The role of representations in the understanding of probabilities in tertiary education

Sofia Anastasiadou
University of Western Macedonia
Florina, Greece
sofan@uom.gr

Theodore Chadjipantelis
Aristotle University of Thessaloniki
Thesaloniki, Greece
chadji@polsci.auth.gr

Summary

This study aims to contribute to the understanding of the approaches students develop and use in solving probabilistic tasks and to examine which approach is more strongly correlated to their success in such tasks. Participants were pre-service teachers from the University of Western Macedonia in Greece. Implicative statistical analysis is applied to evaluate the relation between students' approach and their ability to solve problems. Our results indicate strongly that students tend to avoid graphical representations and use the algebraic approach instead. Students who are able to coordinate multiple representations show better results in problem solving. In addition, the results suggest that flexibility in using multiple representations is a good predictor for success.

Introduction

There is an increasing recognition that statistical and probabilistic concepts are among the most important unifying ideas in mathematics. Nowadays probability and statistics are part of mathematics curricula for primary and secondary schools in many countries. The reasons for this development are related to the usefulness of statistics and probability for daily life (Chadjipadelis, 2003b), its instrumental role in other disciplines, the need for basic stochastic knowledge in many professions, and its key role in developing critical reasoning (Batanero, et al., 2004).

Understanding of probabilistic and statistical concepts does not appear to be easy, given the diversity of representations associated to this concept, and the difficulties inherent in the processes of articulating systems of representation involved in probabilistic and statistical problem solving (SPS) (Anastasiadou, 2007). Probability is difficult to teach for various reasons, including disparity between intuition and conceptual development even as regards apparently elementary concepts (Chadjipadelis & Gastaris, 1995). Since an education that only focuses on technical skills is unlikely to help teachers overcome their erroneous beliefs, it is important to find new ways to teach probability to them, while at the same time bridging their content and pedagogical knowledge (Batanero et al, 2005). There is general consensus in the mathematics education community that teachers need a deep and meaningful understanding of any mathematical content they teach (Estrada et al., 2004). Biehler (1990) suggests that teachers require meta-knowledge about probabilities and statistics, including a historical, philosophical, cultural and epistemological perspective on statistics and its relations to other domains of science.

At primary and secondary levels, probability and statistics is part of the mathematics curriculum; yet primary school teachers and mathematics teachers frequently lack specific preparation in statistics education (Anastasiadou 2007, Chadjipadelis, 2003b). According to Batanero, et al. (2005) probability is increasingly taking an important role in the school mathematics curriculum; yet most teachers have little experience with probability and share with their students a variety of probabilistic misconceptions (Chadjipadelis, 2003b).

The need for a variety of semiotic representations in the teaching and learning of probability is usually explained through reference to the cost of processing, the limited representation affordances for each domain of symbolism and the ability to transfer knowledge from one representation to another (Duval, 1987). A representation is defined as any configuration of characters, images, concrete objects, etc., that can symbolize or "represent" something else (DeWindt-King & Goldin, 2003; Goldin, 1998; Kaput, 1985). In the last decades, great attention has been devoted to the concept of representation and its role in the learning of mathematics. A basic reason for this emphasis is that representations are considered as an “integral” part of mathematics (Kaput, 1987). In certain cases, specific representations are so closely connected to a mathematical concept that it is difficult for the concept to be understood and acquired without the use of these representations. Students experience a wide range of representations from their early childhood years. A main reason for this is that most mathematics textbooks today use a variety of representations in order to enhance understanding. Greeno & Hall (1997) maintain that representations may be considered as useful tools for constructing understanding and for communicating information. They underline how important it is to engage students in activities like choose or construct representations in such forms that help them to see patterns and perform calculations, taking advantage of the fact that different forms provide different support for inference and calculation. Similarly, Kalathil & Sheril (2000) describe ways in which representations may be useful in providing information on how students think about a mathematical issue, and serve as classroom tool for the students and the teacher. In mathematics instruction representations get a crucial role in that teachers can improve conceptual learning if they use or invent effective representations (Cheng, 2000).

The use of multiple representations, such as pictures and text combined, is a main feature of mathematics education, which deals with a wide range of representations of ideas in order to enhance understanding. Generally, there is strong support in the mathematics education community that students can grasp the meaning of mathematical concepts by experiencing multiple mathematical representations (e.g., Lesh, Post, & Behr 1987; Sierpinska, 1992). Principles and standards for school mathematics (NCTM, 2000) include a standard referring exclusively to representations and emphasize their value for understanding. Learning from verbal and pictorial information has generally been considered beneficial for learning (Carney & Levin, 2002; Schnotz, 2002). For example, Ainsworth, Wood, & Bibby (1997) suggest that the use of multiple representations may help students develop different ideas and processes, constrain meanings, and promote deeper understanding. Furthermore, a second representation may be provided to support the interpretation of a more complicated or less familiar representation (Gagatsis & Michaelidou, 2002). According to Ainsworth et al. (1997), by combining representations students are no longer limited by the strengths and weaknesses of one particular representation.

In the field of statistics instruction, representations play an important role as an aid for supporting reflection and as a means for communicating statistical ideas. In this study, we revisited the role of representations in an effort to understand the nature and structure of representations in developing statistical concepts. We investigated the ability to use multiple representations and translate from one representation to another.

Representations have been classified into two interrelated classes: external and internal (Goldin, 1998). Internal representations refer to mental images corresponding to internal formulations that we construct of reality. External representations concern the external symbolic organizations representing externally a certain mathematical reality. In this study the term “representation” is interpreted as “external” tool used for representing statistical ideas such as tables and graphs (Confrey & Smith, 1991). By a translation process, we denominate the psychological processes involved moving from one mode of representation to another (Janvier, 1987). Several researchers in the last two decades addressed the critical problem of translation between and within representations, and emphasized the importance of moving among multiple representations and connecting them (Karaolis, Neofytides, Charalambous, & Gagatsis, 2006; Gagatsis & Elia, 2004, 2005; Gagatsis, Elia, & Mougi, 2002; Hitt, 1998; Yerushalmy, 1997). Duval (2002) claimed that the conversion of a mathematical concept from one representation to another is a presupposition for successful problem solving. According to Elia & Gagatsis (2006), the role of representations is a central issue in the teaching of mathematics. The most important aspect of this issue refers to the diversity of representations for the same concept, the connection between them and the conversion from one mode of representation to others. Gagatsis & Shiakalli (2004) and Ainsworth (2006) suggest that different representations of the same concept complement each other and contribute to a more global and deeper understanding of it.

Understanding a mathematical concept presupposes the ability to recognise this concept when it is presented by a series of qualitatively different representation systems, the ability to flexibly handle this concept in the specific representation systems, and finally, the ability to translate the concept from one system to another (Dufur-Janvier, Bednarz & Belanger, 1987; Lesh, Post, & Behr, 1987). In statistical education, the interest focuses both on the various types of representation and on the translations between them.

The focus of this study is to evaluate the approach pre-service teachers use in order to solve simple probability tasks. It is of interest to know whether these teachers are flexible in using algebraic, graphical and verbal representations in probabilistic problems. Most of the teachers in our study use an algebraic approach in order to solve simple probabilistic tasks. This study analyzes the role of different modes of representation on understanding of some basic probabilistic concepts. Teachers’ performance is investigated in two aspects of probabilistic understanding: the flexibility in using multiple representations and the ability to solve the problems posed.

Participants – Tasks – Data analysis

The sample consisted of 243 pre-service students from the University of Western Macedonia. For the analysis of the collected data we use the statistical implicative analysis. The tasks consist of 12 exercises related to the concept and definition of probability, to Venn diagrams, and to probability problem solving.

The test was developed by the authors. The 12 items covered the concepts of population, sample, and mean, frequency, frequency tables, bar and pie charts and their application to solving everyday problems. Students’ responses to the tasks comprise the variables of the study which are codified by an uppercase letter followed by the number of the item and two letters. The uppercase letters denote the concept involved in the task; V stands for items concerning Venn diagrams, P for probability problems, R for items with a concept definition (e.g. event). The lower case letters following denote the type of representation: r = representation, t = table, g = graphic, v = verbal, s = symbol; the first letter stands for the initial, the last for the final representation. Correct answers are encoded as 1, wrong or no answers as 0, and partial solutions are encoded as 0.5.

For example, the first and second tasks are the following:

Task 1. Given two events A and B of a chance experiment and with the help of set theory we have the following event $A' \cap B'$. Present this event by a Venn diagram (encoded as V1sg).

Task 2. Given two events A and B of a chance experiment; with the help of set theory we define the event $(A' \cap B) \cup (A \cap B')$. Express this event verbally (encoded as V2sv).

For the analysis of the collected data the method of statistical similarity (Lerman, 1981) is conducted using the software C.H.I.C. (Classification Hiérarchique, Implicative et Cohésitive, Bodin, Coutourier, & Gras, 2000). This method determines similarity relations between variables. In particular, the similarity analysis is a classification method, which aims to identify in a set V of variables, thicker and thicker partitions of V, arranged in an ascending order. These partitions, when fitting together, are represented in a hierarchically constructed tree diagram using a statistical criterion for the similarity among the given variables. Similarity is defined by the cross-comparison between a group V of the variables and a group E of the individuals (or objects). This kind of analysis allows for studying and interpreting clusters of variables in terms of a typology determined by decreasing similarity; clusters of variables, which are established at particular levels of the diagram can be opposed to others, at the same level. It should be noted that statistical similarities do not necessarily imply logical or cognitive similarities.

For this study’s needs, a similarity and an implicative diagram are produced. The construction of the similarity diagram is based on the following process: At the first (highest) similarity level, those two of the variables that are most similar to each other with respect to the similarity indices of the method are joined together to form a group. In the next step, this group may be linked to one variable in a lower similarity level or two other variables that are already combined together and establish another group at a lower level, etc. This grouping process goes on until the similarity or the cohesion between the variables or the groups of variables gets very weak. Based on this process, it is evident that similarity is stronger between groups the shorter the vertical lines are between them in the diagram The thick (red) horizontal lines represent significant relations of similarity. The implicative diagram, which is derived by the application of Gras's statistical implicative method, contains implicative relations that indicate whether success in a specific task "implies" success in another task related to the former one.

Results

The descriptive results are shown in Table 1

Table 1: Students' success rates of pre-service students in the tasks in all types of conversions.

Tasks	**Type of translation**	**Success rate (%)**	**Tasks**	**Type of translation**	**Success rate (%)**
V1sg	Symbolic – Graphic	52.4	P7va	Verbal – Algebraic	42.5
V2sv	Symbolic – Verbal	50.7	P8vg	Verbal – Graphic	37.9
V3gs	Graphic – Symbolic	38.9	P9vs	Verbal – Symbolic	31.1
V4gv	Graphic – Verbal	36.9	P10vv	Verbal – Verbal	28.9
V5vg	Verbal – Graphic	49.4	R11vv	Verbal – Verbal	28.8
V6vs	Verbal – Symbolic	51.4	R12vs	Verbal – Symbolic	32.5

Results of similarity analysis

The similarity diagram (Figure 1) allows the arrangement of the responses to the tasks into groups according to their homogeneity. The similarity lines with (thick) red colour are significant at a significance level of 99%.

Two clusters (Cluster *A* and *B*) of variables are identified in the similarity diagram of pre-service students' responses as shown in Figure 1. Cluster *A* involves three pairs of variables V1sg-V2sv, V3gs-V4gv, V5vg-V6vs and comprises representations of events with the aid of Venn diagrams. Cluster *B* involves also three pairs of variables, namely R11vv- R12vs, P7va-P8vg, P9vs-P10vv and involves variables relating to probability problem solving. This grouping suggests that students dealt similarly with the conversions involving probability problems of the same cluster.

The structure of the diagram reveals a cognitive difficulty that arises from the need to accomplish flexible and competent conversion back and forth between different types of probabilistic representations. Thus, this particular structure of the diagram indicates a compartmentalization of the tasks of the tests. On the one hand there are the tasks associated to Venn diagrams and on the other those tasks, which involve probability problems. Students approached the two groups of tasks in a completely distinct way, Therefore, possible instructive activities would focus on the identification of the two different groups. The strongest similarity (almost 1) occurs between variables (V3gs-V4gv) (Figure 1). Furthermore the similarity (V1sg-V2sv, V3gs-V4gv) is also important (0.923).

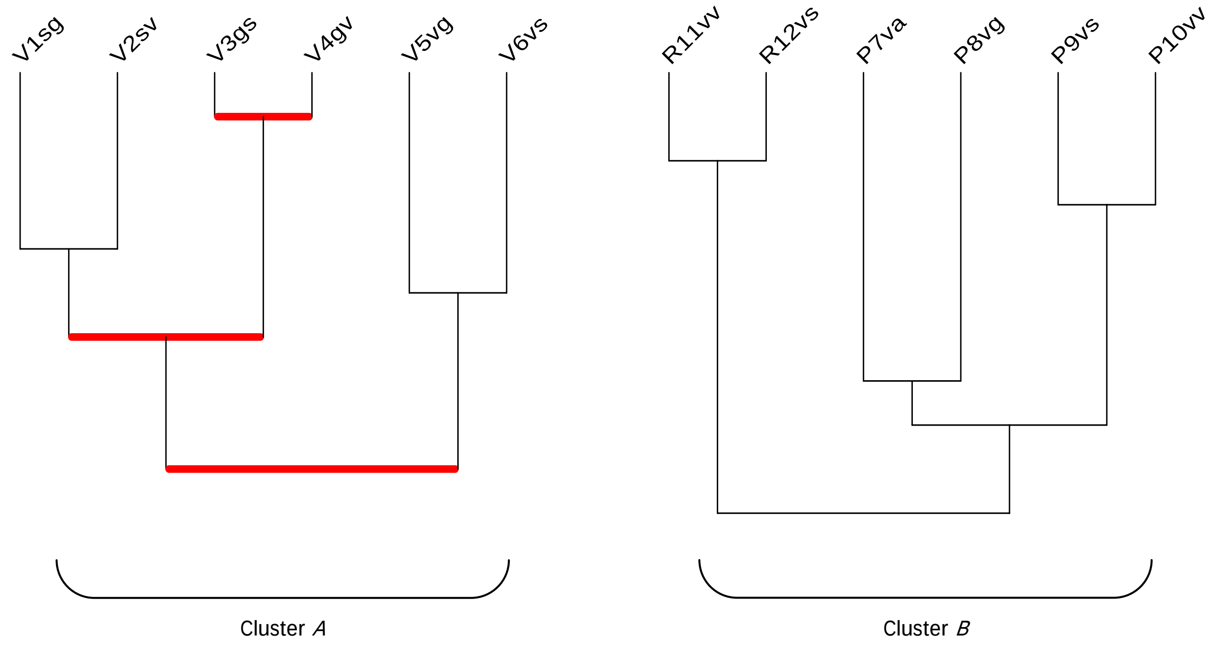

Figure 1: Similarity Diagram – thick connections are significant at 99%.

Comments on the implicative diagram

The implicative diagram shows the implicative relations between the variables (Figure 2). According to this diagram, not all the tasks of the test are connected by implicative relations. The implications represented by straight, dashed, or dotted lines represent relations significant at levels of 99%, 95%, or 90% respectively.

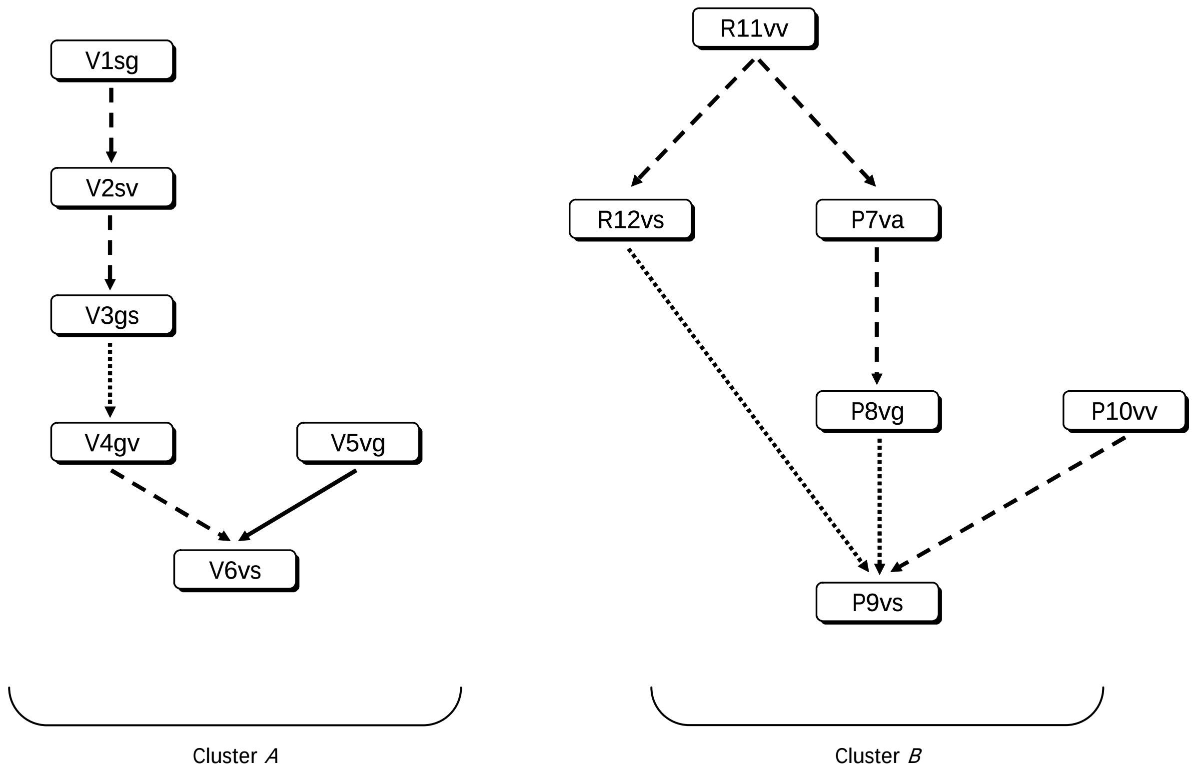

Figure 2: Implicative Diagram: arrows 99%, broken arrows 95%, dashed arrows 90% significance.

Two distinct chains of variables maybe seen from the implicative diagram in Figure 2:

(V1sg $\rightarrow$ V2sv $\rightarrow$ V3gs $\rightarrow$ V4gv, V4gv, V5vg $\rightarrow$ V6vs) and

(R11vv $\rightarrow$ R12vs $\rightarrow$ P7va $\rightarrow$ P8vg $\rightarrow$ P9vs, R12vs, P8vg, P10vv $\rightarrow$ P9vs)

The first one in cluster *A* concerns representation of events with the aid of Venn diagrams. The second one in cluster *B* involves variables related to probability problem solving. This grouping suggests that students dealt similarly with the conversions involving probability problems. The implicative diagram of the pre-service students' responses is exactly in accordance to the right similarity diagram as shown in Figure 1.

Conclusions

Representations enable students to interpret situations and to comprehend relevant relations embedded in probabilistic problems. Thus, we consider representations to be extremely important with respect to cognitive processes in developing probabilistic concepts. The main contribution of the present study is the identification of pre-service teachers' abilities to handle various representations and to translate among representations related to the same probabilistic relationship. Our findings provide a strong case for the role of different modes of representation on pre-service teachers' performance to tasks. At the same time they enable a developmental interpretation of students' difficulties in relation to representations of Venn diagrams. Lack of connections among different modes of representations in the similarity diagram indicates the difficulty in handling two or more representations in probabilistic tasks. This incompetence is the main feature of the phenomenon of compartmentalization in representations, which is detected in this study. This inconsistent behaviour may be seen as an indication of students' conception that different representations of the same concept are completely distinct like if they were autonomous mathematical objects and not just different ways of expressing the meaning of the same notion.

Probability instruction needs to engage students in activities including translations between different modes of representation. As a result, students will be able to overcome the compartmentalization difficulties and develop their flexibility in understanding and using a concept within various contexts or modes of representation and in moving from one mode of representation to another.

It seems that there is a need for further investigation into the subject with the inclusion of a more extended qualitative and quantitative analysis. In the future, it is interesting to compare strategies and modes of representations students use in order to solve the problems. Besides, longitudinal investigations might reveal new insights how the flexibility in using the multiple representations grows.

References

Ainsworth, S. E., Wood, D. J., & Bibby, P. A.: 1997, *Evaluating Principles for Multi-Representational Learning Environments.* Paper presented at the 7th EARLI Conference. Athens.

Ainsworth, S.: 2006, DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction* 16, 183–198.

Anastasiadou, S.: 2007, The phenomenon of compartmentalization in students' processing of statistical concepts, in *Intern. Symposium Elementary Mathematics Teaching*, Prague, 61–68.

Biehler, R.: 1990, Changing conceptions of statistics: A problem area for teacher education, in Hawkins, A. (Ed.), *Training Teachers to Teach Statistics. Proc. Intern. Statistical Institute Round Table Conference*, Intern. Statistical Institute, Voorburg, 20–38.

Batanero C., Codino D. J., & Roa R.: 2004, Training teachers to teach probabilities. *Journal of Statistics Education* 12(1). http://www.amstat.org/publications/jse/v12n1/batanero.htm1.

Batanero, C., Biehler, R., Maxara, C., Engel, J., & Vogel, M.: 2005, Using Simulation to Bridge Teachers´ Content and Pedagogical Knowledge in Probability, in *15th ICMI Study Conference: The Professional Education and Development of Teachers of Mathematics.* http://stwww.weizmann.ac.il/G-math/ICMI/strand1.html.

Bodin, A., Coutourier, R., & Gras, R. : 2000, *CHIC : Classification Hierarchique Implicative et Cohésitive-Version sous Windows – CHIC 1.2,* Association pour le Recherche en Didactique des Mathematiques, Rennes.

Carney, R. N. & Levin, J. R.: 2002, Pictorial illustrations still improve students' learning from text. *Educational Psychology Review* 14(1), 101–120.

Chadjipadelis, T. & Gastaris, P.: 1995, Difficulties of understanding and misconceptions in Probabilities and in Statistics. *Eucledes C* 43, 35–68. Online: http://www.stat.auckland.ac.nz/~iase/publications/17/8C3_CHAD.pdf

Chadjipadelis T.: 2003a, Children's walk in probability and statistics from childhood to youth, in *Proc. 19th Panhellenic Conf. on Mathematical Education*, Veria, 27–40.

Chadjipadelis T.: 2003b, Data, methods and results, in *Proc. 16th Panhellenic Conf. on Statistics*, Kavala, 37–44.

Cheng, P.C.H.: 2000, Unlocking conceptual learning in mathematics and science with effective representational systems. *Computers and Education* 33, 109–130.

Christou, C, Gagatsis, A. & Zachariades, T.: 2001, The difficulty level of representations in mathematical relationships. *Scientia Paedagogica Experimentalis* XXXVIII(1), 25–38.

Confrey, J. & Smith, E.: 1991, A framework for functions: Prototypes, multiple representations and transformations. In Underhill, R. G. (Ed.), in *Proc. 13th annual meeting North American Chapter of the Intern. Group Psychology of Mathematics Education*, Virginia Polytechnic Institute and State University, Blacksburg, 57–63.

Duval, R.: 1987, *The role of interpretation in learning mathematics. Dimension*, 2, 56–74.

DeWindt-King, A. M. & Goldin, G. A.: 2003, Children's visual imagery: Aspects of cognitive representation in solving problems with fractions. *Mediterranean Journal for Research in Mathematics Education* 2(1), 1–42.

Dufur-Janvier, B., Bednarz, N., & Belanger, M.: 1987, Pedagogical Considerations Concerning the Problem of representation. In Janvier, C. (Ed.), *Problem of Representation in the Teaching and Learning of Mathematics* (57–-63), Lawrence Erlbaum, Hillsdale, NJ.

Duval, R.: 2002, The cognitive analysis of problems of comprehension in the learning of mathematics. *Mediterranean Journal for Research in Mathematics Education* 1(2), 1–16.

Elia, I. & Gagatsis, A. (2006). The effects of different modes of representation on problem solving: Two experimental programs, in Novotná, J., e. a. (Eds)., Proc. 30th Conf. Intern. Group Psychology of Mathematics Education, Vol. 3 (pp. 3–25), Prague.

Estrada, A., Batanero, C.,Fortuny, J. M., & Díaz., C.: 2004, A structural study of future teachers' attitude towards statistics, in Bosch, M. (Ed.), *Proc. 4th Congress European Society for Research in Mathematics Education* (pp. 518–527), ERME, Saint Feliu de Guixols, http://ermeweb.free.fr/CERME4/.

Gal, I.: 2002, Adult's statistical literacy. Meanings, components, responsibilities, *International Statistical Review* 70(1), 1–25.

Gagatsis, A. & Elia, I.: 2004, The effects of different modes of representations on mathematical problem solving, in Johnsen Hoines, M. & Berit Fuglestad, A. (Eds.), *Proc. 28th Conf. Intern. Group for the Psychology of Mathematics Education*, Vol. 2 (pp. 447–454), Bergen University College, Bergen.

Gagatsis, A. & Elia, I.: 2005, A review of some recent studies on the role of representations in mathematics education in Cyprus and Greece, in Bosch, M. (Ed.), *Proc. 4th Congress European Society for Research in Mathematics Education* (pp. 102–111). ERME, Saint Feliu de Guixols, http://ermeweb.free.fr/CERME4/.

Gagatsis, A., Elia, I., & Mougi, A.: 2002, Multiple Representations in Developing Mathematical Relationships. *Scientia Paedagogica Experimentalis* XXXIX (1), 9–24.

Gagatsis, A. & Michaelidou, E. 2002. Le relazioni tra le diverse rappresentazioni del concetto di funzione e la compresione del concetto stesso. Une ricerca riferita a gli studenti della scula secondaria superiore. *La mathematica e la sua didattica* 4, 411–429.

Gagatsis, A. & Shiakalli, M.: 2004, Ability to translate from one representation of the concept of function to another and mathematical problem solving. *Educational Psychology* 24(5), 645-657.

Goldin, G. A.: 1998, Representational systems, learning, and problem solving in mathematics. *Journal of Mathematical Behavior* 17(2), 137–165.

Greeno, J. G. & Hall, R. P.: 1997, Practicing representation: Learning with and about representational forms, *Phi Delta Kappa* 78, 361–367.

Hitt, F. 1998. Difficulties in the articulation of different representations linked to the concept of function. *Journal of Mathematical Behavior* 17(1), 123–134.

Janvier, C.: 1987, Representation and understanding: The notion of functions as an example, in Janvier, C. (ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 67–72), Lawrence Erlbaum Associates, Hillsdale, N.J.

Kalathil, R. R. & Sheril, M. G.: 2000, Role of students' representations in the mathematics classroom, in Fishman, B. & O'Connor-Divelbiss, S. (Eds.), *4th Intern. Conf. of the Learning of Sciences* (pp. 27–28), Erlbaum, Mahwah, NJ.

Kaput, J.: 1987, Representation and problem solving: methodological issues related to modelling, in Silver, E. A. (Ed.), *Teaching and Learning Mathematical Problem Solving: Multiple Research Perspectives* (pp. 381–398). Erlbaum, Hillsdale, NJ.

Lesh, R., Post, T. & Behr, M.: 1987, 'Representations and translation among representations in mathematics learning and problem solving', in Janvier, C. (ed.), *Problems of Representation in the Teaching and Learning of Mathematics*, Lawrence Erlbaum, Hillsdale, NJ, 33–40.

Lerman, I. C.: 1981, *Classification et analyse ordinale des données*, Dunod, Paris.

NCTM.: 2000, *Principles and standards for school mathematics*, NCTM, Reston, Va.

Schnotz, W.: 2002. towards an integrated view of learning from text and visual displays. *Educational Psychology Review* 14(1), 101–120.

Sierpinska, A.: 1992, On understanding the notion of function, in Dubinsky, E. & Harel, G. (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (pp. 25–28), The Mathematical Association of America.

Yerushalmy, M.: 1997, Designing representations: Reasoning about functions of two variables. *Journal for Research in Mathematics Education* 27(4), 431–466.