A Clustering Method for Multiple-Answer Questions on Pre-service Primary Teachers’ Views of Mathematics

Published Version:
A Clustering Method for Multiple-Answer Questions on Pre-service Primary Teachers’ Views of Mathematics / Maffia, Andrea; Rossi Tisbeni, Simone; Ferretti, Federica; Lemmo, Alice. - ELETTRONICO. - (2020), pp. 129-138. [10.1007/978-3-030-50526-4_13]

Availability:
This version is available at: https://hdl.handle.net/11585/849373 since: 2022-01-31

Published:
DOI: http://doi.org/10.1007/978-3-030-50526-4_13

A CLUSTERING METHOD FOR MULTIPLE-ANSWER QUESTIONS ON PRE-SERVICE PRIMARY TEACHERS' VIEWS OF MATHEMATICS

Andrea Maffia°, Simone Rossi Tisbeni°°, Federica Ferretti*, Alice Lemmo**,

°University of Pavia, °°University of Bologna, *Free University of Bozen, **University of L'Aquila,

In the last years, research has paid strong attention to pre-service primary teachers' views of mathematics. Interviews and questionnaires to pre-service teachers during their academic studies are the mainly used tools for collecting data. Qualitative and quantitative approaches may give different insights. In this paper, after a review of the different methods used in the literature to face the topic of pre-service primary teachers' views of mathematics, we propose a new method. A clustering technique is applied to data collected with multiple-answer questions about pre-service primary teachers' views of mathematical ability. Obtained clusters are interpreted and compared.

METHODS FOR STUDYING PRE-SERVICE TEACHERS' VIEWS OF MATHEMATICS

It is recognized internationally how the teachers' beliefs are decisive in the process of teaching-learning mathematics; since several years, research in mathematics education follows this direction of investigation (e.g. Schoenfeld, 1989). As pointed out by Pajares (1992), teachers' beliefs are in most cases already developed during pre-service university courses and for this reason many studies in literature focus on the affective sphere of prospective mathematics teachers (e.g. Hannula, Liljedahl, Kaasila, & Roesken, 2007; Brady & Bowd, 2005). In line with this perspective, our research focuses on pre-service primary teachers' views of mathematics.

Due to the epistemological and cultural nature of the research on affect, most studies employ qualitative methods mainly using ethnographical and linguistic methods (e.g Ebbelind, 2015) or using open questionnaires (see Hart, 2002). As highlighted by Kloosterman and Stages (1992), closed instruments (like Likert-scale) can suggest the researcher's ideas, thus influencing the respondent in a social desirability perspective. To overcome this, some authors use open questionnaires in which respondents are free to express their emotions, beliefs and memories by using their own words and they are not forced to align their opinion on a ready-made list chosen by the researcher (Di Martino & Sabena, 2011; Di Martino, Coppola, Mollo, Pacelli, & Sabena, 2013). Again, even if the respondents are free to use their own words, we could argue that the formulation of a question may still influence the answers, especially in the case of inverse formulated items. The data gathered by the questionnaire are often analysed through an inductive content analysis (Patton, 2002); in some cases, descriptive statistics are also used (Di Martino & Sabena, 2011**)**.

Other authors use closed-ended tools that may be adapted or inspired by already validated scales. For example, Zollman and Mason (1992) created the Standard Beliefs Instrument (SBI). Such instrument consists of a battery of Likert-scale questions aimed at determining the consistency of teachers' beliefs with NCTM Standards (1989). Obviously, that one is not the only study on the topic of teachers' believes using Likert-scale (another well-known example is given by the Problem-Solving Project, Schoenfeld, 1989).

In this paper, we present a clustering method for analysing multiple-answer questions. Cluster analysis was also used by Hannula, Kaasila, Laine and Pehoken (2005) in the case of a different format of questions: they clustered the respondents of two questionnaires about pre-service teachers' views of mathematics (see also Roesken, Hannula, & Pehkonen, 2011). Such analysis produced three main types of belief profiles: positive, neutral, and negative view, each one then divided into subclasses (Hannula et al., 2005). Similarly, we use cluster analysis to obtain different profiles of respondents, but we cluster the answers to only one multiple-answer question.

CLUSTERING OF MULTIPLE ANSWERS

Multiple-answer questions

As we have seen in the previous section, research about pre-service teachers' views of mathematics often favours qualitative methods over quantitative ones and open questions rather than multiple-choice questions. Open questions give to the respondents the possibility to express freely their own opinion, using personal wording and so giving thicker descriptions (Geertz, 1973). However, the analysis becomes more time-consuming limiting the amount of processable data. In contrast, when administering a multiple-choice question, the designer of the questionnaire forces the respondent to select just one among a limited number of given answers (so rising many possible critiques, see Cohen, Manion, & Morrison, 2002), but making it easier to perform quantitative analysis. We get just a restricted image of human behaviour when "social scientists concentrate on the repetitive, predictable and invariant aspects of the person; on 'visible externalities' to the exclusion of the subjective world; and on the parts of the person in their endeavours to understand the whole" (Cohen, Manion, & Morrison, 2002, p.18).

Several solutions could be found "in between" the fully open-question and the multiple-choice. One possibility is the use of rating scales (like Likert-scales). This kind of scales allow to know more than just a "yes/no" agreement on a statement giving some shadows to an otherwise black/white picture of the results. As also seen before, Likert-scales are largely used in psychological and educational research and many researchers developed this kinds of method (*ibidem*). It is much rarer, especially in research on affect in mathematics education, to see the use of *multiple-answer questions*. A multiple-answer question is similar to a multiple-choice one, with the only difference that the respondent can select more than one answer option. Research has paid attention on identifying the optimal number of options in a multiple-choice question (e.g. Baghaei & Amrahi, 2011), but less is known about the optimal number of choices for a multiple-answer question. When the respondents are forced to select between a yes/no answer (or even an agree/disagree choice) they tend to express more positive opinions rather when they have a “select-all-that-apply” option (Dillman, Smyth, Christian, & Stern, 2003); hence, multiple-answer questions allow the respondents to express better their opinion. Rasinski, Mingay and Bradburn (1994) show that it is possible that respondents do not really mark “all-that-apply” when answering to this kind of question. Furthermore, there is evidence that the first given answers receive more selections than the other ones (Dillman, Smyth, Christian, & Stern, 2003; Rasinski, Mingay & Bradburn, 1994). We conjecture that, by giving a fixed amount of answers to select, we induce the respondents to read all the possible answers. Thus they will opt just for those that mostly represent their own opinion. It could be argued that, in this way, we restrict the possibility of expression of the respondent, getting closer to the case of multiple-choice format. This is true, but a following interpretative process may help in getting still a not-superficial picture of the respondents’ opinions.

Usually, multiple-answer questions are analysed computing the relative frequency of each answer. In this way, the obtained results are not different from those obtainable by a multiple-choice question. Much more insight is given when some of the answers are “linked”. This could be realized by a correspondence analysis (Greenacre, 2017) of the answers or by clustering techniques. In this paper we take the latter option, as it is presented in the following sections.

Clustering

The answers are encoded in a binary vector of size n, with n the number of possible options. The i-th position of the vector has a value of 1 if the i-th option was selected, 0 otherwise. For example, if a respondent in a 6-options multiple answer question selected options A, D and F, the resulting vector would be $(1, 0, 0, 1, 0, 1)$. The vectors are concatenated in matrix $\mathrm{M}_{p\times n}$ with p rows, one for each participant.

A similarity matrix is computed as $\mathrm{A}_{p\times p} = \mathrm{M} \cdot \mathrm{M}^T$, where the element $a_{i,j}$ corresponds to the number of options selected by participant i also selected by participant j. The resulting matrix is a symmetric matrix, with 3 in every element of its diagonal (each participant has 3 answers in common with herself).

A new matrix is obtained by calculating the pair-wise Euclidean distance between each row of the matrix A. We then use classical *multidimensional scaling* to map the points into a 3-dimensional space (e.g. Borg & Groenen, 1997).

The resulting points are then clustered using an agglomerative hierarchical clustering algorithm, meaning that at each steps the cluster with the *shortest distance* are merged (fig. 2). The distance D between two clusters X, Y can be computed in multiple ways (e.g. Backhaus, Erichson, Plinke & Weiber, 2016). Single linkage may function to determine the outliers in the data, and then performing the Ward algorithm classifies the remaining elements. While this algorithm can result in a valid clustering, in this work its performance was reduced, due to the absence of isolated data points (a more detailed discussion is reported in the final section).

The complete linkage rule was then chosen since it tends to find compact clusters of similar diameters, avoiding chaining phenomena (Everitt, Landau, Leese & Sthal, 2011). This distance is defined as:

$$D(X, Y) = \max_{\boldsymbol{x} \in X, \boldsymbol{y} \in Y} d(\boldsymbol{x}, \boldsymbol{y})$$

where d is the Euclidean distance. In this way, cluster are joined where the distance between the furthest members of the clusters is the lowest.

The chosen number of clusters is the one that minimizes the absolute maximum deviation from the median.

AN EXAMPLE: PRE-SERVICE TEACHERS' VIEW OF MATHEMATICAL ABILITY

In this paper we are presenting a method for analysing multiple-answer questions by means of clustering. Rather than presenting a huge amount of results, we prefer to show clearly how the method is applied and what kind of results is possible to get. For this reason, we will analyse the results of just one question that was inserted in a questionnaire administered to students attending the course of "Mathematics Education" within the last year of the master-degree in Primary Education in the University of Bologna, Italy. The sample consists of all the students enrolled in the last year of the master-degree (N=207). They all answered to a questionnaire containing some mathematical tasks, some Likert-scale items about their emotions toward mathematics and mathematics teaching, and two multiple-answer items; one about their views of mathematics ability and one on their beliefs about teaching mathematics. These questions were always administered in the same order. In the following, we refer to one of these questions: *Select among the following options the three features that, according to you, are fundamental for having success in mathematic*s. Twelve options were available (Fig. 1).

We selected the options according to the model of mathematical giftedness described by Pitta-Pantazi, Christou, Kontoyianni and Kattou (2011). Following these authors, mathematical ability is the result of *learned mathematical abilities* (like verbal, spatial, quantitative abilities and other) and *creativity* (defined as a combination of fluency, flexibility and originality). Both learned abilities and creativity are supported by *natural abilities* (comprehending working memory, control and speed of processing). We decided to integrate this model adding the dimension of *affect*. Indeed, research has shown strong evidence of the influence of emotions and motivation on mathematical performances (e.g. Zan, Brown, Evans & Hannula, 2006).

Figure 1 shows how each of the available options for our question is related to one of these dimensions. The capital letter before each option corresponds to the order in which the twelve options were given in the questionnaire: A is the first option and N is the last one (the letters J and K are not commonly used in the Italian alphabet). As explained above, we asked the respondents to select only a fixed number of options. The choice of the number 'three' makes possible to select only the options belonging to one dimension. Pre-service teachers responding to the questionnaire do not know the framework and so their interpretation of the terms could differ from the formal definition. For this reason, we changed some terms into more colloquial synonyms or phrases. However, the problem is not so easily solved and so an interpretative process of the given answers is necessary. In the following, we will discuss those cases in which a word (in the original language of administration of the questionnaire, Italian) may be interpreted in more than one way.

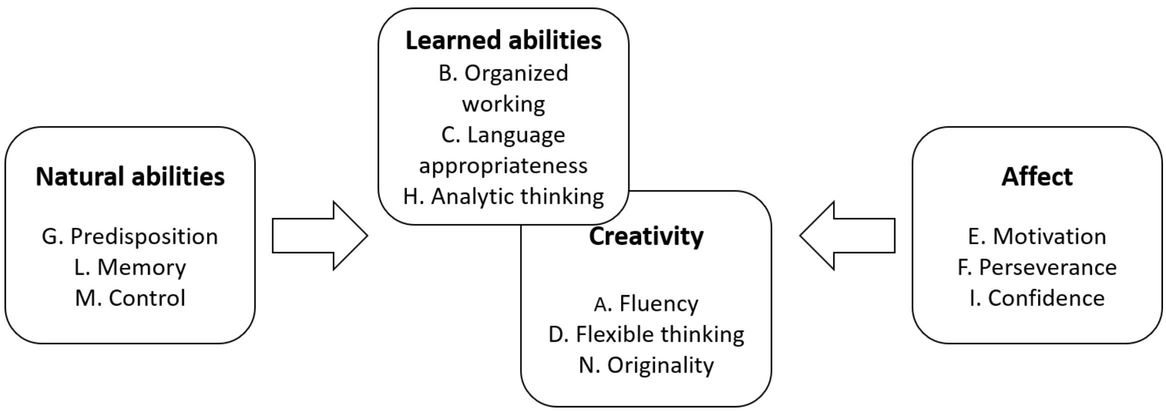

Figure 1: Classification of options within to the adopted theoretical framework.

The questionnaires were administered by means of paper and pencil, dividing the participants in groups of a maximum size of 33. Three researchers (among them the second and third authors of this contribution) administered the questionnaire to each group. A time limit was given, and the question analyzed was in the central part of the questionnaire. We analyze only the complete answers, so we are not considering blank answers and those cases in which less than three options were selected. In this way, the amount of analysed answers is p=178.

Figure 2 shows a dendrogram representing the process of clustering described in the previous section. Starting from the bottom, elements with a distance smaller than the value reported on the ordinate axis are merged. We decided to take 6 clusters even if the distance from the median value is minimized also in the cases of 14 or 15 clusters. This is because a high number of clusters creates difficulties in the interpretative process. Figure 3 shows how the choice of the options is distributed in the six clusters; the last column reports the number of respondents in each of the obtained clusters

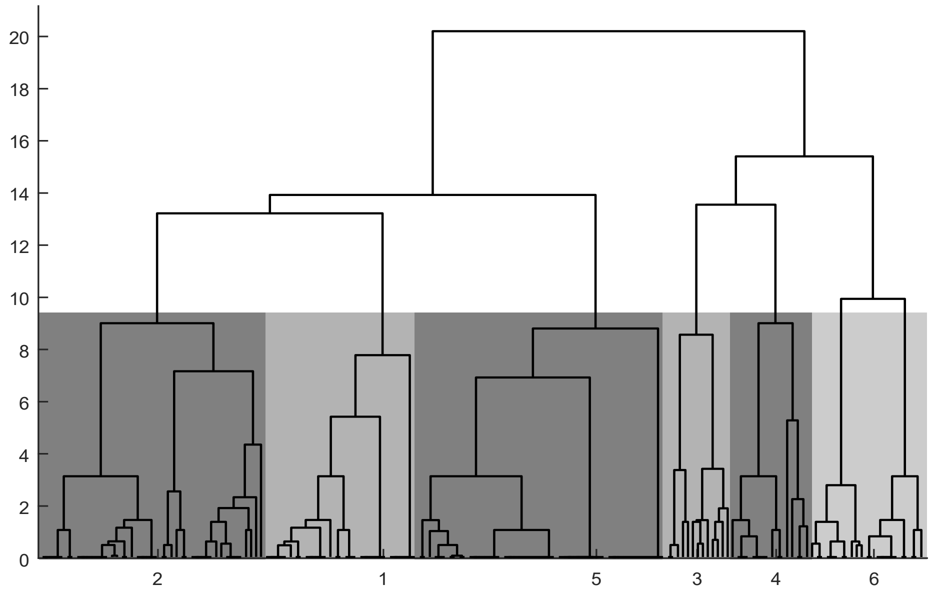

Figure 2: Dendrogram representing the clustering process. The obtained clusters are highlighted in grey. The y-axis reports the distance at which clustering occurs.

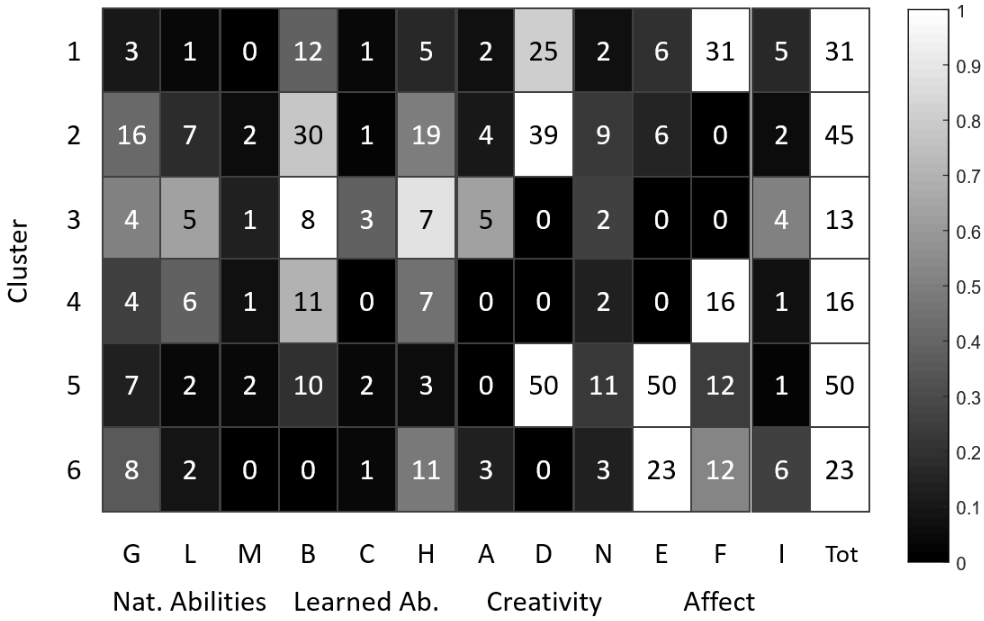

Figure 3: Frequency of choices of options (A-N) in the six clusters. Colours represent the relative frequency compared to the maximum frequency (tot) within the cluster.

The *first cluster* comprehends pre-service teachers who selected mainly the answers D and F (flexible thinking and perseverance). 12 over the 30 members of this cluster

(40%) selected also option B (organized working). This cluster contains those pre-service teachers giving relevance to the relation between creativity and affect, taking also in consideration learned abilities. They do not give importance to natural abilities.

According to the members of the *second cluster*, mathematical ability is related mainly to learned abilities and creativity: they give relevance to the relation between organized working (option B) and flexible thinking (option D). There are also high percentages dedicated to analytic thinking (42%) and predisposition (36%). Apparently, members of this cluster see flexible thinking as the main component of mathematical ability, supported by a certain behaviour and, maybe, a predisposition. They give less relevance to the dimension of affect.

The *third cluster* is less populated; most of its members (62%) select the option B (organized working) often paired with analytic thinking (54%), predisposition (31%), memory (38%) or fluency (38%). Pre-service teachers in this cluster give strong importance to innate abilities and learned behaviour. Indeed, they often select one between the options G, L or A. In our model, we related option A (fluency) to the dimension of creativity, but the Italian translation could also be interpreted by the participant as "quickness", so relating to the speed of processing that is a natural ability in the model of Pitta-Pantazi and colleagues (2011).

The dimension of creativity is ignored also by the *fourth cluster*. They give relevance to keeping perseverance (100%) on an organized work (69%), eventually with the support of natural abilities. The other aspects of affect are not relevant for this group.

The *fifth cluster* is clearly characterized by the choice of option D and E, respectively flexible thinking (100%) and motivation (100%). This is not the only group giving relevance the role of motivation, indeed also the *sixth cluster* does (100%). The difference between these two last clusters is that the former put affect mainly in relation to the creative dimension, while the latter give more relevance to the other dimensions.

Table 1 summarizes the relevance given to the different dimensions of mathematical ability by the members of each cluster according to the interpretation we gave in this section. The positive sign indicates a high relevance, while the negative sign represents a low relevance. No sign is indicated when we did not see a strong tendency in the selection of options by respondents.

DISCUSSION

Large scale studies about pre-service teachers' views of mathematics provide a huge amount of data that is unrealistically analysable only by means of qualitative methods. Quantitative analysis has its limits, as many researchers noticed (Cohen, Manion, & Morrison, 2002). The availability of several different instruments allowing the respondents to express their opinion as freely as possible appears relevant. In this paper we have studied the use of multiple-answer questions within a questionnaire about pre-service primary students' views of mathematics. We used a clustering technique to analyse the answers to a multiple-answer question about respondents' opinions on mathematical ability. In our literature review we did not find other works using these methods in our field of research.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Natural abilities	–		+		–	
Learned abilities		+	+	+	+	
Creativity	+	+		–		
Affect	+	–		+	+	+

Table 1: Numerosity of respondents in each cluster and relevance given to the four dimensions of mathematical ability.

The used hierarchical clustering algorithm does not explicitly consider the existence of links (common options) between two answers. To verify that this information is not lost, for each answer we measured the average number of common options with other answers in the same cluster. The selected cluster for an answer, is considered valid if the maximum average of common option is within the selected cluster while, for the other clusters, the average is lower. In our results, using a complete linkage, 161 answers over 178 belonged to clusters with highest average number of shared options, and 9 of the 17 remaining where paired to the cluster with the second highest average. We tested also single linkage (obtaining a ratio of 63/178) and Ward linkage (153/178), so concluding that the chosen algorithm was the best option in this case. As shown in the previous section, some limits of closed questions still remain, in particular the possible different interpretation of the terms used in the answer options (for instance in the analysis of the third cluster). However, we can observe that the choice of allowing the selection of only three options resulted in avoiding one of the usual limits of multiple-answer question: differently than observed in previous research (Dillman, Smyth, Christian, & Stern, 2003; Rasinski, Mingay & Bradburn, 1994), the respondents did not show a preference for the first options in the list, indeed option A and C were among the less selected. This observation suggests that while answering to this kind of questions, respondents pay more attention to the selection of those options that mainly represent their opinion. As it was already done with multiple-choice questions (Baghaei & Amrahi, 2011), further research is needed to identify the effects of changing the number of options to list and to select.

The example of analysis in the previous section shows that it is possible to cluster respondents according to their answers to a question about their views of mathematical ability. Such clustering revealed different groups that we were able to describe according to an interpretative process based on a reference theory (Pitta-Pantazi et al., 2011). We can see that the cluster of pre-service teachers giving more importance to natural abilities is the smaller while much attention is paid to learned abilities and affect. Concerning the affect dimension, we have two clusters giving more relevance to perseverance and two clusters paying more attention to motivation.

Clustering appears as a suitable method in analysing multiple-answer questions giving more insights than just frequency calculation. Furthermore, assigning each respondent to a cluster, we can deepen the analysis comparing the results of a multiple-answer question with those of different type of questions. For instance, in our questionnaire we can compare the membership to a specific cluster with the emotions that the respondent associate with mathematics and/or mathematics teaching. It will be possible to look for interdependence between emotions and a certain views of mathematics.

Acknowledgements

We wish to thank Prof. Ira Vannini, Dr. Andrea Ciani and Carla Provitera for their fundamental collaboration in making possible the realization of this study.

References

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2016). Clusteranalyse. In K. Backhaus, B. Erichson, W. Plinke, & R. Weiber (Eds.) *Multivariate Analysemethoden* (pp. 453-516). Berlin, DE: Springer Gabler.

Baghaei, P., & Amrahi, N. (2011). The effects of the number of options on the psychometric characteristics of multiple-choice items. *Psychological Test and Assessment Modelling, 53*(2), 192.

Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling: theory and applications*. New York, NY: Springer.

Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education*. Routledge.

Di Martino, P., Coppola, C., Mollo, M., Pacelli, T., & Sabena, C. (2013). Pre-service primary teachers' emotions: The math-redemption phenomenon. In A. M. Lindmeier, & A. Heinze (Eds.) *Proceedings of the 37th Annual Conference for the Psychology of Mathematics Education, vol. 2* (pp. 225-232). Kiel, DE: IGPME.

Di Martino, P., & Sabena, C. (2011). Elementary pre-service teachers' emotions: shadows from the past to the future. *Current state of research on mathematical beliefs XVI*, 89-105.

Ebbelind, A. (2015). We Think so, Me and My Mom–Considering External Participation inside Teacher Training. In C. Bernack-Schüler, R. Erens, T. Leuders, A. Eichler (Eds.) *Views and Beliefs in Mathematics Education* (pp. 109-120). Wiesbaden, DE: Springer Spektrum.

Everitt, B. S., Landau, S., Leese, M., & Sthal, D. (2011). *Cluster Analysis*. Hoboken, NJ: Wiley.

Geertz, C. (1973). Thick description: towards an interpretive theory of culture. In. C. Geertz (Ed.) *The Interpretation of Cultures*. New York, NY: Basic Books.

Greenacre, M. J. (2017). *Correspondence analysis in practice*. New York, NY: Chapman and Hall.

Dillman, D. A., Smyth, J. D., Christian, L. M., & Stern, M. J. (2003) Multiple answer questions in self-administered surveys: the use of check-all-that-apply and forced-choice question formats. *Paper presented at the American Statistical Association*, San Francisco, CA.

Hannula, M. S., Kaasila, R., Laine, A., & Pehkonen, E. (2005). Structure and Typical Profiles of Elementary Teacher Students' View of Mathematics. *International Group for the Psychology of Mathematics Education, 3,* 89-96.

Hannula, M., Liljedahl, P., Kaasila, R., & Roesken, B. (2007). Researching relief of mathematics anxiety among pre-service elementary school teachers. In J. H. Woo, H. C. Lew, K. S. Park, & D.Y. Seo (Eds.) *Proceedings of the 31st Annual Conference for the Psychology of Mathematics Education, vol. 1* (153-156). Seoul, KS: IGPME.

Hart, L. C. (2002). Preservice teachers' beliefs and practice after participating in an integrated content/methods course. *School science and mathematics, 102*(1), 4-14.

Kloosterman, P. & Stage, F. (1992). Measuring beliefs about mathematical problem solving. *School Science and Mathematics, 92*(3), 109-115.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

Pajares, M.F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(1), 307-332.

Patton, M. (2002). *Qualitative research and evaluation methods*. London, UK: Sage Publications.

Pitta-Pantazi, D., Christou, C., Kontoyianni, K., & Kattou, M. (2011). A model of mathematical giftedness: Integrating natural, creative, and mathematical abilities. *Canadian Journal of Science, Mathematics and Technology Education, 11*(1), 39-54.

Rasinski, K. A., Mingay, D., & Bradburn, N. M. (1994). Do respondents really "mark all that apply" on self-administered questions? *Public Opinion Quarterly, 58*(3), 400-408.

Roesken, B., Hannula, M. S., & Pehkonen, E. (2011). Dimensions of students' views of themselves as learners of mathematics. *ZDM, 43*(4), 497-506.

Schoenfeld, A. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education, 20*(4), 338-355.

Zan, R., Brown, L., Evans, J., & Hannula, M. S. (2006). Affect in mathematics education: An introduction. *Educational studies in mathematics, 63*(2), 113-121.