

Psychological Effects of Peer Tutoring in Mathematics in Primary Education

Lidon Moliner^a, Francisco Alegre^b, Gil Lorenzo^c ^{a,b,c}Universitat Jaume I, Castellon, Spain, Email: mmoliner@uji.es, falegre@uji.es, valentin@uji.es

In this work, a literature review of peer tutoring studies in Elementary Education Mathematics has been carried out. This review has focused only on those studies in which psychological variables are analysed. The main objectives of this work have been to compile these studies, to analyse the influence of peer tutoring at a psychological level on the students and to identify those variables that influence the results obtained (type of tutoring, ages of the participants, duration of the sessions...). The final review is made up of 22 studies. Among the main conclusions obtained from its analysis, it should be noted that 16 of these 22 studies reported important improvements in the psychological variables analysed (73% effectiveness), which is a sign of the potential of this methodology. The psychological variable on which the greatest improvements were detected was that of self-concept (89% of studies with evident improvements). Although the conclusions obtained can give a good sample of the potential of this methodology on a psychological level, a series of limitations are provided in this work to be taken into account. These limitations must be taken into account when analysing these conclusions from a critical perspective.

Key words: *peer tutoring, mathematics, primary education, psychological variables*

Introduction

The main objective of this work is to analyse the influence of peer tutoring on psychological variables of elementary school students, specifically in the subject of Mathematics. For decades, most meta-analyses and systematic reviews in the field of peer tutoring have focused exclusively on the variable of academic performance. This fact is motivated, in turn, by the fact that most research in this field has academic achievement as its main or only variable of study. In this way, the influence of this methodology on different psychological variables (anxiety, self-concept, attitude...) remains in the background, and there is practically no review of this style in the literature. Thus, although the effect of this methodology on academic performance

in Mathematics has been sufficiently proven, its effect on the psychological, emotional or social level has not yet been widely studied. Therefore, in this End of Grade Project, a review of the existing literature on peer tutoring in Elementary Education in the subject of Mathematics will be carried out. First, a search will be made by consulting the available papers and studies in English and Spanish related to this field. Once this search is done, these studies will be recorded by means of an orderly list of them. Only those studies that meet the following conditions will form part of this list: that the student body to which it refers belongs to Primary Education, that the interaction between the students is direct and in pairs, that the subject of study is Mathematics and that at least one variable analysed is psychological. The following variables will be included in this relationship: types of tutoring (fixed or reciprocal and between students of the same or different ages), condition of the participating students (with or without disabilities), type of study design (quantitative, qualitative or mixed), variables analysed (self-concept, attitude, anxiety...), number of participants, frequency of tutoring sessions among peers and total duration of the sessions in minutes. Likewise, the main results obtained in each study will be attached in a brief way. Finally, this relationship will be analysed with the aim of analysing the effect of this methodology at the psychological level, trying to identify those variables that can play a key role in this effect.

Objectives of the study

The following objectives are proposed for this work:

- Objective 1: To search studies in the peer tutoring field in mathematics primary education in which psychological variables are analysed, detailing diverse variables like the type of tutoring (fixed or reciprocal), ages of the participants (same age, different ages) or condition of the participants.
- Objective 2: To carry out a simple quantitative study based on the relationship obtained in objective 1 with the different variables analysed.
- Objective 3: To analyse the main results obtained in each study and draw conclusions from them regarding the influence of peer tutoring on different psychological variables in the subject of Mathematics.

Theoretical framework

The interaction of students with their classmates is one of the main elements that have a direct impact on learning. Through the interaction with the students it generates a series of cognitive conflicts that favour their learning. In the traditional classroom, such learning takes place through unidirectional communication through interaction with the teacher who plays the role of mediator. However, it has been proven that in various circumstances the students themselves can play the role of mediators. Thus, the vast majority of us have learned from colleagues at

school, high school, university or in our respective jobs. Some authors such as Duran (2006) state that, on certain occasions, students can become better mediators than adults themselves. This statement is justified by the fact that students are recent learners of the curricular content, so they are both knowledgeable and sensitive to the conceptual sections in which their peers may require greater help. In addition, students tend to use more direct language and have the added advantage of sharing linguistic and cultural references. Also, the closeness between classmates can also play a key role in this regard. Therefore, for a large number of students it is easier and less embarrassing to ask a question directly to a classmate than to ask it aloud to a teacher, facilitating communication in this context.

This research has as its central axis, peer tutoring. Duran and Vidal (2004) define peer tutoring as a cooperative learning method based on the creation of couples. These pairs are characterised by having an asymmetric relationship derived from the respective role of tutor or mentor. There is a common and shared role and objective which is the acquisition of a certain curricular competence. This objective is achieved through a relationship framework structured and planned by the teacher. In addition to the above, Topping (2005) indicates that over the years peer tutoring has evolved, so that it is increasingly difficult to give a precise definition of it. Nevertheless, an approach to such a definition could be the following: people from similar social groups who do not have professional teaching skills and who help each other to learn at the same time that they learn by themselves through teaching. Other authors such as Gordon (2009) are less precise and define it simply as an instructional method in which one student tutors another in a subject in which the tutor is an expert and the tutor is a novice. Within these definitions, the consideration given by Topping above that participants should not possess professional teaching skills is especially relevant. In this sense, authors such as Boud et al. (1999) or Liu and Carless (2006) emphasise that, among the participants, there should not be people who violate this concept of "equal". In this sense, these authors exclude not only teachers, but also people with university degrees, parents of students or technicians specialised in any field of study. Although there is no single consensus regarding the term "equal", it can be considered that, in general, the authors in the field consider as "equal" the students of the different educational levels, since they share the fact of being academically trained in a learning environment without having concluded their university studies.

Although peer tutoring is widely used in many Anglo-Saxon countries (Peer Tutoring, Peer Learning or Peer Instruction), in Spain it is still not used globally in primary, secondary, high school or university classrooms. Furthermore, although many times in our country this type of experience does take place in an informal way, these practices are rarely collected in a formal way and documented in the academic literature. In fact, this is the case in most Spanish-speaking countries. A simple search in Google Scholar with the words "peer tutoring" and its translation into English shows the differences at the level of formal publication in this sense with 2,600 results in the first case and 61,900 in the second. Thus, it should be borne in mind that a majority of publications in this field will be documented in English.

Materials and methods

The materials used to carry out the study were the different publications formally collected in the literature related to the field of study of this work. First, it was considered essential to include those articles from journals indexed in the Journal Citations Reports and Scopus databases, since in the field of education at the research level are the references worldwide (Bergstrom, 2007; Falagas et al., 2008). As indicated by Godlee et al. (1998) the filters at the editorial and reviewer level of the articles that are published in this type of journals ensure their experimental rigour.

In addition to the above-mentioned articles, as indicated by Cooper, Hedges, and Valentine (2009), there is a large amount of what is known as "gray literature" (also called "fugitive" literature) that has not been published in journals with the indexation indicated in the previous paragraph. This type of literature includes doctoral theses or papers presented at conferences, among others. According to these authors, it is advisable to include this type of research in the different reviews and meta-analyses that are to be carried out. The reason for this is that their inclusion involves a larger and, therefore, more representative sample. In this way, not only are articles published in indexed journals part of this review, but also doctoral theses, reports and communications presented in congresses and all types of publications with sufficient experimental rigour are included. In order to ensure that the publications had sufficient experimental rigour, the criteria proposed by Erickson and Gutierrez (2002) and Smith et al. (2007) were used, so that every publication included in this review meets the criteria proposed by these authors.

Procedure

In order to achieve the objectives previously set out in this paper, a search for peer tutoring articles in Mathematics was carried out. The databases consulted were the following Apa Psycnet (American Psychological Association), ERIC (Educational Resources Information Center) and Google Scholar (the academic version for google searches). The following eight terms were used as descriptors for these searches: 'peer tutoring', 'mentoring', 'peer instruction', 'peer learning', 'math', 'problem solving', 'arithmetic' and 'geometry'. These terms were defined taking into account that they were the most used by the authors in this field after making the theoretical framework previously developed in this work. The search was carried out in Spanish and translating these same terms into English. The terms were combined as follows: each search included two terms, one of the first four and one of the last four previously indicated. Thus, a total of 16 possible search combinations were produced. Since the search was also carried out in English and three databases were consulted, a total of 96 different searches were carried out combining terms, languages and databases. In no case was a restriction made by date of publication. It was decided not to include terms referring to the age of the students such as 'Primary Education', 'School' or 'College', since this could have implied an exclusion of valid articles to be included in this review. In the case of Google Scholar, it was



decided to restrict the search to those articles in which the terms appeared together in the title of the article, since a search with the terms ‘peer tutoring’ and ‘mathematics’ returned about 100,000 different results in the form of publications, making it unfeasible to read and rigorously examine all these publications.

The suitability of the use of these terms and the rigour of these searches were contrasted by means of diverse procedures. In the first place, the bibliographic references found in the previously mentioned databases were contrasted with those published in magazines of high impact index where researches related to the field under study had been previously published. Likewise, it was corroborated that in the previous meta-analyses available in the literature regarding peer tutoring there were no publications different from those obtained with the searches that could be included in this work. The meta-analyses consulted were the following: Alegre-Ansuategui et al. 2018, Bowman-Perrott et al. 2013; Cohen, Kulik and Kulik 1982; Forness et al. 1997; Gersten et al. 2009; Kroesbergen and Van Luit 2003; Leung, 2015; Leung, Marsh and Craven 2005; Rohrbeck et al. 2003; Zeneli, Thurston, and Roseth 2016. In addition, as in the case of the meta-analyses, the literature reviews conducted in peer tutoring were also consulted for the same purpose. Thus, the literature reviews consulted were: Alegre et al. 2018; Britz, 1989; Delquadri et al. 1986; Robinson et al. 2005; Stenhoff and Lignugaris/Kraft, 2007; Topping, 1996.

The searches in the different databases allowed us to obtain a total of 1124 bibliographic references, including repeated articles. Most of the references found with the words ‘mentoring’ or ‘problem solving’ were not related to math peer tutoring, so they were directly discarded. In addition to the above, many articles had no experimental basis or simply did not conduct any kind of analysis, such as manuals, reflections, or reviews. Thus, after excluding the previously mentioned articles, only 154 articles were left that could be included in this bibliographic review. These articles still passed a final filter with the selection criteria. As indicated by Cooper, Hedges and Valentine (2009), the selection criteria should be posed naturally from the questions or objectives of the research work. Thus, taking into account the indications of these authors, four research criteria were raised and used as a final filter prior to the inclusion of the publications in this work. These criteria are indicated below:

First criterion: to include only those publications in which the interaction between tutor and tutor was direct from person to person. In recent years, many publications refer to virtual tutoring systems in which students are guided by different computer software (include references), so that the interaction is not between two people, but between a person and a machine. Since this work is based on peer learning, such publications were excluded. However, it should be noted that those investigations in which the interaction was between two students, even though an electronic device was used as a communication channel, were included. 24 articles were excluded because of this criterion.

Second criterion: include only those publications in which the tutors were not parents, legal guardians, post-graduate staff or other adult persons. There are multiple investigations in which students from different educational levels, such as high school or university, help primary education students through individualised tutoring. However, in this work it is considered that the help coming from parents or post-university staff should not be part of it, since it is understood that the tutorial action comes from someone who violates the concept of "peer" in the context of "peer tutoring". 22 publications were excluded by this criterion.

Third criterion: to include only those publications in which the tutors were Primary Education students, that is, who were enrolled within the educational system in some course between 1st and 6th grades of Primary Education. 56 studies were excluded by this criterion.

Fourth criterion: To include only studies in which some psychological variable was analysed among the participants, since these are the objective of analysis in this work. 30 studies were eliminated by this criterion. Thus, after using these four criteria, 22 studies were part of the present work of bibliographic review.

Analysis of the information

After applying the selection criteria mentioned above and restricting the review to the 22 studies mentioned above, the following information was extracted from each one of them: type of tutoring used (fixed or reciprocal), condition of the participating students (with or without disability), type of design used (quantitative, qualitative or mixed), psychological variables analysed (self-concept, attitude, anxiety...), number of participants in the study, frequency of peer tutoring sessions carried out and total duration of each one of them in minutes. In addition, the most relevant results were extracted with respect to the psychological variables analysed in each article. The most important results and conclusions included in each study were examined. Once all this information was obtained, a simple quantitative descriptive analysis was made (Murgiondo et al., 2005), that is, the percentage of studies that carried out fixed tutoring versus those that carried out reciprocal tutoring was looked at, and so on with the rest of the indicated variables. As indicated by Wayne (1993), in education sciences, it is important to carry out this type of analysis in order to show in a simplified way the information collected in any research. Andrews (2005) also places simple quantitative descriptive analysis as the first fundamental point in any review, since from it one can begin to inform, work, discuss, and reflect on different educational policies and practices. Secondly, based on the results extracted in each study, an exploratory sequential strategy analysis was carried out (Pérez, 2011). The main procedure of this type of analysis is to employ qualitative information trying to quantify it in order to explore a phenomenon. The strategy followed in this type of analysis is the integration of data in the interpretation and the main objective is to explain and interpret relationships between different studies (Creswell, 2003). By using this type of analysis, the aim is to respond to the effectiveness of the 22 studies on psychological variables of the students by analysing the effect of the different variables involved. Thus, it can be considered that the analysis of the

information obtained in this study has been carried out using, on the one hand, quantitative and, on the other hand, qualitative information, which can be considered a model of mixed analysis which is recommended by different experts in the field of education (Onwuegbuzie and Leech, 2006; Tashakkori and Creswell, 2007). In order to facilitate this analysis, data from each of the studies were entered into the statistical program SPSS version 24, obtaining the different percentages indicated in the following sections.

Results and discussion

In this section, first of all, the most relevant results obtained through the quantitative descriptive analysis will be provided, making the relevant discussion and comments in each case. In second place, the most relevant conclusions referring to the analysis of the main results extracted from the studies will be contributed, also providing the corresponding discussions and comments.

At a descriptive level, in the context of Mathematics in Primary Education, one of the most evident conclusions is that in a majority of studies fixed peer tutoring was implemented (14 studies of 22, 63%) as opposed to reciprocal peer tutoring (7 studies of 22, 32%), while in one of the studies both were analysed (5%). The dominance of fixed peer tutoring remains a fact in this field of study. This is confirmed by authors such as Alegre et al. (2018) or Dennis (2013) in their works. These authors indicate that this fact is due, in many cases, to the inclination of teachers and researchers to use this type of tutoring because they feel more comfortable with it. From a logical point of view, it is understandable that they prefer not to invert the roles and that it is always the most competent or best performing students who explain the contents (tutors), while the less competent receive and feed on these explanations (tutorials). However, as will be discussed later in this section, reciprocal tutoring can have much more positive effects at an overall level than fixed tutoring. On the other hand, with regard to the ages of the participants, similarly to what happened between fixed and reciprocal tutoring, a majority of experiences are observed among students of the same age or same-age (17, 77%) compared to those studies in which students of different ages or cross-age participate (5, 23%). This fact, for authors such as Duran (2003) or Alegre et al. (2018) is justified from an organisational point of view, since it is very simple to apply tutoring between students of the same age in a single classroom, while in the other case it already requires considerable teacher coordination and organisation of time and space. In any case, it is concluded that it is important to increase the number of researches with reciprocal peer tutoring, as well as the number of experiences between students with different ages given the low number of researches in this field.

On the other hand, it was observed that practically half of the studies (10, 45%) included students with some kind of disability compared to the rest in which students with no kind of disability participated (12, 55%). As indicated by Topping (2005), in its origins, peer tutoring was mostly implemented in an unfavorable context with disruptive students, with cognitive problems or with some kind of disability. Fuchs et al. (2000) also refer to the "curative" or "reparative" character (remedial in English) of this methodology and how it is a resource

strongly linked to complex educational contexts in which a didactic alternative is sought. Thus, it can be concluded that it is a methodology frequently used in the subject of Mathematics in Primary Education with students who have some kind of disability.

Regarding research designs, the predominance of purely quantitative designs is quite evident (14, 63%) compared to those that are mixed (5, 23%) and those that are purely qualitative (4, 14%). As indicated by Onwuegbuzie and Leech (2007) the low number of qualitative or mixed studies as opposed to exclusively quantitative ones is one of the greatest threats currently facing educational research. Thus, what has happened in this literature review is not an exception in the field and, in line with what Onwuegbuzie and Johnson (2006) contribute, it is concluded that it is necessary to provide research in this field with more mixed methods that allow for the contrast between quantitative and qualitative information.

With regard to the psychological variables analysed, it is concluded that there are three variables that are the most frequently studied in this field and that there is an important balance in terms of studies in each of them. Thus, the percentages are similar in studies that analysed the students' attitude towards the subject of Mathematics (10 studies, 45%), the students' self-concept or self-esteem (9, 40%) and anxiety towards the subject of Mathematics (8.36%). The fact that among the three variables there are more than 22 studies is due to the fact that in some studies more than one of these variables is analysed simultaneously. Some other variables also analysed in an isolated manner were disruptive behaviour, peer helpers, or empathy.

Regarding the analysis of the variables number of participants, frequency of sessions and duration in minutes of each one of them, the following results were obtained. The average number of participants per study was 139.8 with a standard deviation of 77.5. This fact is largely due to the fact that three of the studies had more than 300 participants and one of them had more than 1000. Therefore, it is concluded that there is a great difference in the number of participants per study given the high standard deviation with respect to the average of the sample. On the other hand, the frequency of sessions per week was 3.7 with a standard deviation of 0.8. In this case it is concluded that the frequency of tutoring sessions per week is high and that this type of research in this field is usually intense in terms of implementation. Finally, the duration in minutes of each one of them was 27.7 with a standard deviation of 5.3. The truth is that a majority of studies were around 25 minutes or half an hour of implementation of peer tutoring. This would indicate that in this field of study, approximately half of a teaching session is dedicated to the implementation of this methodology.

The following are the most important conclusions from the analysis of the different results of the studies 16 of the 22 studies documenting significant or evident improvements (73%) in said variables, while 6 of the 22 studies did not report evident improvements (28%). On a global level, the use of peer tutoring in Elementary School Mathematics to improve psychological variables can be considered adequate and, based on the results, it is highly probable that it will be beneficial to the students.

Studies with reciprocal peer tutoring versus fixed peer tutoring showed the greatest improvements in the different psychological variables. The 7 studies with reciprocal peer tutoring (100%) provided results that implied a substantial or significant improvement in the variables to be analysed. However, this was not the case with fixed peer tutoring, with 9 of the 14 studies (64%) showing evident improvements and the other 5 results inconclusive or without important improvements. Thus, it can be concluded that reciprocal peer tutoring is more effective than reciprocal tutoring in producing psychological improvements in elementary school students in the subject of mathematics. As indicated by authors such as Duran and Miquel (2003), the superiority of fixed versus reciprocal peer tutoring has not yet been demonstrated and, as is the case in this bibliographical review, different experiences show that this is not necessarily the case (Alegre et al. 2018). In fact, several authors indicate that this improvement in reciprocal tutoring may be due to several factors. Thus, Topping (1992) indicates that in reciprocal tutoring, it is possible for the mentees to exercise a role (tutor) that they often wish to have and that fixed tutoring does not allow them. In this way, in line with what Fantuzzo et al. (1992) point out, variables such as self-concept would be strongly reinforced in these cases, since the less competent students would be able to explain mathematical contents to their classmates, while with fixed tutoring it would be more difficult for the tutorship to show these improvements.

With respect to student ages, experiences with students of the same ages showed improvements proportionately to students of different ages. Thus, 12 researches out of 17 with students of different ages (70%) showed evident improvements in the psychological variables analysed. Similarly, for the experiences among students of different ages, 4 of the 5 studies (80%) showed a considerable beneficial impact on the students. Therefore, it can be concluded that the effectiveness of tutoring between students of the same age (same-age) and different ages (cross-age) is very similar in the context under study. Such parity between types of tutoring has also been documented, although not in psychological variables, if at the level of improvements in academic performance in previous reviews and meta-analyses such as those of Leung (2015) or Bowman-Perrott et al.

In terms of the students' disability status, 9 out of 12 studies for students without disabilities (75%) and 7 out of 10 studies (70%) that included students with disabilities showed significant improvements in their outcomes. Thus, it can be concluded that this methodology has a similar effectiveness in students regardless of their disability condition. This conclusion would be reinforced by the review carried out by Britz (1989) and by the meta-analysis by Alegre et al. (2018) since, although in the variable of academic performance, similar improvements are reported in both types of students in both investigations.

With respect to the different variables analysed, the results of the improvement of the student's self-concept with respect to the variables of anxiety and attitude should be highlighted. Eight studies out of nine in which the students' self-concept was analysed showed significant improvements (89%), while in the studies which analysed the variables of anxiety and attitude,

said percentages were somewhat lower: seven out of ten for attitude (70%) and five out of eight for anxiety (62.5%). In line with the above, authors such as Fantuzzo et al. (1995) and Fantuzzo et al. (1992) refer to the potential of this methodology in the variable of self-concept or self-esteem and, as indicated above, the fact that students are able to explain content to their peers reinforces their self-esteem and their perception of their own abilities. Thus, it can be concluded that the methodology is beneficial in addressing the psychological variables analysed, but its degree of effectiveness is greater on the self-concept variable.

Regarding the design employed in each of the studies, it can be concluded that the type of design does not exert an important influence on the results obtained. Thus, 10 of the 14 studies with a quantitative design showed evident improvements (71%), 4 of 5 in the case of mixed studies (80%) and 2 of 3 in the case of qualitative studies (66%). Thus, as Zeneli et al. (2016) comment, the design used does not normally condition the results obtained in peer tutoring research. Therefore, as noted in previous paragraphs and reaffirming the provisions of Onwuegbuzie and Leech (2007), it is necessary to bring more research with a qualitative character (ideally mixed) to this field of study.

With regard to the number of participants, it was found that those researches with a greater number of participants showed less improvement than those with a lower number of participants. Thus, when the 22 studies were divided into the 11 with more participants and the 11 with fewer participants, it was observed that, in the case of the most numerous, 6 of the 11 researches contributed evident improvements (54.54%). However, in the case of the less numerous, that same percentage rose to 91%, since 10 of the 11 studies showed such improvements in the psychological variables they analysed. The influence of the sample size or, what is the same, of the number of participants in a study, has been widely discussed in educational research. Thus, Johnson and Christensen (2008) point out that this factor can be key when analysing the effect of educational methodologies. As indicated by Kraemer et al. (2003) when the sample size is very high the level of statistical significance is more easily achieved. However, in peer tutoring interventions the effect seems to be just the opposite. Thus, Alegre et al. (2018) point out that the more students participate in this type of experience, the less improvement is documented at the global level and vice versa. They support this conclusion with the provisions of Slavin and Smith (2009) who, in line with this work, also document a similar effect in terms of the number of participants. The previous authors point out different causes for which this effect may be due. The main one is the fact that the greater the number of participants, the less control is usually had over them overall. In this way, the interactions between students, the use of resources and the implementation of the methodology are seriously affected, causing a considerable decrease in their effects.

Discussion

After indicating the most relevant contributions, a series of limitations are indicated below that should be taken into account when considering the conclusions provided in this paper. The first

of these is the fact that restricting the search to only the English and Spanish languages is in itself a limitation to be taken into account. As pointed out by Kroesbergen and Van Luit (2003), there is a greater tendency in the scientific literature to publish in English and Spanish, but any restriction of languages in a literature review should be understood as a weakness when considering its validity and generalisation at a global level. Similarly, although the inclusion of publications from the so-called 'gray literature' was justified at the beginning of this work, authors such as McAuley et al. (2000) are cautious when considering their inclusion in reviews and meta-analyses, inviting at least a reflection on their suitability in this type of research. In this sense, authors such as Johnson (2001) are quite critical of what he calls "weak designs", that is, those in which there is no control group or which are based on the repetition of measures. Similarly, authors such as Valentine et al. (2010) point out that the number of studies included in this work (22) can in no way be considered as high. In fact, in some subgroups of variables the number of studies was very small (only 7 studies of reciprocal tutoring or only 3 of qualitative design). Drawing conclusions from such a small number of studies is, at least, sufficient to consider them with rigorous caution and understanding that in no case can they be generalised (Cooper and Valentine, 2008). On the other hand, it should also be borne in mind that the division of studies in this work between those that provide significant improvements in psychological variables and those that do not, although following the theoretical guidelines in the field, is undeniably questionable from a qualitative point of view (Marshall and Rossman, 2014). Thus, this classification or division made could have been different under the interpretation of another author who would have carried out this work after analysing the 22 studies that make up this bibliographic review. In this sense, performing a meta-analysis or, at least, calculating the size of the effect of the studies by quantifying the improvements of each one of the studies, would have been much more formal from an experimental point of view (Glass, 1976; Slavin, 1984). In this way, it would have been possible to graduate with greater accuracy the extent to which benefits were obtained at the psychological level in each study, avoiding a single division of the studies and with it possible subjective interpretations of the results of each study. All of the above must be taken into account when considering the validity of this work.

Conclusion

The main conclusion of this paper is that peer tutoring may be of potential interest for practitioners and researchers in mathematics educational psychology. The promising results found in this research suggest that peer tutoring could benefit mathematics education in Primary Education from a psychological perspective. Mathematics self-concept, attitude towards mathematics and mathematics anxiety seemed to be the most analysed variables in the peer tutoring field. Nevertheless, given the above-mentioned limitations of this manuscript, further research is needed in the field to determine in a more accurate way the effectiveness of this methodology on psychological variables in mathematics education. The performance of a meta-analysis is encouraged since not only its effectiveness could be quantified, but also other



variables that could potentially act as moderators (type of tutoring, length of the tutoring program) could be identified.

REFERENCES

- Alegre-Ansuategui, F. J., Moliner, L., Lorenzo, G., & Maroto, A. (2018). Peer Tutoring and Academic Achievement in Mathematics: A Meta-Analysis. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(1), 337-354.
- Alegre, F., Moliner, L., Maroto, A., & Lorenzo-Valentin, G. (2018). Peer tutoring in mathematics in primary education: a systematic review. *Educational Review*, 1-26.
- Andrews, R. (2005). The place of systematic reviews in education research. *British Journal of Educational Studies*, 53(4), 399-416.
- Bergstrom, C. (2007). Eigenfactor: Measuring the value and prestige of scholarly journals. *College & Research Libraries News*, 68(5), 314-316.
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really?. *Educational psychology review*, 15(1), 1-40.
- Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment & evaluation in higher education*, 24(4), 413-426.
- Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, 42(1), 39.
- Britz, M. W. (1989). The Effects of Peer Tutoring on Mathematics Performance: A Recent Review. *BC Journal of Special Education*, 13(1), 17-33.
- Cohen, P. A., Kulik, J. A., & Kulik, C. L. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American educational research journal*, 19(2), 237-248.
- Cooper, H., & Valentine, J. C. (2008). Research synthesis and meta-analysis. In *Handbook of Research on Adult Learning and Development* (pp. 184-202). Routledge.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. Russell Sage Foundation.
- Delquadri, J., Greenwood, C. R., Whorton, D., Carta, J. J., & Hall, R. V. (1986). Classwide peer tutoring. *Exceptional children*, 52(6), 535-542.
- Duran, D. (2006). Tutoría entre iguales: algunas prácticas. *Monográfico de Aula de Innovación Educativa*, 153-154.
- Durán, D., & Miquel, E. (2003). Cooperar para enseñar y aprender. *Cuadernos de pedagogía*, 331, 73-76.
- Duran, D., & Vidal, V. (2004). *Tutoría entre iguales De la teoría a la práctica*. Graó.



- Ellis, S., & Gauvain, M. (2013). Social and cultural influences on children's collaborative interactions. *Children's development within social context*, 2, 155-180.
- Erickson, F., & Gutierrez, K. (2002). Comment: Culture, rigor, and science in educational research. *Educational Researcher*, 31(8), 21-24.
- Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R., & Karageorgopoulos, D. E. (2008). Comparison of SCImago journal rank indicator with journal impact factor. *The FASEB journal*, 22(8), 2623-2628.
- Forness, S. R., Kavale, K. A., Blum, I. M., & Lloyd, J. W. (1997). Mega-analysis of meta-analyses. *Teaching exceptional children*, 29(6), 4.
- Franca, V. M., Kerr, M. M., Reitz, A. L., & Lambert, D. (1990). Peer tutoring among behaviorally disordered students: Academic and social benefits to tutor and tutee. *Education and Treatment of Children*, 109-128.
- Fuchs, D., Fuchs, L. S., & Burish, P. (2000). Peer-assisted learning strategies: An evidence-based practice to promote reading achievement. *Learning Disabilities Research & Practice*, 15(2), 85-91.
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79(3), 1202-1242.
- Gillies, R. M., & Boyle, M. (2010). Teachers' reflections on cooperative learning: Issues of implementation. *Teaching and teacher Education*, 26(4), 933-940.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational researcher*, 5(10), 3-8.
- Godlee, F., Gale, C. R., & Martyn, C. N. (1998). Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *Jama*, 280(3), 237-240.
- González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of informetrics*, 4(3), 379-391.
- Gordon, E. E. (2009). 5 ways to improve tutoring programs. *Phi Delta Kappan*, 90(6), 440-445.
- Johnson, B. (2001). Toward a new classification of nonexperimental quantitative research. *Educational Researcher*, 30(2), 3-13.
- Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches*. Sage.



- Kraemer, H. C., Morgan, G. A., Leech, N. L., Gliner, J. A., Vaske, J. J., & Harmon, R. J. (2003). Measures of clinical significance. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(12), 1524-1529.
- Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and special education*, 24(2), 97-114.
- Leung, K. C., Marsh, H. W., & Craven, R. G. (2005). Are peer tutoring programs effective in promoting academic achievement and self-concept in educational settings: A meta-analytical review. In Annual Conference, Cairns.
- Liu, N. F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher education*, 11(3), 279-290.
- Marshall, C., & Rossman, G. B. (2014). *Designing qualitative research*. Sage publications.
- McAuley, L., Tugwell, P., & Moher, D. (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses?. *The Lancet*, 356(9237), 1228-1231.
- Murgiondo, J. E., Tejedor, F. J. T., & Tejedor, F. J. T. (2005). Análisis descriptivo de datos en educación (Vol. 2). Editorial La Muralla.
- Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, 13(1), 48-63.
- Onwuegbuzie, A. J., & Leech, N. L. (2006). Linking research questions to mixed methods data analysis procedures 1. *The qualitative report*, 11(3), 474-498.
- Onwuegbuzie, A. J., & Leech, N. L. (2007). Validity and qualitative research: An oxymoron?. *Quality & Quantity*, 41(2), 233-249.
- Pérez, Z. P. (2011). Los diseños de método mixto en la investigación en educación: Una experiencia concreta. *Revista electrónica educare*, 15(1), 15-29.
- Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review*, 17(4), 327-362.
- Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, 95(2), 240.
- Slavin, R., & Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic reviews in education. *Educational evaluation and policy analysis*, 31(4), 500-506.



- Slavin, R. E. (1984). Meta-analysis in education: How has it been used?. *Educational researcher*, 13(8), 6-15.
- Smith, S. W., Daunic, A. P., & Taylor, G. G. (2007). Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. *Education and treatment of children*, 30(4), 121-134.
- Stenhoff, D. M., & Lignugaris/Kraft, B. (2007). A review of the effects of peer tutoring on students with mild disabilities in secondary settings. *Exceptional children*, 74(1), 8-30.
- Song, I. S., & Hattie, J. (1984). Home environment, self-concept, and academic achievement: A causal modeling approach. *Journal of Educational psychology*, 76(6), 1269.
- Tashakkori, A., & Creswell, J. W. (2007). The New Era of Mixed Methods. *Journal of Mixed Methods Research*, 1(1), 3-7.
- Topping, K. (1992). Cooperative learning and peer tutoring. *Psychology of Education*, 578.
- Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher education*, 32(3), 321-345.
- Topping, K. J. (2005). Trends in peer learning. *Educational psychology*, 25(6), 631-645.
- Wayne, D. (1993). Estadística con aplicaciones a las ciencias sociales ya la educación. *Ed. Me Crow*.
- Zeneli, M., Thurston, A., & Roseth, C. (2016). The influence of experimental design on the magnitude of the effect size-peer tutoring for elementary, middle and high school settings: A meta-analysis. *International Journal of Educational Research*, 76, 211-223.