Surveying preservice teachers' understanding of aspects of mathematics teaching – a cluster analysis approach

Nils Henry W. Rasmussen, Rune Herheim, Ragnhild Hansen, Troels Lange, Tamsin Meaney and Toril Eskeland Rangnes

Western Norway University of Applied Sciences, Norway

nhwr@hvl.no, rher@hvl.no, rhan@hvl.no, trl@hvl.no, tme@hvl.no, tera@hvl.no

Preservice teachers begin their teacher education with experiences that affect their possibilities for accessing and integrating new learning into their teaching practices. Yet often mathematics teacher education courses treat preservice teachers as a homogenous group. Responses to an electronic survey from the beginning of two compulsory mathematics education courses showed that preservice teachers could be divided into two clusters. The preservice teachers in each cluster give similar responses to different aspects of mathematics teaching, suggesting they share similar sets of views. These differences should be recognised in their future mathematics teacher education courses.

Keywords: Preservice teachers, cluster analysis, argumentation, modelling.

Introduction

Teaching is complicated, with teachers having to simultaneously consider a range of aspects. To ease this complexity, teachers often make assumptions about the groups they teach which then affects what they make available and in what ways and this in turn affects the potential learning of the students in their class (e.g. Hansen-Thomas & Cavagnetto, 2010). Similarly, it is important for teacher educators to know what views preservice teachers (PTs) hold on different aspects of mathematics teaching so they can adapt their teacher education appropriately. Traditionally, teacher educators have focused on what PTs should learn about mathematics and mathematics education (Ponte & Chapman, 2008), rather than what they may already know when entering teacher education. In our wider project, *Learning about teaching argumentation for critical mathematics education in multilingual classrooms* (LATACME), the aim is to improve two compulsory mathematics education courses for teachers of grades 1 to 7 in regard to argumentation, critical mathematics education, multilingual classrooms, mathematical modelling and the use of digital tools. The impetus for this comes from the implementation of a new curriculum in August 2020 in which "reasoning and argumentation" and "modelling and application" are two of six core elements and digital skills is one of five "basic skills" (Utdanningsdirektoratet, 2019), and the requirement that PTs have an "awareness of cultural differences and being able to use these as a positive resource" (National Council for Teacher Education, 2016, p. 9). It is important to identify PTs' existing understanding of these different aspects of mathematics education when they begin these compulsory teacher education courses.

This paper is one of two papers presented at this conference (see also Meaney et al., in press), on the results of an electronic survey about the PTs' initial understandings about the specific aspects of mathematics teaching. The survey was completed at the beginning of the two compulsory courses. In this paper, we focus on the questions that allowed us to identify: 1) if there are clusters in the cohort, and 2) if there were clusters, what were the differences between these clusters. The responses to these questions provide background for the planning of our teacher education courses, which will be discussed in a later paper.

Previous cluster analysis research on preservice teachers

To identify groups requires an investigation into whether there are different sets of understandings about mathematics teaching within a cohort and whether these are held by particular groups of PTs. In mathematics teacher education, cluster analysis has typically been used to identify key aspects of different groups of PTs. For example, using questionnaire data from Finnish elementary PTs' views on mathematics, Hannula, Kaasila, Laine, and Pehkonen (2005) found three main PT profiles. These were characterized by *positive* (43%), *neutral* (36%), or *negative* views (22%) of mathematics. Each group could be split into two sub-groups, based on the amount of encouragement the PTs received from family and on their view of themselves as hard-working. The PTs with a positive view could be split into the two groups *autonomous* (21%) and *encouraged* (22%). The PTs with a neutral view could be split in *pushed* (18%) and *diligent* (18%), and the ones with a negative view could be split in *lazy* (18%) and *hopeless* (4%) – the latter group believing that they could not learn mathematics.

Cluster analysis has also been used to determine whether groups of PTs see language diversity in mathematics classrooms differently (McLeman & Fernandes, 2012; McLeman, Fernandes, & McNulty, 2012). McLeman and Fernandes (2012) investigated the beliefs of 334 PTs from across the USA about the mathematics education of English-as-a-second-language learners (ELs). The PTs completed an online questionnaire in which they ranked 26 items using a five-point Likert scale (strongly disagree to strongly agree). They identified two clusters, which differed according to PTs' beliefs about: parents from some cultures placing a higher value on education than others; limiting mathematics vocabulary to make the content clear to ELs; and creating discussion-rich classrooms as necessary for ELs to learn mathematics. With some variation, the PTs generally expressed deficit views on culture and family and parents' support for their children's education, which seemed to be quite resistant to change by teacher educators and teaching experiences. The PTs did, however, believe mathematics to be an ideal subject to support students who were learning English.

As a follow-up study, McLeman et al. (2012) investigated 292 PTs' conceptions about the mathematics education of ELs, with a similar analysis to that in McLeman and Fernandes (2012). Their results indicated that PTs' exposure to issues to do with ELs affected their conceptions about their mathematics education. They also documented that apart from the item about discussion-rich, mathematics classrooms being beneficial for ELs, the PTs in cluster 1 held views which were not in alignment with research. For instance, the means were low for claims like "Learning English is more important than native language", "Speaking in a language other than English hampers the learning of English", and "Conversational fluency implies capability to learn math like non-English learners".

Methodology

The PTs in this study have two mandatory mathematics courses of 15 ECTS each. The courses are taught in the second and third semesters and integrate mathematics and mathematics education. The questionnaire was administered at the beginning of the second semester, after one semester of teacher education, including a three week practicum period, but before exposure to mathematics teacher education. Thus, the responses to the questionnaire can be assumed to be related to the PTs' school mathematics experiences and early experiences with teacher education.

A digital Likert-scale survey was designed with the PTs being asked to which degree they agreed or disagreed on a total of 51 claims concerning argumentation, digital tools, mathematical modelling, multilingual classrooms and critical mathematics education. Of the approximately 200 PTs in the cohort, 96 completed the questionnaire.

On argumentation, the PTs were presented with Grade-4-student responses to the task: “Why is the sum of two odd numbers always an even number?” Ben’s answer is shown in Figure 1. The PTs were asked to what extent they agreed to three claims: “I can understand and follow Ben’s explanation”; “Ben’s explanation is incomplete”; and “Ben’s explanation is mathematically correct”. The “Don’t know” category, together with the “Neither agree nor disagree” one, reduces the extent of random answers and has been shown not to contribute to lower validity (Lozano, García-Cueto, & Muñiz, 2008; O'Muircheartaigh, Krosnick, & Helic, 2000; Wang & Krosnick, 2020).

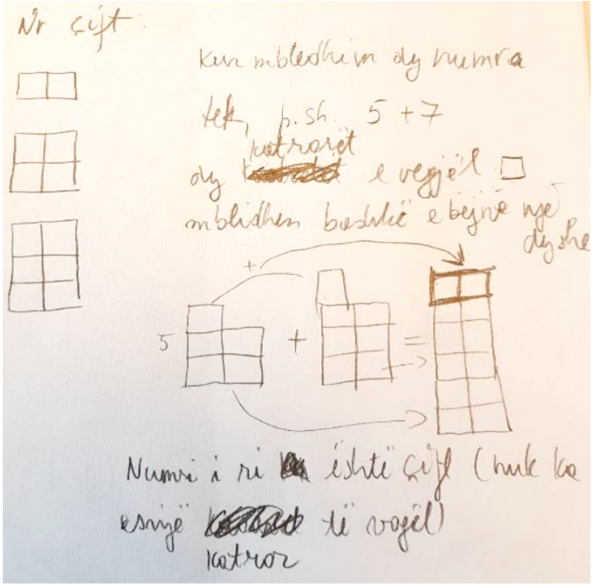

Figure 1: Ben’s explanation

In relationship to modelling, the PTs were also asked to rate claims about a project for Grade 5 on air pollution (Figure 2 shows a diagram from the project). Two examples of claims were: “I would not have used this project in my teaching because the students would find it too extensive" and "The project would take too much time from teaching and would not allow us to get through everything we are supposed to do in the book". The PTs were also presented with eight claims about modelling in mathematics teaching. Among these were “Students in primary school are too young for modelling to contribute to increasing their critical judgment ability about how mathematics is used in society” and “Modelling improves students’ attitudes towards mathematics”.

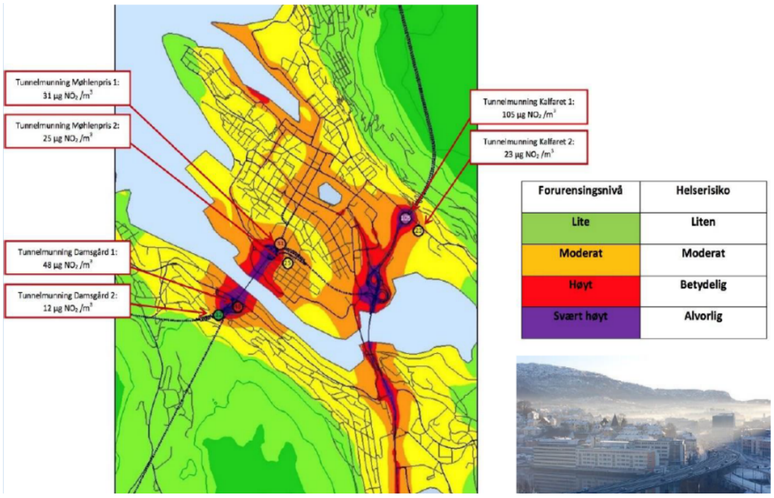

Figure 2: Grade five teaching project on air pollution (Norwegian Public Roads Administration and Bergen Municipality, 2017, p. 15-16)

A cluster-analysis approach was used to identify connections between responses. This approach makes it possible to identify whether there are groups of PTs, i.e. clusters, that respond in similar ways to sets of questions and whether these clusters differ significantly from one another. To do this, each PT can be considered as a point in a 51-dimensional coordinate system, where each axis corresponds to a claim in the questionnaire. For instance, the first axis corresponds to the claim “I can follow and understand Ben’s explanation”. If the PT answered, “Completely agree”, then “Completely agree” is the value of the coordinate of that point. Thus, 96 PTs yield 96 points (some perhaps overlapping) in that coordinate system. If some of the PTs commonly give the same answers, the corresponding points would lie close to one another in the coordinate system and yield a cluster.

Kroh (2006) argued that a respondent replying "Don't know" very often has an opinion on the topic, but has either not understood the question or has not taken the time to think about it. Since clustering algorithms will interpret many "Don't know" responses as if the respondents largely agree on a claim, these responses were replaced with random variables where the probability is equally high for each Likert category. This randomising process was repeated 460 times, with the analysis of differences between how the clusters had responded was repeated for each process. Differences between clusters were interpreted as significant only if they were significant for all 460 versions of the data, ensuring that the probability of false positives is less than 0.01. A check of outliers was made, using the Mahalanobis distance ($p<0.001$). No outliers were found.

To select which clustering algorithm and which number of clusters should be used, we applied the Ratkowsky–Lance criterion for choosing: the clustering method; the similarity measure to use; and the number of clusters to consider (Ratkowsky & Lance, 1978). This criterion has the advantage of also making sense for non-parametric data. Generally, introducing criteria for clustering ensures that a cluster is only split into two smaller clusters if the elements inside the smaller clusters are "sufficiently more" similar to each other compared to the original, larger cluster. However, there must also be a rule for limiting the number of clusters to avoid 96 different clusters with one PT in each.

The similarity measures we tested when applying the Ratkowsky–Lance criterion were the Pearson χ^2 similarity measure and its normalisation, the ϕ^2 similarity measure. For practical purposes, both methods measure the similarity between claims rather than PTs. Two claims are considered similar if the PTs commonly respond equally to both. However, in the χ^2 measure, the similarity index grows with the number of respondents, so that within a cluster, the size of the cluster would affect χ^2. To fix this, the ϕ^2 measure is given by dividing χ^2 by the number of respondents (Anderberg, 1973).

The Ratkowsky–Lance criterion defines the "best clustering" as that which maximises a certain index. This index can be described as: For each claim, the average ϕ^2 similarity between what the PTs have responded and which cluster they belong to, is computed. Then the average for all claims is divided by the square root of the number of clusters. This implies that having too many clusters results in a lower index. Maximising the index yielded that applying the "Within-groups linkage" clustering algorithm together with the ϕ^2 distance metric and two clusters was, by definition of the criterion, the most ideal approach. In within-groups linkage, the clusters are formed so that ϕ^2 within each cluster is as small as possible.

When looking for significant differences between the clusters, we applied the Mann–Whitney U test (Mann & Whitney, 1947; Wilcoxon, 1945).

Findings

Initial findings for the group of PTs as a whole is discussed in the other paper presented at this conference and so the results are only briefly summarised here. The PTs generally chose a positive response to all the claims. They agreed on the need for argumentation in the mathematics classroom, they were largely positive to the idea of modelling, although they placed themselves more in the middle on the claims about the air pollution project. On the other hand, they still showed a tendency towards believing that the project would be good for the students' understanding of modelling.

Overview of responses and characteristics of the two clusters

In this section, we present the two clusters and discuss the statistically significant differences between them. By using the "Within-groups linkage" clustering algorithm together with the Ratkowsky–Lance criterion, two clusters were identified. Cluster 1 (C1) consists of 32 PTs, with the remaining 64 PTs forming cluster 2 (C2). Table 1 provides the responses given by the PTs in each cluster on the claims about Ben's argumentation, as well as some claims about modelling and the air pollution project.

Claims about argumentation (Figure 1)		Disagree	Neither/nor	Agree
1 a) I can follow and understand Ben's explanation.	C1	0.0	3.1	96.9
	C2	25.1	15.6	56.3
1 b) Ben's explanation is incomplete.	C1	81.3	6.3	6.3
	C2	20.4	39.1	34.4
1 c) Ben's explanation is mathematically correct.	C1	0.0	6.3	87.5
	C2	4.7	42.2	42.2
Claims on modelling in general				
3 b) Students in primary school are too young for modelling to contribute to increasing their critical judgment ability towards how mathematics is used in society.	C1	90.6	3.1	6.2
	C2	45.3	39.1	6.3
3 d) Modelling improves students' attitudes towards mathematics.	C1	0.0	15.6	75.0
	C2	4.7	42.2	45.4
Claims on the air pollution project (Figure 2)				
7 a) This project would be too difficult for students who do not speak good Norwegian because they would not be able to justify their answers.	C1	50.0	21.9	18.8
	C2	14.1	34.4	43.8
7 c) I would not have used this project in my teaching because the students would find it too extensive.	C1	53.5	21.9	18.8
	C2	17.2	31.3	34.4
7 f) Students in 5th grade must know or first learn about the concepts of average and spread measures to understand what the project is about.	C1	28.2	31.3	28.1
	C2	6.3	20.3	56.3
7 h) The project would take too much time from teaching and would not allow us to get through everything we are supposed to in the book.	C1	59.4	18.8	9.2
	C2	34.4	28.1	20.4

Table 1: Overview of how the clusters have responded

In Table 1, "disagree" is the combined number of responses from "completely disagree" and "moderately disagree", and similarly for "agree". "Neither/nor" stands for the response "neither agree nor disagree". The percentages for "don't know" are excluded, for the reasons discussed previously, and, as a result, the percentages do not total to 100.

Table 1 shows that cluster 1 is generally more positive towards Ben's explanation, modelling in general and the air pollution project. Cluster 2's responses are more spread across the different Likert categories and there is a higher percentage of "Neither agree nor disagree". Cluster 1 is more positive towards Ben's explanation even though his explanation is written in a language the PTs are unfamiliar with. Not one PT from cluster 1 disagreed with the claim that "I can follow and understand Ben's explanation". Almost everyone agreed with the claim and among these, 62.5% replied "Completely agree". Only 12.5% of the PTs in cluster 2 chose the same response.

Extreme responses, in which PTs "completely agreed or disagreed" were a particular characteristic of cluster 1. On claim 7 h), nearly a third of the PTs in this cluster "completely disagreed" that the air pollution project in topic 7 would take too much time from teaching. The project provided no information about what the time-frame would be for such a project. Only 3.1% of cluster 2 completely disagreed. For claim 7 c), there are 21.9% of the PTs in cluster 1 who completely disagreed that the air pollution project would be too extensive, while in cluster 2, the percentage is 1.6%.

Significant differences between the clusters

We found a significant difference between the clusters on 16 of the 51 claims in the survey. Seven of these were on the claims about student argumentation, two were on modelling in general, and five were on the air pollution project. In addition, one was on argumentation in general and one was on a modelling project concerning composting food waste. Because of limitations on space, we have focused on nine claims where the differences between clusters were the most significant (see Table 1).

As noted in the previous section, the clusters largely differ because the PTs in cluster 1 were more prone to responding on the "completely" end of the scale, while the cluster 2 responses were more evenly spread and with a bigger percentage (compared to cluster 1) who responded, "neither agree nor disagree". This applies in particular for the claims about argumentation. In cluster 1, at least 50% of the PTs answered on the "completely" end of the scale on all except three of the twelve claims. For cluster 2, the same result is only found on two of the claims.

Although the clusters differ, they do largely agree with each other (with only three exceptions among the ones with significant differences), in the sense that if there are more "moderately/completely agree" responses compared to the "moderately/completely disagree" responses in one cluster, then the same is true for the other one. The three claims, where the clusters differ in their responses, are 1 b) on argumentation, and 7 a) and 7 c) on the air pollution project, see Table 1. The difference between the clusters is the largest on 1 b), where 31.3% of cluster 1 completely disagree on the claim that Ben's explanation is incomplete. For cluster 2, the most selected response in 1 b) and 7 c) is "Neither agree nor disagree", and 31% or more selected this for all three claims.

For the remaining 35 claims, the two clusters do not differ significantly. Nevertheless, there is still a trend that indicates that the PTs in cluster 1 were more likely to select the "completely" agree/disagree options, while the PTs in cluster 2 more often selected "Neither agree nor disagree" or "Don't know".

Discussion and concluding remarks

Previous research on clusters of PTs has focused on the PTs' views on their mathematics skills and beliefs regarding multilingual classrooms (see for example, McLeman & Fernandes, 2012). Given the emphases in the new Norwegian curriculum (Utdanningsdirektoratet, 2019), our survey, instead, mapped PTs' views on argumentation, modelling, digital tools, critical mathematics education as well as multilingual classrooms. To our knowledge, there has not been research previously on PTs' views across this spectrum of aspects of mathematics education to determine if their responses showed distinct clusters within a cohort. Our results indicate significant differences between the two clusters on argumentation and modelling, but less so in regard to multilingual aspects, critical mathematics education and ICT. Although McLeman et al. (2012) had found differences between clusters to do with multilingual issues, our questions were different and the experiences of the PTs likely to be considerably different to those of PTs in the USA, which could explain why we did not find similar relationships in our data. McLeman et al. identified in their research that demographic factors contributed to the clustering of the PTs, when studying views on English learners in mathematics classrooms. We did not request this data due to privacy issues, as this information could make the PTs' potentially identifiable. Nevertheless, it may be that further research is needed to determine if similar demographic factors could also affect our clustering.

The results are both interesting in themselves and raise issues that require more thought, particularly to do with how to structure our teacher education to reflect the specific needs of the two clusters. It would seem that, on the whole, the PTs in cluster 1 were much more positive about what students could do with non-textbook tasks, but lacked experience in how to utilise the information they do know to make judgments about what to do in their mathematics teaching. Thus, the positivity displayed by cluster 1 might need to be supplemented with critical skills.

On the other hand, the PTs in cluster 2 appear to display some critical skills which can be developed in relationship to engaging in deep-level discussions, surrounding the complexity of bringing together different aspects of mathematics teaching. These skills could be used, for example, to develop the PTs' critical reflection also in the areas that they are not yet showing awareness of, such as what, when and how digital tools should be incorporated into their future mathematics teaching. As well, the PTs in cluster 2 seemed to show that they are willing to be supportive of students' responses and are ambitious for their learning. These are also good skills to have, which perhaps could also be built on in regard to reflecting on the needs of second-language learners in their mathematics classrooms. Further research on the impact of adjustments that are made to our teaching will be published later, as our larger project develops.

Acknowledgement

This research is part of the LATACME project which is funded by the Research Council of Norway.

References

Anderberg, M. R. (1973). *Cluster analysis for applications*. New York: Academic Press.

Hannula, M. S., Kaasila, R., Laine, A., & Pehkonen, E. (2005). Structure and typical profiles of elementary teacher students' view of mathematics. In H. L. Chick & J. L. Vincent (Eds.), *Proceeding of the 29th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 89–96). Melbourne: PME.

Hansen-Thomas, H., & Cavagnetto, A. (2010). What do mainstream middle school teachers think about their English language learners? A tri-state case study. *Bilingual Research Journal.*, *33*, 249–266.

Kroh, M. (2006). Taking 'don't knows' as valid responses: a multiple complete random imputation of missing data. *Quality and Quantity*, *40*(2), 225–244.

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, *4*(2), 73–79.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, *18*(1), 50–60.

McLeman, L., & Fernandes, A. (2012). Unpacking preservice teachers' beliefs: A look at the language and culture in the context of the mathematics education of English learners. *Journal of Mathematics Education*, *5*(1), 121–135.

McLeman, L., Fernandes, A., & McNulty, M. (2012). Regarding the mathematics education of English learners: Clustering the conceptions of preservice teachers. *Journal of Urban Mathematics Education*, *5*(2), 112–132.

Meaney, T., Lange, T., Hansen, R., Herheim, R., Rangnes, T.E. & Rasmussen, N.H.W. (in press). Surveying mathematics preservice teachers. *Preceedings of the Ninth Nordic Conference on Mathematics Education (Norma20).*

National Council for Teacher Education (2016). *National guidelines for the primary and lower secondary teacher education programme for years 1–7.* Retrieved from https://www.uhr.no/temasider/nasjonale-retningslinjer-for-lererutdanningene/

O'Muircheartaigh, C., Krosnick, J., & Helic, A. (2000). *Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data*. Retrieved from https://EconPapers.repec.org/RePEc:har:wpaper:0103

Norwegian Public Roads Administration and Bergen Municipality (2017). *Luftkvalitet i Bergen 2016 (Air quality in Bergen, 2016)*. Bergen: Author. Retrieved from www.bergen.kommune.no/bk/multimedia/archive/00316/Arsrapport_luftkval_316266a.pdf.

Ponte, J. P. d., & Chapman, O. (2008). Preservice mathematics teachers' knowledge and development. In L. English (Ed.), *Handbook of International Research in Mathematics Education* (pp. 223–261). New York: Routledge.

Ratkowsky, D., & Lance, G. (1978). Criterion for determining the number of groups in a classification. *Biometrics*, *44*(1), 23–34.

Utdanningsdirektoratet (The Norwegian Directorate for Education and Training) (2019). *Læreplan i matematikk 1.–10. trinn (Curriculum for mathematics, grades 1–10)*. Oslo: Author.

Wang, R., & Krosnick, J. A. (2020). Middle alternatives and measurement validity: a recommendation for survey researchers. *International Journal of Social Research Methodology*, *23*(2), 169–184. doi: 10.1080/13645579.2019.1645384

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83.