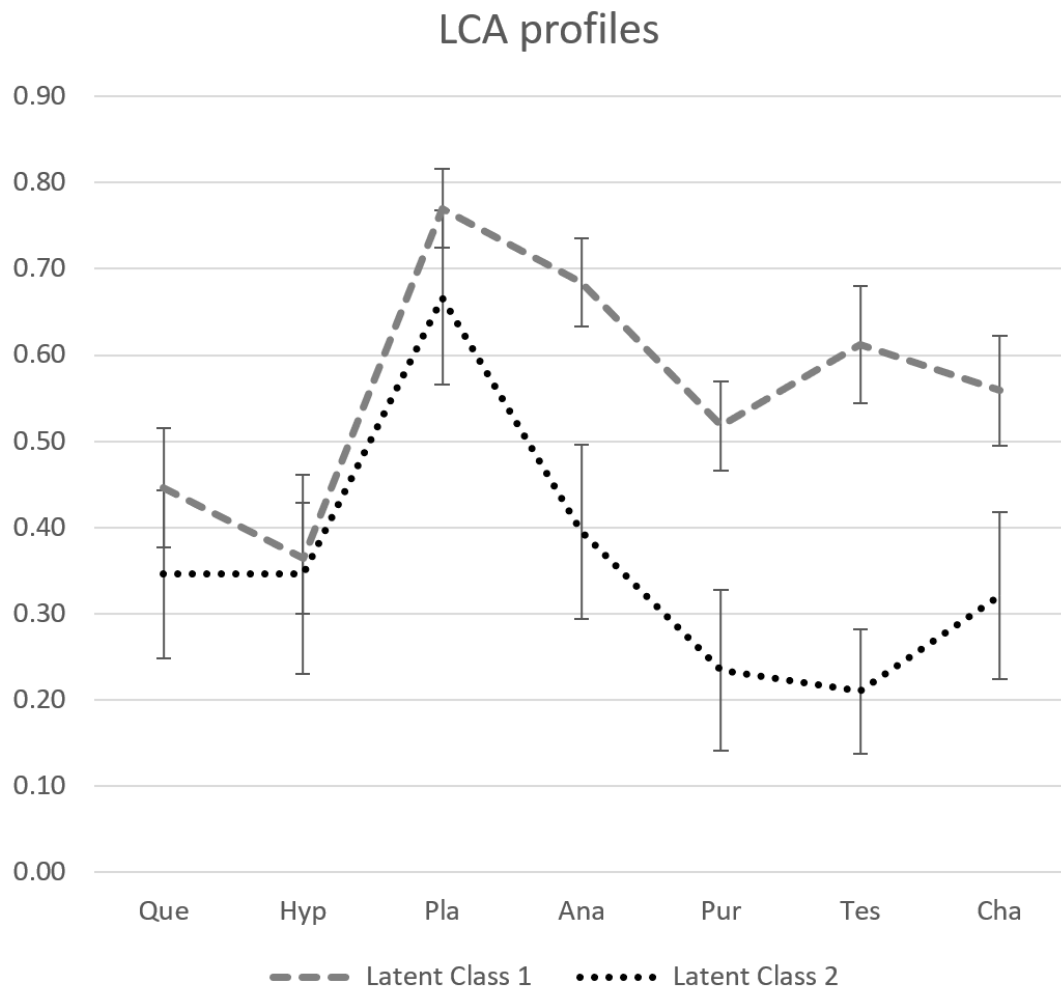


Figure 1. Response profiles for the two latent classes across the seven skills of scientific reasoning (mean score ± 2 * standard error).

For latent class 1 and considering skills related to conducting scientific investigations (Table 1), response probabilities for the skills *formulating research questions* and *generating hypotheses* on the one hand, and *planning investigations* and *analyzing data and drawing conclusions*, on the other hand, are quite similar, even though significant differences with large effect size measures could be found between these two groups of skills. For the skills related to using scientific models (Table 1), correct responses were found significantly more often for the skill *testing models* than for *judging the purpose of models* ($p = 0.02$; $d = 0.36$, small effect size measure).

For latent class 2 and considering skills related to conducting scientific investigations (Table 1), items related to the skill *planning investigations* have been answered correctly significantly more often than the tasks related to the other three skills ($p < 0.001$; $d > 1.00$, large effect size measures). For using scientific models (Table 1), no significant differences between the skills could be found.

In order to better understand the characteristics of the PSTs assigned to latent class 1 and latent class 2, we compared their age, primary majors, and the sum of previous degrees. Independent *t*-tests (Table 3) revealed that there are significantly more PSTs with the primary major of Biology in latent class 1 (about 65%) than in latent class 2 (about 33%). For the primary major of Chemistry, it is quite the reverse (about 15% in latent class 1 and about 33% in latent class 2); also, the number of PSTs with more than one previous degree

is significantly higher in latent class 1 ($n = 11$) than in latent class 2 ($n = 1$). These findings illustrate that the study of Biology as a primary major and a higher number of previous degrees made it more likely to belong to the more proficient latent class 1, whereas the study of Chemistry as a primary major made it more likely to belong to latent class 2.

Table 3. Comparison of the PSTs assigned to latent class (LC) 1 and LC 2 along the variables age, primary major of Biology, Chemistry or Physics, and the sum of previous degrees (the latter as a dichotomized variable with 1 = one previous degree and 2 = more than one previous degree).

Variable	LC Assignment	N	M	SD	t-Test
Age	1	74	26.54	5.35	$t(99) = 0.591; p = 0.556$
	2	27	27.30	6.55	
Biology	1	74	0.65	0.48	$t(99) = 2.918; p = 0.004$
	2	27	0.33	0.48	
Chemistry	1	74	0.15	0.36	$t(37.07) = 1.821; p = 0.077 *$
	2	27	0.33	0.48	
Physics	1	74	0.14	0.34	$t(99) = 0.316; p = 0.753$
	2	27	0.11	0.32	
Previous degrees	1	66	1.24	0.63	$t(78.93) = 2.072; p = 0.042 *$
	2	25	1.08	0.40	

* Adjusted t -statistic and df because of violated assumption of variance homogeneity.

4. Discussion

Using LCA, we revealed that two groups of reasoners emerged amongst the PSTs. One subgroup (latent class 1) had a statistically higher probability of solving scientific reasoning tasks than the other subgroup (latent class 2). Overall, the groups were significantly different on the following five skills out of seven investigated: *planning investigations*, *analyzing data and drawing conclusions*, *judging the purpose of models*, *testing models*, and *changing models*. They were not significantly different from each other on *formulating research questions* and *generating hypotheses*.

The latent class 1 subgroup responded significantly differently from each other on the skills *planning investigations* and *analyzing data and drawing conclusions* in contrast to the skills *formulating research questions* and *generating hypotheses*. Tasks about *testing models* were solved more often than those requiring *judging the purpose of models* within this subgroup. The latent class 2 subgroup responded significantly differently from each other on *planning investigations* compared to the other skills. For using scientific models, no significant differences could be found within this subgroup on the skills related to modeling (*judging the purpose of models*, *testing models*, and *changing models*).

These two subgroups also shared several other key characteristics. In latent class 1, a significant majority had a major in Biology compared to latent class 2, whereas there were far fewer from Chemistry in latent class 1. Moreover, there were significantly more PSTs with more than one previous degree in latent class 1 than in latent class 2. This finding is noteworthy for science teacher education because it suggests that Biology majors were significantly better at *planning investigations*, *analyzing data and drawing conclusions*, *judging the purpose of models*, *testing models*, and *changing models* than Chemistry majors. These findings might have been caused by the dominance of Biology-related items in the instrument; however, as the items require PSTs to apply procedural and epistemic knowledge as shown in Table 1 (and less so content knowledge), the findings lead us towards a renewed emphasis on reasoning tasks for Chemistry teacher education. Nevertheless, future studies could investigate the importance of science content knowledge from specific subjects (such as Biology) for solving the items, for instance, by applying think-aloud studies [25] or statistically investigating difficulty-generating task characteristics [41].

As a 'person-centered' statistical approach, the LCA was particularly powerful in ascertaining subgroups within a science teacher education cohort. This statistical approach

is a departure from traditional variable-centered approaches in education that tend to report on average scores for sample groups [21,23]. The LCA permits statistical cases to emerge from within samples or classrooms and is a recommended approach to generate case studies for further inquiry in science teacher education research.

In combination with relevant epistemic, procedural, and content knowledge, greater attention to *formulating research questions* and *generating hypotheses* would be helpful within science teacher education. Furthermore, reasoning tasks involving *judging the purpose of models* and *changing models* could be a high priority for modeling investigations in pre-service science teacher education. Possible science teacher education activities to support such tasks include the three-phased generating, evaluating, and modifying (GEM) models approach [10]. This approach emphasizes generating hypotheses in the first phase and testing and changing models in the second and third phases [42]. In general, science teacher education courses, Biology majors, or those with additional degrees could be purposefully included within heterogeneous groups for cooperative learning tasks. It was interesting to the authors that Biology majors outperformed other majors in this study, although this might be caused by the dominance of Biology-related items in the instrument; insights into the differences in performance among majors would be a helpful avenue for the design of science teachers education courses and group work in the ways suggested above. By participating in reasoning tasks with such recommendations in mind, future teachers might be able to better support their own students to develop competencies in these areas.

The significance of this study is that it identifies two groups of reasoners who are PSTs with different propensities to reason in science using person-centered statistics. Normally, the classroom would be treated similarly as an entire group; however, with this statistical approach, the researchers are able to show that subgroups of PSTs themselves emerged as competent at very different reasoning tasks. One subgroup is significantly more competent at *planning investigations, analyzing data and drawing conclusions, judging the purpose of models, testing models, and changing models* than the other. The subgroups had approximately equivalent competencies at *formulating research questions* and *generating hypotheses* showing for the first time that among PSTs, different subgroups with specific patterns of scientific reasoning skills exist. This finding can have an impact on science students of these future teachers, who presumably will draw upon their own competencies to demonstrate how to reason in the classroom. Future directions for research could target investigation and model-based reasoning competencies among PSTs and relationships to student reasoning. *Judging the purpose of models, formulating research questions, and generating hypotheses* were areas that PSTs were less competent; researching interventions related to these aspects of modeling and investigation would be worthwhile.

Author Contributions: Conceptualization, M.K., S.K.; methodology, M.K.; investigation, M.K., S.K.; resources, M.K., S.K.; writing—original draft preparation, S.K.; writing—review and editing, M.K., S.K.; visualization, M.K.; funding acquisition, M.K., S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the 2018 UBC-FUB Joint Funding Scheme, grant number FSP-2018-401.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of the University of British Columbia (ID H18-01801, approved 23 July 2018).

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

Data Availability Statement: Data is available upon request to the second author.

Acknowledgments: The authors wish to thank Alexis Gonzalez for support in data collection and tabulation.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

1. Lawson, A.E. The development and validation of a classroom test of formal reasoning. *J. Res. Sci. Teach.* **1978**, *15*, 11–24. [[CrossRef](#)]
2. Osborne, J. The 21st century challenge for science education: Assessing scientific reasoning. *Think. Ski. Creat.* **2013**, *10*, 265–279. [[CrossRef](#)]
3. Khan, S.; Krell, M. Scientific reasoning competencies: A case of preservice teacher education. *Can. J. Sci. Math. Technol. Educ.* **2019**, *19*, 446–464. [[CrossRef](#)]
4. Lawson, A.E. The nature and development of scientific reasoning: A synthetic view. *Int. J. Sci. Math. Educ.* **2004**, *2*, 307–338. [[CrossRef](#)]
5. Zimmerman, C. The development of scientific reasoning skills. *Dev. Rev.* **2000**, *20*, 99–149. [[CrossRef](#)]
6. Ford, M.J. Educational implications of choosing “practice” to describe science in the next generation science standards. *Sci. Educ.* **2015**, *99*, 1041–1048. [[CrossRef](#)]
7. Díaz, C.; Dorner, B.; Hussmann, H.; Strijbos, J.W. Conceptual review on scientific reasoning and scientific thinking. *Curr. Psychol.* **2021**, 1–13.
8. Reith, M.; Nehring, A. Scientific reasoning and views on the nature of scientific inquiry: Testing a new framework to understand and model epistemic cognition in science. *Int. J. Sci. Educ.* **2020**, *42*, 2716–2741. [[CrossRef](#)]
9. Krell, M.; Hergert, S. Modeling strategies. In *Towards a Competence-Based View on Models and Modeling in Science Education*; Upmeier zu Belzen, A., Krüger, D., van Driel, J., Eds.; Springer: Cham, Switzerland, 2020; pp. 147–160.
10. Khan, S. Model-based inquiries in chemistry. *Sci. Educ.* **2007**, *91*, 877–905. [[CrossRef](#)]
11. Babai, R.; Brecher, T.; Stavvy, R.; Tirosh, D. Intuitive interference in probabilistic reasoning. *Int. J. Sci. Math. Educ.* **2006**, *4*, 627–639. [[CrossRef](#)]
12. Upmeier zu Belzen, A.; Engelschalt, P.; Krüger, D. Modeling as scientific reasoning—The role of abductive reasoning for Modeling competence. *Educ. Sci.* **2021**, *11*, 495. [[CrossRef](#)]
13. Holyoak, K.J.; Morrison, R.G. Thinking and reasoning: A reader’s guide. In *Oxford Handbook of Thinking and Reasoning*; Holyoak, K.J., Morrison, R.G., Eds.; Oxford University Press: New York, NY, USA, 2005.
14. Kuhn, D.; Arvidsson, T.S.; Lesperance, R.; Corprew, R. Can engaging in science practices promote deep understanding of them? *Sci. Educ.* **2017**, *101*, 232–250. [[CrossRef](#)]
15. Kind, P.; Osborne, J. Styles of scientific reasoning: A cultural rationale for science education? *Sci. Educ.* **2017**, *101*, 8–31. [[CrossRef](#)]
16. Kuhn, D. What is scientific thinking and how does it develop? In *Handbook of Childhood Cognitive Development*; Goswami, U., Ed.; Blackwell: Oxford, UK, 2002; pp. 371–393.
17. Schauble, L. In the eye of the beholder: Domain-general and domain-specific reasoning in science. In *Scientific Reasoning and Argumentation: The Roles of Domain-Specific and Domain-General Knowledge*; Fischer, F., Chinn, C., Engelmann, K., Osborne, J., Eds.; Routledge: New York, NY, USA, 2018.
18. Dunbar, K.; Klahr, D. *Developmental Differences in Scientific Discovery Processes*; Psychology Press: Hove, UK, 2013; pp. 129–164.
19. Morris, B.J.; Croker, S.; Masnick, A.M.; Zimmerman, C. The emergence of scientific reasoning. In *Current Topics in Children’s Learning and Cognition*; Kloos, H., Morris, B., Amaral, J., Eds.; IntechOpen: London, UK, 2012; pp. 61–82.
20. Krell, M.; Dawborn-Gundlach, M.; van Driel, J. Scientific reasoning competencies in science teaching. *Teach. Sci.* **2020**, *66*, 32–42.
21. Kusurkar, R.A.; Mak-van der Vossen, M.; Kors, J.; Grijpma, J.W.; van der Burgt, S.M.; Koster, A.S.; de la Croix, A. ‘One size does not fit all’: The value of person-centred analysis in health professions education research. *Perspect. Med. Educ.* **2020**, *10*, 245–251. [[CrossRef](#)] [[PubMed](#)]
22. Krell, M.; zu Belzen, A.U.; Krüger, D. Students’ levels of understanding models and modeling in biology: Global or aspect-dependent? *Res. Sci. Educ.* **2014**, *44*, 109–132. [[CrossRef](#)]
23. Watt, H.M.; Parker, P.D. Person-and variable-centred quantitative analyses in educational research: Insights concerning Australian students’ and teachers’ engagement and wellbeing. *Aust. Educ. Res.* **2020**, *47*, 501–515. [[CrossRef](#)]
24. Nersessian, N.J. *Creating Scientific Concepts*; MIT Press: Cambridge, MA, USA, 2010.
25. Krell, M.; Redman, C.; Mathesius, S.; Krüger, D.; van Driel, J. Assessing pre-service science teachers’ scientific reasoning competencies. *Res. Sci. Educ.* **2020**, *50*, 2305–2329. [[CrossRef](#)]
26. Krell, M. Schwierigkeitserzeugende Aufgabenmerkmale bei Multiple-Choice-Aufgaben zur Experimentierkompetenz im Biologieunterricht: Eine Replikationsstudie [Difficulty-creating task characteristics in multiple-choice questions on experimental competence in biology classes: A replication study]. *Z. Didakt. Nat.* **2018**, *24*, 1–15.
27. Krüger, D.; Hartmann, S.; Nordmeier, V.; Upmeier zu Belzen, A. Measuring scientific reasoning competencies. In *Student Learning in German Higher Education*; Springer: Wiesbaden, Germany, 2020; pp. 261–280. [[CrossRef](#)]
28. Opitz, A.; Heene, M.; Fischer, F. Measuring scientific reasoning—A review of test instruments. *Educ. Res. Eval.* **2017**, *23*, 78–101. [[CrossRef](#)]
29. Bicak, B.E.; Borchert, C.E.; Höner, K. Measuring and Fostering Preservice Chemistry Teachers’ Scientific Reasoning Competency. *Educ. Sci.* **2021**, *11*, 496. [[CrossRef](#)]

30. Krell, M.; Mathesius, S.; van Driel, J.; Vergara, C.; Krüger, D. Assessing scientific reasoning competencies of pre-service science teachers: Translating a German multiple-choice instrument into English and Spanish. *Int. J. Sci. Educ.* **2020**, *42*, 2819–2841. [[CrossRef](#)]
31. Hartmann, S.; Upmeier zu Belzen, A.; Krüger, D.; Pant, H.A. Scientific reasoning in higher education. *Z. Psychol.* **2015**, *223*, 47–53. [[CrossRef](#)]
32. Mathesius, S.; Krell, M. Assessing modeling competence with questionnaires. In *Towards a Competence-Based View on Models and Modeling in Science Education*; Upmeier zu Belzen, A., Krüger, D., van Driel, J., Eds.; Springer: Cham, Switzerland, 2020; pp. 117–129.
33. Hagenaars, J.; Halman, L. Searching for ideal types: The potentialities of latent class analysis. *Eur. Sociol. Rev.* **1989**, *5*, 81–96. [[CrossRef](#)]
34. Mathesius, S.; Upmeier zu Belzen, A.; Krüger, D. Competencies of biology students in the field of scientific inquiry: Development of a testing instrument. *Erkenn. Biol.* **2014**, *13*, 73–88.
35. Linzer, D.; Lewis, J. poLCA: Polytomous Variable Latent Class Analysis. R Package Version 1.4. 2013. Available online: <http://dlinzer.github.com/poLCA> (accessed on 7 December 2020).
36. Collins, L.; Lanza, S. *Latent Class and Latent Transition Analysis*; Wiley: Hoboken, NJ, USA, 2010.
37. Langeheine, R.; Rost, J. (Eds.) *Latent Trait and Latent Class Models*; Plenum Press: New York, NY, USA, 1988.
38. Henson, J.; Reise, S.; Kim, K. Detecting mixtures from structural model differences using latent variable mixture modeling: A comparison of relative model fit statistics. *Struct. Equ. Model.* **2007**, *14*, 202–226. [[CrossRef](#)]
39. Nylund, K.; Asparouhov, T.; Muthén, B. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Struct. Equ. Model.* **2007**, *14*, 535–569. [[CrossRef](#)]
40. Spiel, C.; Glück, J. A model-based test of competence profile and competence level in deductive reasoning. In *Assessment of Competencies in Educational Contexts*; Hartig, J., Klieme, E., Leutner, D., Eds.; Hogrefe & Huber: Göttingen, Germany, 2008; pp. 45–68.
41. Krell, M.; Khan, S.; van Driel, J. Analyzing Cognitive Demands of a Scientific Reasoning Test Using the Linear Logistic Test Model (LLTM). *Educ. Sci.* **2021**, *11*, 472. [[CrossRef](#)]
42. Khan, S. New pedagogies on teaching science with computer simulations. *J. Sci. Educ. Technol.* **2011**, *20*, 215–232. [[CrossRef](#)]