



Mixed methods research in service-learning: an integrative systematic review

Celia Camilli Trujillo¹ · Laura Cuervo Calvo² · Desirée García Gil² · Carolina Bonastre Valles³

Accepted: 9 August 2021
© The Author(s) 2021

Abstract

In view of the challenges involved in designing a study of Mixed Methods (MM), as well as the problematics inherent in studying Service-Learning (S-L) from new research perspectives, and considering the lack of systematic reviews of MM in S-L, this study proposes to analyse the use of MM research to evaluate S-L through a systematic integrative review of scientific papers published in international databases (ERIC, DIALNET, SCOPUS, and Web of Science) using the terms ‘mixed methods research’ and ‘service-learning’. The contextual and methodological variables were analysed descriptively and inferentially using Excel and Stata programs. Moreover, 149 predefined codes were created and analysed using the qualitative program Atlas.ti to identify which terms associated with the research methodology were used most frequently in relation to the MM for S-L methodology, and how and why they were used. Of the 192 papers found, only 93 met the inclusion criteria. The results show that very few investigations specify the MM they used in sufficient detail. Only 28% provide information to estimate the effect size, suggest interventions based on S-L, and collect measures for post-intervention evaluation. This demonstrates that the relationship between MM and stages of the S-L are not complemented in the methodological design. The results demonstrated that researchers should continue working with MM in S-L for the integration of qualitative and quantitative results. This research can be a methodological guide for professionals and academics who want to investigate MM in S-L because it identifies methodological deficiencies and strengths and offers and alternative designs to evaluate the service.

Keywords Systematic review · Mixed methods research · Service-learning · Research method

✉ Celia Camilli Trujillo
ccamilli@ucm.es

Extended author information available on the last page of the article

1 Introduction

This article analyses the application of Mixed Methods (MM) to the study of didactic initiatives based on service-learning (S-L) methodology, since the study of a phenomenon as complex as S-L requires a series of complimentary approaches and research strategies. S-L is “an approach to teaching and learning that integrates community service with academic study in order to enrich the learning process, teach civic responsibility, and strengthen the community” (Puig et al. 2007: 15). Service becomes truly meaningful when the learning experience achieved while performing community service is perceived as truly useful within and outside the classroom; at the same time, apart from its educational advantages, S-L has positive repercussions on the social community. Research on MM in S-L should thus resort to several inter-coordinated strategies (Capella et al. 2019).

Within this framework, it is important to point out that MM utilizes and combines quantitative and qualitative perspectives within a single study to answer research questions and test hypotheses (Driessnack, et al. 2007; Frels and Onwuegbuzie 2013; Johnson and Onwuegbuzie 2004). Hamui-Sutton (2013: 212) states: “[MM] is more than just the sum of [these two research models], [given that] during the interphase process [...] the limitations of both models are remedied, while at the same time creating a broader panorama that strengthens the validity of the [...] results”. According to Tashakkori and Teddie (2010), MM represents the highest degree of integration and/or combination of qualitative and quantitative approaches.

This integration is what characterizes and differentiates MM from other multi-methods that apply more than one data collection method in a study or in a set of interrelated studies (Hesse-Biber 2016). Within this context, Anguera et al. (2018: 2758) point out that “the last 20 years have witnessed an explosion of studies that have, in many instances, shown that confusion regarding the meaning of multimethod and mixed methods is still rife”. They agree with the definition put forward by Bazeley (2015), for whom multi-methods are those which employ different methods or approximations in parallel or sequentially, but do not integrate them to the ultimate degree implied by their inferences.

According to scientific literature, MM have their roots in the 1960s, especially in the field of medical criminology. A decade later, the first case study based on MM emerged, in which questionnaires were used as an additional means to collect data (Maxwell 2016). As a result of these practices, from the 1980s onwards, researchers began to explore and emphasize the viability and effectiveness of data triangulation, as it was considered to be a process that went beyond the comparison of qualitative and quantitative data (Lucero et al. 2018; Moorkens 2015; Pereira 2011). One example was the study carried out by Greene et al. (1989), who observed that evaluators of educational and social programmes had recently extended the repertoire of research methodologies by adding designs that included qualitative and quantitative methods in combination. The use of that approach came to be promoted in the educational field during the 1990s. Along those lines, Nuñez (2017) traced the extended trajectory that MM had taken in the field of Anglo-Saxon education research as well as in Spanish-language studies, although in the latter, paradoxically, they are widely applied as methodologies but rarely become the object of methodological reflection, with the notable exception of studies carried out by Ponce (2011), Castañer et al. (2013), Hamui-Sutton (2013), and Pereira (2011).

The primary criticisms raised against MM have been targeted towards the ontological incompatibility and differing natures of the two methodologies (positivism—quantitative versus hermeneutic—qualitative) and the lack of explicitness with regard to the

essential elements of MM (Nuñez 2017). This has been corroborated by Bryman (2006), who affirmed that in a review of 232 scientific articles, which supposedly used MM, only 10 of the articles provided specific information on the use of qualitative and quantitative elements.

Various authors have developed specific types of mixed design research approaches based on the studies that have been carried out in different social, scientific, and humanistic disciplines. For example, according to Greene (2007), MM models can be summarised to include: triangulation design, explanatory design, exploratory design, and embedded design approaches. Tashakkori and Teddlie (2010) classify MM into concurrent triangulation (in which quantitative and qualitative data are collected at the same time) and sequential triangulation (in which data are collected consecutively). Hanson et al. (2005) explain that the above stated categories in turn include within them concurrent nested design, sequential exploratory design and concurrent transformative design. Moreover, each of these subcategories vary with regard to: the use of theoretical/practical perspectives, the procedure used to collect data (whether sequential or concurrent), the priority given to the qualitative and quantitative data, and the stage in which data are analysed or integrated, among other considerations. Eventually, the qualitative and quantitative approaches can either be made to retain their original design or structure, or they can be adapted, altered, or synthesized.

In order to ensure that MM is consistent and of use to researchers, it is necessary that they move away from the idea that this methodology involves the use of two irreconcilable paradigms. This in turn requires more emphasis to be placed on the questions, objectives, or hypothesis, as these would play a key role in deciding which research method to use. In order to achieve genuine empirical complementarity, it has become increasingly important to be aware of the similarities and differences between the two methodological approaches (Punch 2014).

MM are of particular relevance because they permit investigators to achieve an improved understanding of research problems, and to obtain a better grasp of complex phenomena than any of the two approaches would not have achieved if applied alone (Creswell and Plano 2017). MM also provide information regarding the manner in which one method evolves from the other; applied sequentially, they can be used to improve construct validity and they can uncover inconsistencies stemming from analysis results, thereby opening up new research possibilities. The triangulation of a data set can likewise improve the validity of inferences (Molina-Azorin 2016). In the opinion of Greene et al. (1989) the primordial importance of MM is found in: a) their complementarity (elaboration or extension of results from one method thanks to the findings elicited by the other one), b) their contribution to research progress and evolution (when a researcher uses one method's results to find the most appropriate use for the other one), and c) their potential for further expansion (researchers using MM seek to expand the depth and range of the research question by applying different methods to different components of the question).

Thanks to these qualities, MM have branched out from the social sciences, where they had their origins, to the fields of medicine, psychology, the health sciences, and education (Creswell et al. 2013). Education research has run into a series of controversies in its attempt to achieve a better grasp of the educational phenomenon, but MM provide researchers with a thoroughly useful alternative to face this challenge (Ponce and Pagán-Maldonado 2015). Martínez-Usarralde et al. (2019) point out that MM are the second most common methodology used to study S-L: MM are used in 37.49% of studies, superseded by qualitative studies (54.17%); quantitative methods (4.17%) and action research (4.17%) are employed to a much lesser degree. This finding is similar to that of Bukas

et al. (2020), who found that case studies were the most common (20%) followed by MM (15%), whereby only 5% of them applied action research, and only another 5% consisted in qualitative studies that did not specify which method they had applied; meanwhile, 55% of the studies they analysed did not even mention which methodology they had used. At any rate, these two overview articles merely identified the research methodologies employed in the studies they investigated, without analysing what kind of methodological benefits MM could provide for the comprehension of S-L.

Up to this point, we have provided an overview of what is understood by the concepts of S-L and MM, the differences between multimethods and MM, the history of MM and their application in the field of education, the reasons why MM have not often been the object of methodological reflection, different classifications of MM, what is needed to achieve true methodological complementarity, the potential benefits of MM, and which research methodologies are most prevalent in the study of S-L. In the following sections we go further in depth regarding the theoretical justification of S-L; we then proceed to describe the research methods applied to S-L methodology and demonstrate that it is necessary to carry out a systematic revision of MM in S-L, thereby validating the present study's objectives.

1.1 S-L methodology: key coordinates

According to Furco (2003), one of the major challenges to S-L methodology is the impossibility of formulating a universally accepted definition. This is because of the fact that the design adapts to the idiosyncrasies of the human collectives that apply it, which thereby limits its scope for generalization.

Nonetheless, there is a broad consensus (Butin 2010; Puig 2009; Speck and Hoppe 2004) on the need to identify S-L methodology as an educational methodology that challenges traditional pedagogical discourses; a methodology that expands the scope of teaching activity beyond the classroom (Wang et al. 2020), generates spaces for the development of social engagement (Marco-Gardoqui et al. 2020) and ethical behaviours (Lee et al. 2020), promotes autonomy, self-efficacy (Gerholz 2017), engagement (Folgueiras et al. 2018) and critical thinking (Kuntjara 2019), generates satisfaction with the learning achieved (Lorenzo-Moledo et al. 2021), and facilitates meaningful learning (Celio et al. 2011). This methodology “is based on the idea that mutual assistance is a better mechanism for personal, economic, and social progress than competition and the obsessive pursuit of success” (Puig 2009: 10).

In their search for epistemological positions, some authors place this methodology within the following models:

- (1) Philanthropic model: pays attention to social issues and enables students to explore their own proposals and solutions (Abel 2004). This model posits that the individuals who provide the service enjoy social and economic advantages that those who receive the service do not, which consequently may give rise to charity, altruism, or social justice (Sementelli 2004).
- (2) Civic model: is based on the fulfilment of citizens' rights and responsibilities, and the use of education as a tool to create a fair society (Waters and Anderson-Lain 2014). This model takes the existence of social injustice as its starting point and is therefore directly related to public policy and the erosion of democracy.
- (3) The socio-communitarian model: promotes a balance between individual rights and social responsibilities, and proposes a set of common values (Codispoti 2004). This

implies an individual transformation in favour of a commitment to being socially active and aware.

Deeley (2015) places S-L methodology within the 'Critical Pedagogy' model (Ortega 2009), as she believes that the model requires deep critical reflection on the part of students. Guided by the empirical evidences she has come across in her own research, she affirms that this methodology is able to engender transformation in students' manner of exercising responsibilities, in students' levels of self-esteem, and in the promotion of social cooperation among students. Owing to the intensity of the learning process, this approach also favours the acquisition of those skills, which can be transferred to students' future professional roles.

Criticisms pertaining to S-L methodology have focused on teachers' loss of control over the learning process (Deeley 2015), the consequences of potential indoctrination, and the provision of training to the teachers in question. However, it is true that, over the last decade, S-L has attracted increasing attention in the academic world (Puig 2009; Speck and Hoppe, 2004).

2 Research methods in S-L methodology

A panoramic view of the research methods adopted with regard to S-L methodology helps us understand that there is a gradual proliferation of academic, institutional, and association-related proposals concerning this type of learning, and that there is a need to identify rigorous, high-quality lines of research (Billig and Waterman 2003; Clayton et al. 2013). García and Lalueza (2019) identify two major models conditioned by the diversity of conceptual approaches with regard to this type of learning: a positivist model, based on the variables related to the implementation of S-L programmes, and a constructivist model that focuses on understanding the learning process. Eyler (2011) locates a majority of S-L research designs within the first model, and more specifically within the field of quasi-experimental designs. However, she also identifies certain deficiencies in the designs: they are not explicitly results-oriented and, it is difficult to randomly assign samples in an educational setting where such designs are applied. In contrast, she and other authors (Billig and Eyler 2003; Bringle et al. 2004) have underlined the importance of producing designs of holistic nature, which are less common in S-L research. This is because they are inclined towards studying how students process the challenges of service activity and how they develop their capacity for critical thinking and their cognitive, inter- and intrapersonal capacities, among other aspects. Holistic designs also enable researchers to outline the learning experience and anticipate results with greater precision. Billig and Eyler (2003) also recommend the use of longitudinal experimental designs with randomized samples.

Bringle et al. (2011) are supported by a wide range of authors (Billig and Waterman 2003; Lewellyn and Kiser 2014) in their assertion that S-L researchers should:

- (1) Be guided by a rigorous theoretical framework;
- (2) Describe the S-L experience clearly and in detail;
- (3) Control or explain the differences between the groups that form part of the research project;

- (4) Take measurements that are psychometrically adequate, with multiple indicators including self-reports from participants and assessments from external observers to counteract the bias of expected results in a social setting (Camilli et al. 2018);
- (5) Use various research methods and promote the convergence of results to increase the understanding, reliability, and generalizability of results (Aramburuzabala et al. 2018);
- (6) Use designs that support the conclusions reached, so that these conclusions are not arbitrary; and
- (7) Foster involvement with the learning process in general.

Speck and Hoppe (2004) asserted that a majority of the studies on S-L methodology had been published in the previous decade; other authors pointed out that one of the limitations of S-L research is the lack of studies that justify the use of specific methodologies and emphasise the importance of adopting rigorous criteria to assess the quality of the chosen research focus (Billig and Waterman 2003; Bryman 2006; Higher Education Quality Committee [HEQC] 2006; Kiely and Hartman 2011). For example, after analysing 232 articles published between 1994 and 2003 in which the authors claimed to have applied MM in their research, Bryman (2006) concluded that a majority of those articles were unable to justify their use of that particular design, apart from the lack of knowledge that generally characterizes the theoretical framework of such studies.

Consequently, the benefits of using MM in S-L are numerous. They include: (a) reinforcement of research, since the combination of quantitative and qualitative strategies and tools can cancel out the individual limitations they each might separately have; (b) support for the comprehension of socially complex phenomena, such as those pertaining to the world of education; and (c) the encouragement of data triangulation, which delivers more reliable results (Salvador-García et al. 2018).

The present study is thus justified, because: (a) the increase in studies of S-L proves that the S-L methodology is becoming more and more widespread, particularly in Europe and Latin America (Hayward and Li 2017; Nuñez 2017); (b) this growth in S-L methodology is promoting primary investigations in areas that eventually require systematic quantitative and qualitative reviews in the future, in view of the fact that (c) the amount of systematic reviews of S-L studies is scant. One of them focused on a university setting (Salam et al. 2019), but did not study the MM that were applied; the same can be said of a second S-L study in Physical Education programmes (Pérez-Ordás et al. 2021), and of a third one focusing on Spanish scientific production (Redondo-Corcobado and Fuentes 2020); the two latter ones limited themselves to identifying the methodological approximations, but none of the three systematic reviews aimed to analyse the contribution of MM to S-L.

In view of the challenge involved in designing a study of MM (Schoonenboom and Johnson 2017), as well as of the need to study S-L from other research approaches to evaluate S-L whilst applying the entire gamut of analysis parameters (Capella et al. 2019; Nuñez 2017; Salvador-García et al. 2018), and in view of the lack of systematic reviews of MM applied to S-L, the general objective of our study is to analyse how MM is used in S-L methodology via an integrative systematic review. Our study's specific objectives are the following: to identify the scientific articles in which MM is applied for the purpose of studying S-L; to analyse the contextual and methodological variables of the primary documents that meet the criteria of inclusion; to identify the methodological terms most frequently associated with MM in S-L and to assess the complementarity of MM in a research design where S-L methodology is studied.

3 Method

We have used the *integrative review* method in this study. The method aims to synthesize theories, analyse methodological problems, and examine the empirical findings of primary studies of experimental and non-experimental natures (Hopia et al. 2016). This is characterised by the integration of quantitative and qualitative primary studies that combine theory-based information with empirical literature. The research process of this study includes the following stages: formulation of the problem or the review question literature review; evaluation of the MM; utilisation of quantitative and qualitative data, tools and techniques, and interpretation and discussion of results.

4 Research questions that guide our systematic review (review questions)

What types of MM are used to study S-L in scientific articles? How are they incorporated into the methodological design of these primary studies?

4.1 Literature review

The literature review was based on a search of bibliographic databases such as *Education Resources Information Center*, DIALNET, SCOPUS, and *Web of Science*. We searched the terms ‘Mixed Methods’ and ‘Service-Learning’ in the title, keyword list, or abstract, depending on the requirements of the database in question.

4.2 Inclusion criteria

The primary studies were included based on the following criteria: (1) empirical; (2) not time-related; (3) written in English or Spanish; (4) include abstracts; (5) have complete texts; (6) are research articles; (7) MM is used as a research methodology; and (8) related to S-L, regardless of discipline, field, or context. English and Spanish are the languages in which most S-L studies have been published (> 99%) in order to ensure greater representability of results, with English publications being the most numerous. Specifically, the reason we included Spanish in our search was due to the marked influence of Anglo-Saxon research on the Latin American S-L scene (Gezuraga and García, 2020; Lotti and Betti 2019), and because in Spain one can observe a notable increase in publications dealing with S-L. From 2015 to 2018, three times as many articles on S-L were published in Spain than from 2003 to 2014 (Redondo-Corcobado and Fuentes 2020).

4.3 Review process: stages and flow diagram

For our review process, we elaborated a manual in which the studies were classified (by author, year of publication, title, etc.). This enabled us to establish common criteria to search and analyse the documents. All gathered information was stored in an Excel spreadsheet, shared on Google Drive by the four researchers involved in this study. Each research team member initially reviewed a database: if certain documents could not be retrieved

in fulltext, the entire team attempted to locate it. We also paid special heed to provide detailed justification for the inclusion or exclusion of a study from our selection. If one of the research team members had any doubts or left a field without a response in the Excel spreadsheet, the entire group reviewed and consulted together regarding the item in question.

In the *first round* of the review process, we found 192 documents; of these, 103 met the inclusion criteria. Consequently, of the 89 studies eliminated, 31 studies were eliminated due to repetition, 2 due to absence of abstract, 15 due to them being theory-based, 7 due to them being related to volunteering, 18 due to them being written in a language other than Spanish or English, 7 due to them not being related to S-L, and 9 due to them not being related to MM (inter-rater reliability = 87.2%).

During the *second round* of the review process, a much more detailed analysis of the 103 remaining documents was carried out, in consequence, 10 studies were eliminated because they focused on the experiences of a service or a project that was not related to S-L methodology (inter-rater reliability = 91.4%).

The remaining 93 documents were once again reviewed by the researchers during the last round of the review process (inter-rater reliability = 98.7%). The discrepancies were discussed by the four researchers, and ultimately, the 93 documents in question made up the final sample of the literature that met the inclusion criteria.

The review process and its constituent stages are represented in Fig. 1, which thereby meet the PRISMA criteria (Hutton et al. 2016).

4.4 Coding system and type of analysis

A coding system based on *contextual variables* was established based on the analysis of the 93 documents. The variables included (1) publishing year, (2) disciplines, (3) publishers, (4) universities, and (5) references. The coding system also took into account *methodological variables* related to (6) the objectives of the study, (7) the MM used, (8) the sample, (9) the data-collection tools or techniques used, (10) the data-analysis technique used, and (11) the effect size.

In quantitative terms, the contextual and methodological variables were analysed descriptively (frequency and percentage) and inferentially (analysis of time series and comparison of averages for separate groups) via the use of Excel and Stata 13.1. With respect to the methodological variables, we only performed a qualitative analysis of the MM used and the classification thereof using the qualitative Atlas.ti 8 program, as recommended by Gibbs (2018). The process involved the identification of the patterns of similar ideas, concepts, or themes to subsequently establish relationships and integrate information in line with the theoretical foundations of the study (Miles and Huberman 1994).

We created 149 predefined codes in English and Spanish in order to identify which among the terms associated with the research methodology were used most frequently in relation to MM for S-L methodology, and how and why they were used. These codes were created in accordance with two criteria: a review of the keywords for the articles selected; and a review of the subject indices in the books by Creswell and Plano (2017) and Tashakkori and Teddlie (2010).

Once they had been identified, the codes were divided into four categories: method, sample, data collection, and data analysis. They were then processed in Atlas.ti via the use of the ‘auto-coding’ function. The methodological information analysed was obtained exclusively from the corresponding section on research methods or methodologies in each

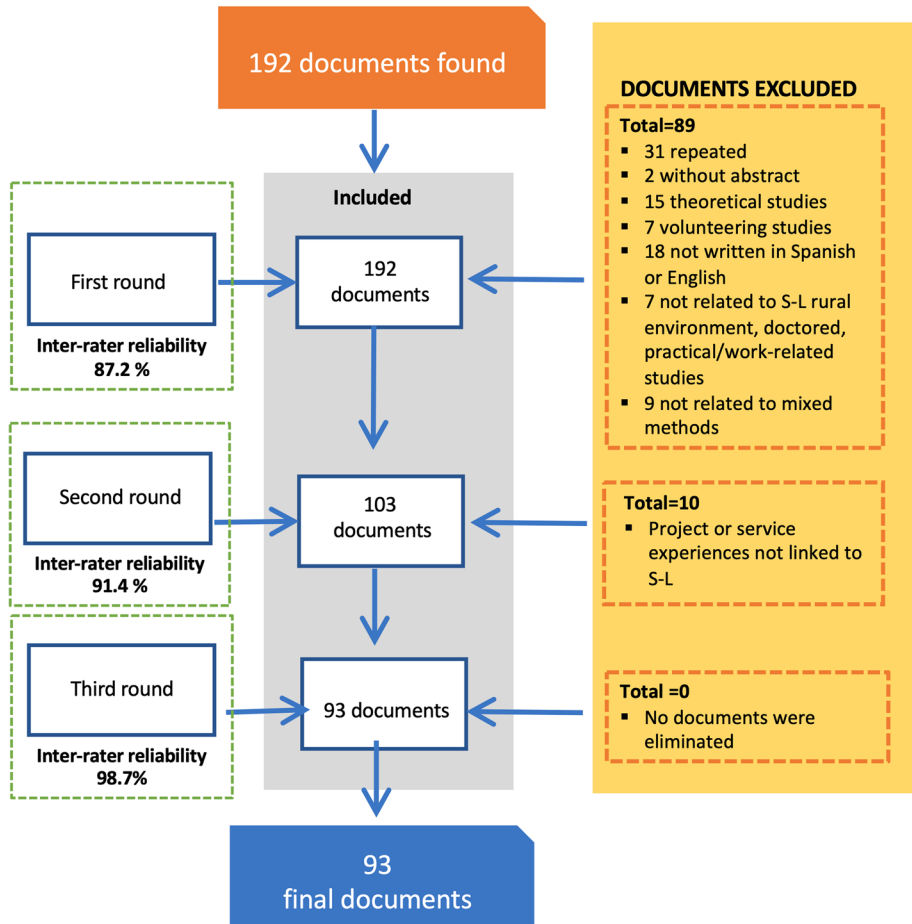


Fig. 1 Flow diagram

of the selected studies. We also obtained a co-occurrence index when two codes were found in the same quote, which helped us identify the interactions between codes. Codes are conceptual constructs that allow us to assign units of meaning to descriptive or inferential information compiled in the course of a study, thereby providing explanations and elucidating processes and perspectives with regard to the object of study. Quotes, on the other hand, are text fragments: codes are thus conceptual formulations, summaries or groupings of the quotes (Eldh et al. 2020).

5 Results

In order to fulfil our study's specific objectives, e.g., 'to identify the scientific articles in which MM is applied for the purpose of studying S-L' and 'to analyse the contextual and methodological variables of the primary documents that meet the criteria of inclusion', while at the same time fulfilling the criteria of a systematic review, we produced a general

description of the studies by carrying out a statistical analysis of contextual variables (year of publication, country, institution, speciality, journal, and bibliographical references directly related to the research methodology referred to in the studies) and methodological variables (study design, characteristics of the sample, data gathering techniques, data analysis techniques applied in the studies). We then carried out a quantitative analysis with two goals. The first goal was to fulfill our specific objective ‘to identify the methodological terms most frequently associated with MM in S-L’. The second goal was to have an analysis that allowed us to grasp which types of MM are used to study S-L in scientific articles, and how MM are used in the studies featured in this review, with several examples that serve to illustrate their use: both questions respond to the specific objective ‘to assess the complementarity of MM in a research design where S-L methodology is studied’.

5.1 General description of the studies: statistical analysis

5.1.1 Contextual variables

The percentage of studies published per year within the selected period is as follows: 1.08% (2005, 2007); 3.23% (2009, 2013); 4.30% (2010), 6.45% (2011, 2012); 7.53% (2014); 8.60% (2017); 12.90% (2015, 2019); 15.05% (2018), and 17.20% (2016). The year in which the most articles meeting our criteria were published was 2016 (17.20%), followed by 2018 (15.05%). Two thirds of the publications (66.67%) are from the years 2015–2019, and the remaining third (33.33%) were published over a much longer period, namely 2005–2014.

When analysing the rising trend in the number of publications on the subject, we observed a clear pattern of progressive incremental growth: coefficient of 1.05, $z=4.84$, $p<0.001$ (CI at 95%: 0.62, 1.48). This linear trend, with a certain degree of variation, can be observed in the graph in Fig. 2.

Of the 93 articles analysed, a majority of them came from American universities (67.74%), followed by those from higher-educational institutions in Spain (8.60%) and Canada (4.30%). Experiences with S-L in Spanish universities come from documents published in English or Spanish, but appear in the databases indistinctly with descriptors in both languages, with the exception of three studies that were only recovered via descriptors in Spanish. This criterion affects that percentage, but not to the point of changing the position of Spain in the overall ranking.

As shown in Table 1 (below), by grouping data based on the income level of the authors’ countries of origin (in accordance with the classification produced by the World Bank, n.d.), we found more studies (89.8%) to be produced in the 11 countries with higher incomes (Table 1) (mean = 8.1; $SD=19.7$; median = 1; interquartile range: 1–4) in comparison to the eight countries with middle to lower incomes (Table 1) (mean = 1.3; $SD=0.5$; median = 1; interquartile range: 1–1.5). According to the Wilcoxon test ($z=1.15$; $p=0.250$) and the parametric comparison ($t(17)=0.98$; $p=0.171$), the difference was not statistically significant. While the effect size was moderate ($d=0.45$; 95% CI: -0.48 ; 1.37) (Table 1), this may, however, be attributed to the low statistical power.

With regard to disciplines, the studies mostly focused on the fields of Humanities and Social Sciences (66.67%), followed by Biology and Biomedicine (16.13%), Physics and Technological Sciences (9.68%), and Agricultural Sciences (specifically, Horticulture) (1.08%). There were also a series of studies that carried out research on a range of different disciplines (6.45%).

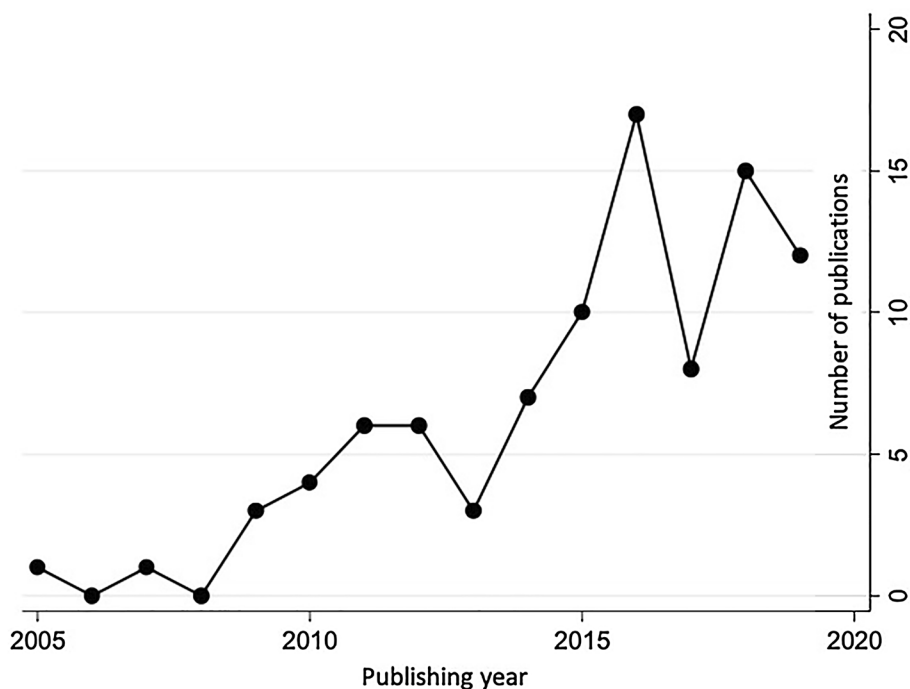


Fig. 2 Rising trend in the number of publications

Table 1 Percentage of publications from the authors' university of origin by the countries' economic category

Countries with higher incomes

USA	Spain	Canada	Australia	UK	Germany	Chile	Singapore	Cyprus	UAE	Israel
67.74	8.6	4.30	1.61	2.15	1.08	1.08	1.08	1.08	0.54	0.54

Countries with middle to lower incomes

China	Georgia	Turkey	South Africa	Ghana	India	Malaysia	Pakistan
2.15	2.15	1.08	1.08	1.08	1.08	1.08	0.54

The total percentage was adjusted accordingly for publications with authors from two countries

A significant percentage of the articles analysed were published by universities (27.9%), specifically the University of Michigan, Georgia University, Louisiana University, University of Alicante, Loyola Marymont University, University of Technology Sydney, Indiana University, Autonomous University of Madrid, and Adventist University of the Plata. These universities were followed by a number of renowned and well-indexed publishers such as SAGE (18.27%), Taylor & Francis (11.82%) and Elsevier (8.6%).

Further, 9.67% of the articles analysed were published by scientific associations, such as the 'Spanish Inter-University Association for Teaching Research', the 'American

Society for Horticultural Science', 'American Dental Education Association', 'Australian Chiropractors Association', 'Education Association of South Africa', 'American Society for Engineering Education', and the Turkish 'Association for Social Studies Educators'. The remaining 19.35% of the articles analysed were published by private companies and institutions.

With regard to the references used by these primary studies on research methodologies, only 16.42% ($n=23$) of these studies made reference to the specialist documentation on MM. Of the remaining 83.57% ($n=117$) of the primary studies, 52.14% ($n=73$) made reference to qualitative research methodologies, 19.28% ($n=27$) made reference to research methodologies in general, and 12.14% ($n=17$) made reference to quantitative research. Moreover, the authors most often mentioned are Miles and Huberman (12 mentions), Creswell and Strauss (30 mentions), and Corbin (11 mentions).

5.1.2 Methodological variables

A majority of the selected studies had a major focus on exploring the repercussions or development of the S-L methodology. The research objectives were found to be clear, well-integrated, and equipped with a methodological framework for the MM in only 21% of the studies analysed.

Of the 93 documents that apply MM, only 16.12% offer a detailed description of which MM was applied, as opposed to 'sample characteristics', in which 83.87% do indeed include detailed information. With regard to the samples of the studies analysed, 65.38% of the studies had samples with 150 participants or fewer, while 34.62% had samples with more than 150 participants. The most frequently used data collection tools and techniques in the S-L methodology were questionnaires (20.43%), surveys (18.27%), and interviews (15.05%). These were followed by personal reflection (7.52%), focus groups (7.52%), open-ended questions (7.52%), Likert scale (6.45%), diaries (5.37%), and observation (3.22%). These tools and techniques to a lesser extent included inventories (1.07%), life-history (1.07%), tests (1.07%), self-reporting (1.07%), blogs (1.07%), academic performance (1.07%), and academic exams (2.15%). Finally, 92.48% of the documents studied specified the data analysis techniques they used.

With respect to the information included in the articles, we must bear in mind, from a quantitative perspective, that not all of the articles selected were of sufficient quality to help us estimate the standardized effect size, and to then compare their effect size to that of other studies. In fact, only 28% of the studies analysed provided sufficient information to help us make such an estimate. These studies proposed interventions based on S-L and performed baseline measurements, followed by new assessments once the intervention had taken place. The measurements in question were quantitative, and the studies in the results sections provided sufficient information to enable the estimation of the effect size (this generally included sample size, mean, and standard deviation for both, pre- and post-intervention assessments. Although this information was partially available in some cases, we were able to reliably estimate the effect size, thus, the effect size to be estimated corresponded to Cohen's d , thereby providing a standardized estimate for the different mean values). Further, although the assessments included a wide variety of variables and measurement methods thereof, very few (22.58%) assessments used standardized and validated instruments. Consequently, a number of different approaches were adopted to measure the effectiveness of the interventions.

Table 2 Codes and quotes

	Codes		Quotes	
	No	%	No	%
Method	29	13.03	909	13.05
Sample	10	10.62	741	10.46
Data collection	46	43.29	3020	43.37
Data analysis	61	33.06	2306	33.12
Total*	146	100	6976	100

*Does not include the terms 'MM', 'qualitative', and 'quantitative', as they can be found in any of the four categories studied. This analysis corresponds exclusively to the 'Methodology' section in each of the selected articles

With respect to the 67 articles that failed to meet the necessary quality requirements for quantitative synthesis, we observed the following attributes among them: (1) They did not provide quantitative data, or if they did, it was only in a descriptive form (26.9%), (2) They did not include a pre-post assessment, or the design used by these studies was exclusively cross-sectional (19.4%), (3) They only provided data regarding proportion (13.4%), (4) They did not give the standard deviation or the exact value of p (10.4%), (5) The results for the studies were presented only in the form of graphs (7.5%), (6) While the results of the statistical pre-post comparison test were provided, the raw data were not; consequently, although it was possible to estimate the approximate effect size, it was not possible to estimate the weighting of the study, which could have caused a bias in the results (7.5%), (7) The type of design prevented us from evaluating the effect of the intervention itself (6.0%), (8) Only multivariate analyses such as linear regression and MANOVAs were performed (4.5%), (9) Only the p value was provided (3.0%) and (10) The variable assessed was irrelevant to S-L methodology (1.5%).

5.2 MM research in S-L methodology: qualitative analysis

The 149 codes obtained from the terms associated with MM were subjected to qualitative analysis for the 93 documents selected. These codes generated 6,976 quotes, which were distributed across four major categories (Table 2). These are largely accounted for by codes and quotes related to data collection (43.29% and 43.37%) and data analysis (33.06% and 33.12%). A smaller number of codes (19.82%) and quotes (13.05%) were linked to research method and design, while just 6.85% of codes and 10.46% of quotes were linked to study sample.

Of these 149 codes, the first 10 terms were 'participants' (6.06%), 'significance' (5.92%), 'survey' (5.45%), 'questions' (4.93%), 'interviews' (4.38%), 'method' (3.66%), 'scale' (3.64%), 'item' (3.34%), 'qualitative' (3.26%) and 'design' (2.98%). The last 10 codes (0.01%) were applied for the terms 'confidentiality', 'factorial analysis', 'generalise', 'inference', 'inquire', 'paradigm', 'pattern', 'procedure', 'steps', and 'synthesis'. Terms such as 'action-research', 'causality', 'checks', 'cluster', and 'confirmatory' did not appear. Interestingly, the MM used in each of the 93 documents selected (e.g., 'explanatory design', 'exploratory design', 'embedded design', 'triangulation design', etc.) were not among the most frequently used terms.

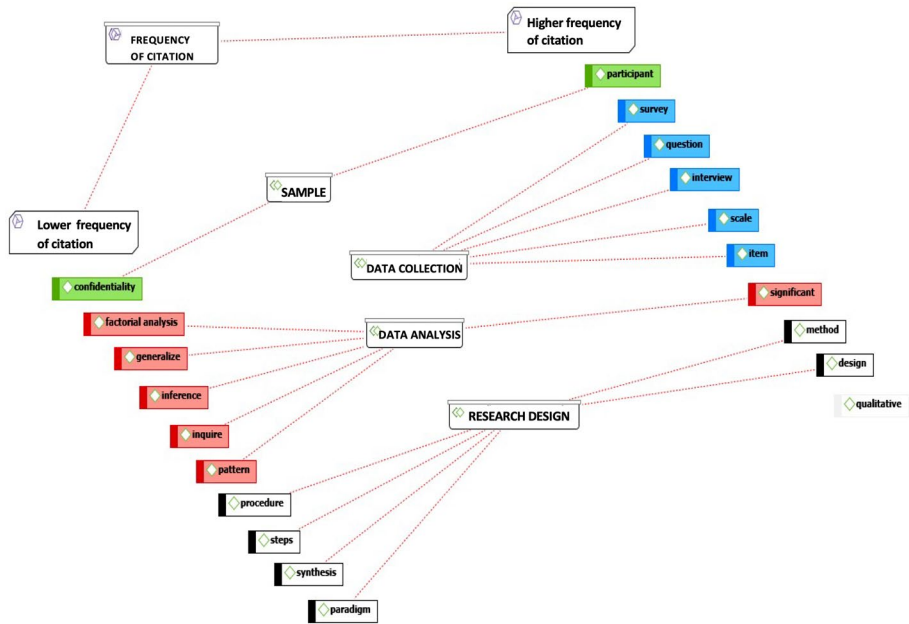


Fig. 3 Percentage of co-occurrences between codes

Table 3 Percentage of co-occurrences between codes

	MM (21 quotes)		Qualitative (240 quotes)		Quantitative (160 quotes)	
	Codes	%	Codes	%	Codes	%
Method	50	56.18	116	19.90	83	22.19
Sample	4	4.49	40	6.86	22	5.88
Data collection	14	15.73	247	42.37	160	42.78
Data analysis	21	23.60	180	30.87	109	29.14
Total*	89	100	583	100	374	100

*This analysis corresponds exclusively to the 'Methodology' section in each of the selected articles

When the keywords related to research methodology (e.g., 'mixed methods', 'quantitative' and 'qualitative') were analysed individually, the top results were for the term 'qualitative', with 240 quotes (3.26%). This was followed by the terms 'quantitative', with 160 quotes (2.17%), and 'mixed methods', with 21 quotes (0.29%).

A more specific analysis (Fig. 3) indicates that the methodological terms most frequently used in the studies of S-L methodology were related to data collection, while the codes that appeared the least were linked to research method and design, and data analysis. It seems that when it comes to working with MM, the greatest difficulties lie in the design and in finding a way to analyse the data that enables the integration of MM. Conversely, it is much easier to select techniques for the collection of quantitative and qualitative data.

Despite the foregoing observations, when we analysed the research methodology as a whole (Table 3), we observed a higher degree of co-occurrence (56.18%) between the

terms ‘mixed methods’ and ‘study design’. However, the level of co-occurrence between codes became weaker when we linked them to ‘data collection’ (15.73%) and ‘study sample’ (4.49%). This indicates that the term ‘mixed methods’ is much more closely related to design, rather than the remaining elements that comprise the research methodology. Table 3 demonstrates that a significant number of quotes were associated with qualitative methodology; it had 80 more quotes in comparison to quantitative methodology. It seems that ‘qualitative approach’ has a greater presence in studies on S-L methodology than ‘quantitative approach’ on the whole.

However, there was a greater degree of co-occurrence among the terms ‘qualitative’ (42.37%), ‘quantitative’ (42.78%) and ‘data collection’ (Table 3), followed by ‘data analysis’ (30.87% and 29.14%, respectively). Both results indicate that in S-L methodology, a significant emphasis is placed on identifying and describing the tools and techniques for collecting qualitative and quantitative data, as well as the techniques used for analysis.

5.3 Classification of MM research: qualitative analysis

For the qualitative analysis, the 93 documents have been identified with a number or, failing that, the name of the authors is mentioned and can be consulted in the supplementary file. The verbatim from the primary documents have been indicated with quotation marks.

In relation with our research question ‘which types of MM are used to study S-L in scientific articles’, quantitative analysis shows that the most frequently employed methods (15.05%) in which the scientific articles offered meticulous detail about their application were: *concurrent design* (12,22,29,51,64), *embedded design* (35,54), *explanatory sequential design* (17,26,28,37,41,67), and *triangulation design* (12,31,66).

Concurrent design is an integrative design methodology that has been used by Mampane and Huddle (2017) to bring into account the different sub-systems within the service experience. This method has also been used in combination with a specific type of concurrent design, namely the *concurrent transformative strategy approach*, in which the qualitative and quantitative foci are given equal weighting within the study (29). Seemingly, in S-L methodology, it is a common practice to combine concurrent design with other methods such as *longitudinal studies* or *cross-sectional studies* (51, 64).

Continuing with our analysis of these MM approaches, we found two studies to have used the *embedded design method*, in which a single set of data is insufficient to fulfil the objectives of the study. Consequently, the research questions are required to be addressed via the use of different types of data. Huff et al. (2016) adopted this approach to address the question of how engineering students perceive the impact of an S-L programme on their professional development. Likewise, the study by McWhorter et al. (2016) confirmed that study designs involving multiple cases were becoming increasingly popular in the field of business research.

At the same time, the *explanatory sequential design method* was applied with increasing frequency in S-L methodology, with there being a total of seven studies in this regard. One example is the study by Cone (2009). The prior quantitative stage consisted of a pre- and post-test study to measure the effects of community-based S-L methodology. This was followed by interviews to round out or add depth to the most relevant data, or clarify any data that was considered unclear or noteworthy with regard to the various debates and activities that were carried out.

Finally, although Creswell et al. (2013) indicate the most common and well-known MM design is *triangulation design*, which consists in directly comparing and contrasting

quantitative statistical results with qualitative findings, or validating or extending quantitative results via qualitative data. Only three studies applied it. Li et al. (2019) studied the effect of different service experiences in adolescents suffering from cerebral palsy. For their part, Hirschinger-Blank et al. (2009) concluded that levels of political awareness are raised among students when they are immersed in S-L projects. Packard et al. (2016) evaluated the inter-professional education experiences of health sciences students throughout university.

However, to continue with our research question regarding which types of MM are used to study S-L in scientific articles, authors provided less details regarding their application (84.95%) when the methods were applied in studies using experimental designs (12,33,34,39,52,55,59,61,62,65), quasi-experimental designs (1,6,20,21,39,46,59,71,72,73,77,79,80), case studies (28,36,47,54,88), longitudinal studies (28,64,81), cross-sectional studies (21,39,51), cross-case studies (28,54), descriptive studies (15,25,70), thematic analysis (4,8,16,22,35,39,41,75,78,81,87), content analysis (7,12,13,21,33,50,59,84,86,87), discourse analysis (45), and Grounded Theory (2,3,4,9,10,46,48,63,68,85,88).

For example, a total of 11 studies had adopted *experimental designs* and applied the same pre- and post-test techniques to measure intervention (52,61). In these studies, the participants were randomly assigned to evaluate the effect of S-L methodology on students' sensitivity to social issues, in addition to its effect on civic and social skills (34), among other factors. For the experimental group, the specific tasks chiefly included the reading of materials related to S-L methodology (55), specific activities, and weekly monitoring assessments (12,55). Without any exception, these experimental designs began with the application of quantitative research. Later, once the information was gathered and analysed, the designs proceeded to apply a qualitative approach as semi-structured interviews (33) or narrative responses to open-ended survey questions (52). There were also three *non-experimental descriptive studies*, whether ex post-facto or survey-based (15,25,70).

The *quasi-experimental designs* were characterized using intentional groups. These groups were not always equivalent to one another (6,20) and comprised an experimental and a control group with pre- and post-measurements. Other examples included were the survey-based quasi-experimental studies (52,59). Only in the study conducted by Chiva-Bartoll et al. (2018) did the independent variable incorporate two levels based on the type of intervention (in line with the chosen S-L model). Generally speaking, the S-L control group was an inactive group in which students were taught in accordance with a methodology based on the model of traditional classes, theory-based lectures, and/or research projects (73).

Continuing with this type of design, we found the *cross-sectional* approach to have been utilized as a quantitative focus. Authors utilized 'a mixed-methods design consisting of a quasi-experimental design to evaluate differences in test scores and a cross-sectional design to describe students' attitudes' (39), and they also used an observational method as part of the qualitative focus (51).

Longitudinal studies enabled researchers to conduct a more profound analysis of the progress made by students during the S-L experience over an extended period. It also enabled them to study how experience influenced students' roles (22) and responsibilities as future professionals (64), and conduct retrospective assessments of the years of experience in S-L methodology (81).

For their part, the *case studies* related to S-L methodology were either standalone cases, multiple cases, or exploratory case studies. Although the case study method is claimed to have one of the best theoretical groundings and explanations, the studies did not specify how the method was integrated into the MM approach and its constituent stages (36).

In S-L methodology, the case studies are usually subjected to the process of *cross-case analysis*, a method of analysis that involves the exploration of similarities and differences between cases with the aim of lending support to empirical generalization and theoretical predictions (28). Given the enormous variation and complexity of contexts and environments within the field of S-L, this cross-sectionality is present throughout the entire structure of the study (54).

Content analysis and *discourse analysis* were also used in relation to S-L methodology as qualitative methods for analysing personal reflections (13) and interviews (33) in which the number of words was counted, and the proportion of their occurrence in the text was analysed (21). *Grounded theory* was also used as an additional qualitative method: 'We used grounded theory to code and analyse qualitative data from the interviews and the critical reflections' (10).

To finish, Table 4 presents several examples that illustrate who the selected scientific articles applied MM and combined the results of qualitative and quantitative data, thus likewise answering the question as to how they are incorporated into the methodological design of these primary studies.

6 Discussion and conclusions¹

This study's general objective was to analyse how MM are applied in S-L by carrying out an integrative systematic review.

Regarding our first specific objective, a total of 93 scientific articles were quantitatively and qualitatively analysed through a series of contextual and methodological variables reflecting how MM are applied to S-L. Regarding our second objective, the obtained results of contextual variables confirm that the amount of studies of the application of MM in S-L notably increased from 2015 to 2019: the leading country was the US, and Spain was in second place. This observation is confirmed by further authors who report a significant increase in publications on S-L in Latin America (Gezuraga and García 2020; Lotti and Betti 2019).

Studies about S-L with MM are likewise most prevalent in the field of Humanities and Social Sciences, whereby Education Sciences play a particularly prominent role (Creswell et al. 2013), as well as in programmes carried out within a university context (Redondo-Corcobado and Fuentes 2020). In the primary documents we analysed, the number of references relative to MM are less frequent, probably because, as Nuñez (2017) points out, very few studies actually provide a true methodological reflection regarding how and why MM should be applied in S-L.

Concerning the subject of methodological variables, the number of articles that provide a detailed explanation of the MM they applied is very low. This finding corroborates the above-mentioned explanation given by Nuñez (2017); at the same time, it reinforces the need to continue working on research objectives as a methodological framework of MM and their design in the study of S-L (Schoonenboom and Johnson 2017).

Questionnaires are the tool most often employed, and the documents we analysed are characterized by their specificity in describing their sample. Both of these observations

¹ Documents that met the inclusion criteria which are mentioned in this section are in the supplementary file.

Table 4 Examples of the integration of quantitative and qualitative results

Document No	Examples	Page
1	<p>Students' total correct Parkinson's disease (PD) knowledge scores significantly increased after participation in service with a large effect size (<i>pre-test mean</i> = 14.77, [<i>SD</i> = 2.57]; <i>post-test mean</i> = 19.69 [<i>SD</i> = 2.06], <i>Cohen's d</i> = 1.64) and a paired t-test indicated a significant change in students' Parkinson's Attitude Scale scores (<i>t</i> (34) = 2.22, <i>p</i> < .05). Ninety-one percent of students (31) indicated they would recommend the program and 82% (29) indicated they would participate again. During focus groups, students reflected on the relationships they formed with their buddies, indicating the program provided a support system while helping them learn about PD. Patients indicated the program expanded their social circle and meetings with first year medical students were beneficial</p>	1
6	<p>The results of the quantitative analysis revealed that the programme resulted in significant effective personality improvement, as much in the comparison of pre- and post-test measurements of the experimental group, as well as the post-test measurements between the experimental and control groups. The qualitative analysis complemented these results by showing that pre-service teachers also made allusion to improvements in dimensions such as social self-realisation, self-esteem and problem solving self-efficacy. In addition, their discourse defended the suitability of service-learning to strengthen and implement the theoretical lessons learned during this training</p>	327
17	<p>Movement integration (MI) is a strategy within comprehensive school physical activity programs (CSPAP). School-university partnerships are recommended to leverage teachers' capacity to use MI. A mixed method process evaluation was conducted of the first year of implementing Partnerships for Active Children in Elementary Schools (PACES). There were no significant differences between intervention classrooms and control classrooms MI promotion. Differences approaching significance (<i>U</i> = 5, <i>p</i> = 0.04, <i>d</i> = 1.2) were observed when comparing classrooms that received two (community of practice, community-based participatory research) or three components (two components plus service learning) of the intervention and classrooms that received one (community of practice) or no components. Qualitative findings revealed that teachers in classrooms that were more successful responded more favorably to the intervention components than teachers in classrooms that were less successful. Quantitative and qualitative results supported the effectiveness of community-based participatory research as a component of PACES</p>	
45	<p>When examining the data, we found mixed results. In the qualitative data-set, the majority of the participants described thoughts and attitudes that align with the six learning outcomes we identified for social justice service-learning. However, there were some participants who expressed ideas that led us to believe that the experience had fostered paternalistic attitudes for them, and the experience also seemed to support rather than contradict stereotypes for others. When examining the quantitative data-set, we found that there was no statistically significant difference on response items from the beginning to the end of the semester; however, the differences, while not significant, did yield some interesting observations. Though the statistical analyses yielded a non-finding, the open-ended response item at the end of the Munroe questionnaire elicited responses that suggest that the experience did foster social justice perspectives for some participants</p>	90–91

Table 4 (continued)

Document No	Examples	Page
51	In this study, we sought to learn if adult undergraduate students who participated in online service and those who participated in service on-site identify similar service-learning experiences. According to the data, the participants in this study experienced comparable outcomes and experiences in their service-learning opportunities. Both the quantitative and qualitative data support the fact that students, in each setting, had an overall positive experience. It is evident that each population benefited from the service-learning opportunities	16
67	Qualitative data were used to assist in explaining and interpreting the findings of the quantitative data in regard to Personal Science Teaching Efficacy (PSTE) and Science Teaching Outcome Expectancy (STOE) In summary, it was evident from the quantitative and qualitative data that the confidence levels of most participants increased during the course. Interview participants attributed changes, or lack thereof, in confidence to various factors. However, when mastery experiences such as community-based service-learning (CBSL) were present (i.e., the opportunity to teach science to diverse student groups) and supplemented with explicit discussions and activities about diversity, such experiences played an important role in reinforcing or raising participants' confidence levels (PSTE) and expectations of student success (STOE) to varying degrees	374 378
79	"The quantitative study shows that, unlike the control group, both experimental groups improved their social skills and attitudes after de service-learning program. The qualitative study helped to complement and enhance the understanding about the effects of the program to refine the skills and attitudes achieved. The three categories obtained were: group consciousness, implication and group organization skills and communication skills."	278
86	"Significant effects regarding students' development of their self-efficacy, self-concept and attitude to being engaged were found. The qualitative results provide a deeper understanding of these changes, including the different perspectives from students and from charitable organizations [For example] Regarding self-concept, the quantitative data revealed effects for time in the case of change in self-concept as well as effects for the different projects. The findings in the qualitative data have shown, that on one hand, the students perceive their relevance in society. The students realized that their capabilities can make a valuable contribution in the society. On the other hand, students received personal insights to their own strengths and weaknesses	1 10–11

were also made in the systematic review provided by Pérez-Ordás et al. (2021), while Redondo-Corcobado and Fuentes (2020) specified that the most common questionnaires had been elaborated ad hoc.

One of our study's limitations is that less than one-third of the studies met the criteria for inclusion in the meta-analysis. Even amongst those that do, the samples are very small, the methodology is unclear, and the valid assessment methods are not used. Although many studies propose new scales, a great majority of these scales have not been psychometrically verified, nor is any information provided in relation to their validity or reliability beyond the simple fact of internal consistency. When viewed as a whole, the studies are highly heterogeneous in terms of their objectives and methods, and the representability of their samples.

Fulfilling our study's third specific objective, our analysis of which methodological terms are most frequently associated with the use of MM in S-L confirms that the two greatest difficulties lie in study design and in how data should be analysed to achieve an MM-specific integration. It was easier, on the other hand, to identify and describe which tools and techniques were used in quantitative and qualitative data collection. This finding agrees with the review by Schoonenboom and Johnson (2017), who affirm that it is essential that study designs should be adjusted to the concrete situation and to each study's specific research questions. We likewise find a greater quantity of terms related with qualitative research, which, indeed, is the methodological approximation most often employed in S-L. This result is confirmed by Martínez-Usarralde et al. (2019), Bukas et al. (2020), and Redondo-Corcobado and Fuentes (2020).

In terms of what types of MM are used to study S-L in scientific articles and how are they incorporated into the methodological design of these primary studies, it is confirmed that there are few studies that refer to the methodological complementarity of the quantitative–qualitative approach (1,6,19,20,34,37,80) and its importance to the research process (34), in the sense of complementarity between stakeholder voices (19) or data collection techniques (12,22) or in the data analysis (1) or occasions in which the researchers themselves move between inductive and deductive modes of thought (6) or when both modes are combined in the discussion.

For example, only the study by Mampane and Huddle (2017) alludes to the fact that the MM approach forms part of the pragmatic paradigm. For its part, the study by Gil-Gómez et al. (2016) presents the quantitative and qualitative paradigm to contextualize and justify the experimental design and the questions and techniques that are used: 'Likewise, taking the qualitative paradigm as our starting point, we formulated a research question in order to complement the quantitative findings' (34). Another example is the study by Brizee (2014), which clearly demonstrates how descriptive statistical analysis favours data triangulation: 'With Data analysis, I used descriptive statistics and grounded theory to analyse and triangulate data' (46). It also shows how qualitative analysis complements quantitative analysis: 'As the interview transcripts were reviewed and coded, categories were developed, comparisons made, and connections identified that were able to further illuminate the quantitative findings and integrate the data' (38).

These methodological deficiencies are also evident in the integration of the various cycles or stages of S-L methodology with the stages of the research design, inasmuch as they are carried out separately (27). For example, although the studies describe how the various stages of S-L methodology helped the researchers achieve their aims, 'Sensitivity to the community partner's culture during all phases of planning, delivering, and evaluation services is needed for successful international health efforts' (78). Further factors, such as how the teaching and learning methodologies (e.g. project-based methodologies)

were applied during these stages of the service, how community needs were assessed (27,76), and how the sample was selected (6), were not integrated into the research design or explained within the context of the said design and vice versa. According to Greene (2007), in order for the various approaches to feed into one another, and thereby generate an improved understanding of the phenomenon being studied, it is essential to bring together the ‘various mental models’ in the same search space to foment a respectful dialogue between them.

Nevertheless, in view of the complex methodological elaboration involved in MM design (Schoonenboom and Johnson 2017), and taking the contributions of Creswell and Plano (2017) as well as of Molina-Azorin (2016) into account, we can affirm that studies using MM in S-L have benefited from the methodological complementarity that emerges from amplifying the results of one method with those of another in a synergetic feedback loop—one of the three great advantages of MM.

Reflecting our study’s fourth specific objective, an example of such complementarity can be found in the study by Cumberland et al. (2019), who made unexpected discoveries of further variables that could fulfil a certain function in S-L, but which had not been emerged from the quasi-experimental design alone. The role played by students of medicine in the relations they forged with their service partners turned out to be crucial in dealing with patients suffering from neurological disorders such as Parkinson’s disease. That role was so outstanding that the study authors decided to incorporate a questionnaire scale on the subject of friendship and social roles in their future studies of S-L, since friendship and social roles turned out to be a variable that bore a positive influence on the experiment’s results.

Qualitative analysis has likewise confirmed the results of the quasi-experimental study carried out by Chiva-Bartoll et al. (2019), showing that teacher trainees participating in S-L activities tend to improve along dimensions such as social self-fulfilment, self-confidence, and self-efficacy when they face the task of resolving problems: this information could only be obtained thanks to the methodological integration that is typically specific to MM. Li et al. (2019) affirmed that although they found significant differences in favour of the two service groups (adolescents at risk and adolescents with cerebral palsy) compared with subjects who had not carried out the service, only “qualitative analysis provided evidence of the mechanism underlying service-learning” (Li et al. 2019: 1).

Therefore, it can be concluded that there is a methodological vacuum in MM research projects that study S-L methodology as a formative methodology for teaching and learning to foster the development of personal and social skills. Paradoxically, while the MM approach provides a methodological framework for the studies analysed, very few studies provide a detailed description of the MM they apply, or of the potential benefits of MM for analysing and comprehending S-L. Such deficiencies can be observed in the need for a greater degree of integration between service phases, the way they are processed, and MM design. Learning should not only be produced in the classroom and should not be of use exclusively therein: S-L only makes sense if it is applied in the learning circumstances of real society. In the same manner, the use of MM cannot be truly grasped without ensuring a serious quantitative and qualitative methodological reflection that accompanies the service throughout its application (Pérez-Ordás et al. 2021).

This research is a methodological guide for professionals and academics who want to investigate MM in S-L, because it identifies methodological deficiencies and strengths, and offers alternative designs to evaluate the service. Researchers are advised to continue to use MM for S-L methodology, provided the methods in question are oriented towards the integration of quantitative and qualitative results, wherein the methodological complementarity

of the approach is made clear from the initial planning stage to the discussion and conclusions of the study.

One of the findings that emerge from our study, and which can provide perspectives and guidelines for S-L research using MM, is the observation that qualitative analysis, thanks to results obtained from experimental and quasi-experimental design, can provide insights that help us grasp the underlying mechanisms involved in service; qualitative analysis likewise helps us comprehend significant differences that are observed between pre-test and post-test scores, as well as new, unexpected categories that emerge as possible variables which could have an influence on service and are certainly worth further study. Another useful finding is the observation that descriptive statistical analysis favours data triangulation; qualitative analysis, for its part, complements quantitative analysis, which, in turn, facilitates an evaluation of what has occurred in the service process, i.e. how the diverse agents involved in S-L perceive the process, what relevance they attribute to it in a particular case, which elements emerge, and how the teacher and the students value the competential, conceptual, social and/or critical learning achieved with this methodology. S-L thus requires holistic methodological approaches to increase the problematization of the services to bring out any prevailing theoretical inconsistencies and triangulate or corroborate the results using different types of data.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11135-021-01218-3>.

Acknowledgements This research belongs to the Call for Projects of Learning Service at the Complutense University of Madrid (UCM) 2018–2019, by means of the *Moving Music Project, Moving through Music* (n° 17), carried out in Madrid Town Hall and sponsored by the Delegate of the Rector for Diversity and Environment of the UCM.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. No funding was received for conducting this study.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abel, C.F.: A Justification of the Philanthropic Model. In: Speck, S., Hoppe, Sh. (eds.) *Service-Learning: History, Theory, and Issues*, pp. 45–59. Praeger (2004)
- Anguera, M.T., Blanco-Villaseñor, A., Losada, J.L., Sánchez-Algarra, P., Onwuegbuzie, A.J.: Revisiting the difference between mixed methods and multimethods: is it all in the name? *Qual. Quant. Int. J. Methodol.* (2018). <https://doi.org/10.1007/s11135-018-0700-2>
- Aramburuzabala, P., Geruzaba, M., López-de-Arana, E.: Cómo abordar la evaluación en los proyectos de Aprendizaje-Servicio. In: Ruiz-Corbella, M., García-Gutiérrez, J. (eds.). *Aprendizaje-Servicio. Los retos de la evaluación*, pp. 27–38. Narcea, Madrid (2018)

- Bazeley, P.: Writing up multimethod and mixed methods research for diverse audiences. In: Hesse-Biber, S.N., Johnson, R.B. (eds.) *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*, pp. 296–313. Oxford University Press, Oxford (2015)
- Billig, S.H., Eyler, J. (eds.): *Deconstructing service-learning: Research exploring context, participation, and impacts*. Information Age (2003)
- Billig, S.H., Waterman, A. (eds.): *Studying Service-Learning: Innovation in Education Research Methodology*. Lawrence Erlbaum (2003)
- Bringle, R.G., Phillips, M.A., Hudson, M.: *The Measure of Service Learning: Research Scales to Assess Student Experiences*. American Psychological Association (2004)
- Bringle, R.G., Hatcher, J., Williams, M.: Quantitative approaches to research on international service learning. In: Bringle, R.G., Hatcher, J.A., Jones, S.G. (eds.) *International Service Learning: Conceptual Frameworks and Research*, pp. 275–291. Stylus, Virginia (2011)
- Brizee, A.: Toward participatory civic engagement: Findings and implications of a three-year community-based research study. *Comput. Compos.* (2014). <https://doi.org/10.1016/j.compcom.2014.04.007>
- Bryman, A.: Integrating quantitative and qualitative research: how is it done? *Qual. Res.* (2006). <https://doi.org/10.1177/1468794106058877>
- Bukas, V., Azean, N., Mohd, S., Tahir, L.A.: Systematic review of e-service learning in higher education. *Int. J. Interact. Mobile Technol.* (2020). <https://doi.org/10.3991/ijim.v14i06.13395>
- Butin, D.W.: *Service-Learning in theory and practice. The future of community engagement in higher education*. Palgrave Macmillan, USA (2010)
- Camilli, C., García, M., Galán, A.: ¿Es posible evaluar los resultados de proyectos de Aprendizaje-Servicio? In: Ruíz-Corbella, M., García-Gutiérrez, J. (eds.) *Aprendizaje-Servicio Los retos de la evaluación*, pp. 13–26. Narcea, Madrid (2018)
- Capella, C., Gil-Gómez, J., Chiva-Bartoll, O.: Innovative analysis of service-learning effects in physical education: a mixed methods approach. *J. Teach. Phys. Educ.* (2019). <https://doi.org/10.1123/jtpe.2019-0030>
- Castañer, M., Camerino, O., Anguera, M.T.: Métodos mixtos en la investigación de las ciencias de la actividad física y el deporte. *Apunts Educación Física y Deportes* **112**, 31–36 (2013)
- Celio, C., Durlak, J., Dymnicki, A.: A meta-analysis of the impact of service-learning on students. *J. Exp. Edu.* (2011). <https://doi.org/10.1177/105382591103400205>
- Chiva-Bartoll, O., Gil-Gomez, J., Zorrilla-Silvestre, L.: Improving the effective personality of pre-service teachers through service-learning: a physical education approach. *Rev. Inv. Educ.* (2019). <https://doi.org/10.6018/rie.37.2.303331>
- Clayton, P., Bringle, R., Hatcher, J. (eds.): *Research on service-learning. Conceptual frameworks and assessment*. Stylus (2013)
- Codispoti, F.R.: A justification of the communitarian model. In: Speck, B., Hoppe, S.H. (eds.) *Service-Learning: History, Theory, and Issues*, pp. 99–119. Praeger, London (2004)
- Cone, N.: A bridge to developing efficacious science teachers of All students: Community-based service-learning supplemented with explicit discussions and activities about diversity. *J. Sci. Teach. Educ.* (2009). <https://doi.org/10.1007/s10972-009-9137-8>
- Creswell, J., Plano, V.: *Designing and Conducting Mixed Methods Research*. SAGE (2017)
- Creswell, J., Plano, V., Gutmann, M., Hanson, W.: Advances in mixed methods research designs. In: Tashakkori, A., Teddlie, C. (eds.) *Handbook of Mixed Methods in Social and Behavioral Research*, pp. 209–240. SAGE (2013)
- Cumberland, D.M., Sawning, S., Church-Nally, M., Shaw, M. A., Branch, E., LaFaver, K.: Experiential learning: transforming theory into practice through the Parkinson's disease buddy program. *Teach. Learn. Med.* (2019). <https://doi.org/10.1080/10401334.2019.1580583>
- Deeley, S.J.: *Critical Perspectives on Service-Learning in Higher Education*. Palgrave MacMillan (2015)
- Driessnack, M., Sousa, V., Costa, L.: Revisión de los diseños de investigación relevantes para la enfermería: parte 3: métodos mixtos y múltiples. *Rev. Lat. Am. Enfermagem* (2007). <https://doi.org/10.1590/S0104-11692007000500025>
- Eldh, A.C., Arestedt, L., Berterö, C.: Quotations in qualitative studies: reflections on constituents, custom, and purpose. *Int. J. Qualit. Methods* (2020). <https://doi.org/10.1177/1609406920969268>
- Eyler, J.: What international service-learning research can learn from research on international learning. In: Bringle, R.G., Hatcher, J.A., Jones, S.G. (eds.) *International Service Learning: Conceptual Frameworks and Research*, pp. 225–243. Stylus, Virginia (2011)
- Folgueiras, P., Aramburuzabala, P., Opazo, H., Mugarra, A., Ruiz, A.: Service-learning: a survey of experiences in Spain. *Edu. Citizsh. Social Justice* (2018). <https://doi.org/10.1177/1746197918803857>
- Frels, R., Onwuegbuzie, A.: Administering quantitative instruments with qualitative interviews: a mixed research approach. *J. Couns. Dev.* (2013). <https://doi.org/10.1002/j.1556-6676.2013.00085.x>





- Furco, A.: Issues of definition and program diversity in the study of Service-Learning. In: Billig, S., Waterman, A. (eds.) *Studying Service-Learning: Innovation in Education Research Methodology*, pp. 13–33. Lawrence Erlbaum, New York (2003)
- García, D., Lalueza, J.L.: Procesos de aprendizaje e identidad en aprendizaje-servicio universitario: una revisión teórica. *Educación XXI* (2019). <https://doi.org/10.5944/educxx1.22716>
- Gerholz, K.H., Liszt, V., Klingsleck, K.: Effects of learning design patterns in service learning courses. *Act. Learn. High. Educ.* (2017). <https://doi.org/10.1177/1469787417721420>
- Gezuraga, M., García, Á.: Recepciones de la pedagogía experiencial de Dewey en diversos enfoques metodológicos: el valor añadido del aprendizaje-servicio. *Educatio Siglo XXI* (2020). <https://doi.org/10.6018/educatio.452921>
- Gibbs, G.R.: *Analyzing qualitative data*. SAGE (2018)
- Gil-Gómez, J., Moliner-García, O., Chiva-Bartoll, García López, R.: A service-learning experience in future teachers: development of the social and civic competence. *Rev. Complutense Educ.* (2016)
- Greene, J., Caracelli, V.J., Graham, W.: *Toward a Conceptual Framework for Mixed-Method. Evaluat. Des. Edu. Evaluat. Policy Anal.* (1988). <https://doi.org/10.3102/01623737011003255>
- Greene, J.C.: *Mixed methods in social inquiry*. Jossey-Bass Ed (2007)
- Greene, J., Caracelli, V.J., Graham, W.: *Toward a Conceptual Framework for Mixed-Method. Evaluation Designs Educational Evaluation and Policy Analysis* (1989). <https://doi.org/10.3102/01623737011003255>
- Hamui-Sutton, A.: *Un acercamiento a los métodos mixtos de investigación en educación médica. Investigación En Educación Médica* (2013). [https://doi.org/10.1016/S2007-5057\(13\)72714-5](https://doi.org/10.1016/S2007-5057(13)72714-5)
- Hanson, W., Creswell, J.W., Plano, V., Petska, K., Creswell, J.D.: Mixed methods research designs in counseling psychology. *J. Couns. Psychol.* (2005). <https://doi.org/10.1037/0022-0167.52.2.224>
- Hayward, L., Li, L.: Sustaining and improving an international service-learning partnership: evaluation of an evidence-based service delivery model. *Physiother. Theory Pract.* (2017). <https://doi.org/10.1080/09593985.2017.1318425>
- HEQC (Higher Education Quality Committee): a good practice guide for self-evaluation instruments for managing the quality of Service-Learning. Council on Higher Education. <https://www.ufs.ac.za/docs/librariesprovider43/service-learning-documents/articles-documents/heqc-good-practice-guide-677-eng.pdf?Status=Master&sfvrsn=0> (2006). Accessed 30 July 2020
- Hesse-Biber, S.N.: Introduction: navigating a turbulent research landscape: working the boundaries, tensions, diversity, and contradictions of multimethod and mixed methods inquiry. In: Hesse-Biber, S.N., Johnson, B. (eds.) *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*, pp. XXXIII–LIII. Oxford University Press (2016)
- Hirschinger-Blank, N.B., Simons, L., Kenyon, A.: An evaluation of a service-learning model for criminal justice undergraduate students. *J. Exp. Educ.* (2009). <https://doi.org/10.1177/105382590903200106>
- Hopia, H., Latvala, E., Liimatainen, L.: Reviewing the methodology of an integrative review. *Scand. J. Car. ing Sci.* (2016). <https://doi.org/10.1111/scs.12327>
- Huff, J.L., Zoltowski, C.B., Oakes, W C.: Preparing engineers for the workplace through service learning: perceptions of EPICS alumni. *J. Eng. Educ.* (2016). <https://doi.org/10.1002/jee.20107>
- Hutton, B., Catalá-López, F., Moher, D.: La extensión de la declaración PRISMA para revisiones sistemáticas que incorporaran metaanálisis en red: PRISMA-NMA. *Med. Clin.* (2016). <https://doi.org/10.1016/j.medcli.2016.02.025>
- Johnson, B., Onwuegbuzie, A.: Mixed methods research: a research paradigm whose time has come. *Educ. Res.* (2004). <https://doi.org/10.3102/0013189X033007014>
- Kiely, R., Hartman, E.: Qualitative research methodology and international service learning. In: Bringle, R.G., Hatcher, J.A., Jones, S.G. (eds.) *International Service Learning: Conceptual Frameworks and Research*, pp. 291–317. Stylus, Virginia (2011)
- Kuntjara, E.H.: Students' reflection on their service-learning experience as a way of fostering critical thinking and as a peace building initiative. *Citizensh. Teach. Learn.* (2019). https://doi.org/10.1386/ctl_00007_1
- Lee, J., Chang, S.H., Haegele, J.A.: Does service-learning work? Attitudes of college students toward individuals with disabilities. *J. Exp. Edu.* (2020). <https://doi.org/10.1177/1053825920918887>
- Lewellyn, A., Kiser, P.M.: Conceptualizing critically as a guiding principle for high quality academic service learning. *Int. J. Teach. Learn. Higher Edu.* **26**(1), 147–156 (2014)
- Li, Y., Yao, M., Song, F., Fu, J., Chen, X.: Building a just world: the effects of service-learning on social justice beliefs of Chinese college students. *Educational Psychology. Int. J. Exp. Educ. Psychol.* (2019). <https://doi.org/10.1080/01443410.2018.1530733>

- Lorenzo-Moledo, M.D., Sáez-Gambín, D., Ferraces, M.J., Varela, C.: Reflection and quality assessment in service-learning projects: when with whom and why. *Front. Edu.* (2021). <https://doi.org/10.3389/educ.2020.605099>
- Lotti, P., Betti, Fr.: Aprendizaje-Servicio al servicio de la comunidad educadora. *Revista Iberoamericana de Aprendizaje-Servicio* (2019). <https://doi.org/10.1344/RIDAS2019.7.5>
- Lucero, J., Wallerstein, N., Duran, B., Alegria, M., Greene-Moton, E., Israel, B., Kastelic, S., Magarati, M., Oetzel, J., Pearson, C., Schulz, A., Villegas, M., White Hat, E.R.: Development of a mixed methods investigation of process and outcomes of community-based participatory research. *J. Mixed Methods Res.* (2018). <https://doi.org/10.1177/1558689816633309>
- Mampane, R., Huddle, C.: Assessing the outcomes of school-based partnership resilience intervention. *S. Afr. J. Educ.* (2017). <https://doi.org/10.15700/saje.v37n1a1412>
- Marco-Gardoqui, M., Eizaguirre, A., García-Feijoo, M.: The impact of service-learning methodology on business schools' students worldwide: a systematic literature review. *PLoS ONE* (2020). <https://doi.org/10.1371/journal.pone.0244389>
- Martínez-Usarralde, M.J., Gil-Salom, D., Macías-Mendoza, D.: Revisión sistemática de Responsabilidad Social Universitaria y Aprendizaje Servicio: Análisis para su institucionalización. *Revista Mexicana de Investigación Educativa* **80**, 149–172 (2019)
- Maxwell, J.: Expanding the history and range of mixed methods research. *J. Mixed Methods Res.* (2016). <https://doi.org/10.1177/1558689815571132>
- McWhorter, R.R., Delello, J.A., Roberts, P.B.: Giving back: exploring service-learning in an online learning environment. *J. Interact. Online Learn.* **14**(2), 80–99 (2016)
- Miles, M.B., Huberman, A.M.: *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd end. SAGE (1994)
- Molina-Azorin, J.F.: Mixed methods research: an opportunity to improve our studies and our research skills. *Eur. J. Manag. Bus. Econ.* (2016). <https://doi.org/10.1016/j.redeen.2016.05.001>
- Moorkens, J.: Consistency in translation memory corpora: a mixed methods case study. *J. Mixed Methods Res.* (2015). <https://doi.org/10.1177/1558689813508226>
- Núñez, J.: Los métodos mixtos en la investigación en educación: hacia un uso reflexivo. *Cuadernos De Pesquisa* (2017). <https://doi.org/10.1590/198053143763>
- Ortega, P.: La pedagogía crítica: Reflexiones en torno a sus prácticas y desafíos. *Pedagogía y Saberes* (2009). <https://doi.org/10.17227/01212494.31pys26.33>
- Packard, K., Ryan-Haddad, A., Monaghan, M.S., Doll, J., Qi, Y.: Application of validated instruments to assess university-wide interprofessional service-learning experiences. *J. Interprof. Educ. Pract.* (2016). <https://doi.org/10.1016/j.xjep.2016.06.005>
- Pereira, Z.: Los diseños de método mixto en la investigación en educación. Una experiencia concreta. *Revista Electrónica Educare* (2011). <https://doi.org/10.15359/ree.15-1.2>
- Pérez-Ordás, R., Nuviala, A., Grao-Cruces, A., Fernández-Martínez, A.: Implementing service-learning programs in physical education: teacher education as teaching and learning models for all the agents involved: a systematic review. *Int. J. Environ. Res. Public Heath* (2021). <https://doi.org/10.3390/ijerph18020669>
- Ponce, O.: Investigación de métodos mixtos en educación. Publicaciones Puertorriqueñas, San Juan (2011)
- Ponce, O.A., Pagán-Maldonado, N.: Mixed methods research in education: capturing the complexity of the profession. *Int. J. Edu. Excell.* (2015). <https://doi.org/10.18562/IJEE.2015.0005>
- Puig, M.P., Battle, R., Bosch, C., Palos, J.: Aprendizaje Servicio. Educar para la ciudadanía. Octaedro (2007)
- Puig, J.M.: Aprendizaje Servicio. Educación y compromiso cívico. Graó (2009)
- Punch, K.F.: *Introduction to Research Methods Education*. SAGE (2014)
- Redondo-Corcobado, P., Fuentes, J.L.: La investigación sobre el Aprendizaje-Servicio en la producción científica española: una revisión sistemática. *Revista Complutense De Educación* (2020). <https://doi.org/10.5209/rced.61836>
- Salam, M., Awang, D.N., Ibrahim, D.H.A., Farooq, M.S.: Service learning in higher education: A systematic literature review. *Asia Pac. Educ. Rev* (2019). <https://doi.org/10.1007/s12564-019-09580-6>
- Salvador-García, C., Santágueda-Villanueva, M., Ruiz-Montero, P.J.: Posibilidades de los métodos cuantitativos y mixtos en la investigación del Aprendizaje-Servicio. In: Chiva-Bartoll, O., Gil, J. (eds.). *Aprendizaje-servicio en la formación inicial docente*, pp. 107–118. Octaedro, Barcelona (2018).
- Schoonenboom, J., Johnson, R.B.: How to construct a mixed methods research design. *Köln Zeitschrift Für Soziologie Und Sozialpsychologie* (2017). <https://doi.org/10.1007/s11577-017-0454-1>
- Sementelli, A.: A critique of the philanthropic model. In: Speck, B., Hoppe, S.H. (eds.) *Service-Learning: History, Theory, and Issues*, pp. 45–59. Praeger, London (2004)
- Speck, B., Hoppe, S. (eds.): *Service-Learning: History, Theory and Issues*. Praeger (2004)

- Tashakkori, A., Teddlie, C.: Handbook of Mixed Methods in Social and Behavioral Research. SAGE (2010)
- Wang, C., Yan, W., Guo, F., Li, Y., Yao, M.: Service-learning and chinese college students' knowledge transfer development. *Front. Psychol.* (2020). <https://doi.org/10.3389/fpsyg.2020.606334>
- Waters, S., Anderson-Lain, K.: Assessing the student, faculty, and community partner in academic Service-Learning. *J. High. Educ. Outreach Engagem.* **18**(1), 89–122 (2014)
- World Bank: World Bank Country and Lending Groups. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> (n.d.) Accessed 10 August 2020

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Celia Camilli Trujillo¹  · Laura Cuervo Calvo²  · Desirée García Gil²  ·
Carolina Bonastre Valles³ 

Laura Cuervo Calvo
lcuervo@ucm.es

Desirée García Gil
desiree.garcia@edu.ucm.es

Carolina Bonastre Valles
carolina.bonastre@uam.es

- ¹ Department of Research and Psychology in Education, Education Faculty, Complutense University of Madrid (UCM), Madrid, España
- ² Department of Didactics of Languages, Arts and Physical Education, Education Faculty, Complutense University of Madrid (UCM), Madrid, España
- ³ Interfacultative Music Department, Faculty of Teacher Learning and Education, Autónoma University of Madrid (UAM), Madrid, España