

# **What is Evidence for Engagement in Argumentation in Qualitative Studies? From Early Childhood Education to Teacher Education**

María Pilar Jiménez-Aleixandre  
Universidade de Santiago de Compostela, Spain  
marilarj.aleixandre@usc.es

## **Abstract:**

This paper examines the processes of gathering and analysing evidence for engagement in argumentation in qualitative studies in science education. The overarching question is: What is empirical evidence about engagement in argumentation in qualitative studies and how can modifications in the design process contribute to identifying it? There is an underlying concern about the coherence among research questions, research design, construction of rubrics and tools for analysis, findings and conclusions; in other words about sound methodological decisions. We suggest that in studies about practices and processes the evidence should be about participants' performances, rather than about learning products. The paper's rationale draws from the sociocultural perspective. Processes, and in particular changes, involved in research studies and papers, such as refining research questions; modifying appropriate designs to gather data; elaborating specific tools to analyse particular dimensions of argumentation; and drawing conclusions grounded in evidence, are discussed, with examples from two studies, in Early Childhood Education (Monteira & Jiménez-Aleixandre, 2016) and in Teacher Education (Jiménez-Aleixandre & Brocos, 2017). Implications for research and school practice are discussed.

**Keywords:** research evidence; argumentation; methodological issues; qualitative studies.

## **Introduction: What is Evidence for Engagement in Argumentation?**

This paper examines the processes of gathering and analysing evidence for engagement in argumentation in qualitative studies. The focus is on methodological issues related to classroom studies. Our argument is that methodological issues are relevant and should perhaps command more attention, in papers and conferences, from the science education research community. The overarching question is:

*What is empirical evidence about engagement in argumentation in qualitative studies and how can modifications in the design process contribute to identifying it?*

Underlying the question, there are two concerns, the first one about the type of evidence that is appropriate for qualitative studies; the second one about the coherence among the different elements in a study: research questions, research design, construction of rubrics and tools for analysis, findings and conclusions; in other words about methodological decisions. The question is further broken down in the examination of decisions related to modifications in these steps: 1) asking and refining research questions about argumentation; 2) planning and modifying appropriate designs to gather data that constitute quality evidence; 3) elaborating specific tools to analyse particular dimensions of argumentation; and 4) drawing conclusions grounded in evidence. These issues are discussed by debriefing the processes leading to quality research studies and papers about argumentation, through a reflection on the decisions made while carrying out, and reporting about, two studies whose quality is contrasted by their publication in: a) the *Journal of Research of Science Teaching*, a high impact journal indexed in the web of science (Monteira & Jiménez-Aleixandre, 2016), its content is argumentation in Early Childhood Education; and b) a chapter of a Springer volume about interpersonal argumentation (Jiménez-Aleixandre & Brocos, 2017), it is a study about the processes of negotiation involved in building a shared argument in a decision-making context about vegetarian or non-vegetarian diets. The first section discusses the rationale and the issue of evidence in qualitative studies; then the decisions related to asking questions, planning research design, elaborating rubrics and drawing conclusions are discussed.

### **1. Rationale: Warranted Accounts and Evidence in Qualitative Studies**

Educational research belongs to social sciences; therefore the type of evidence necessary in order to support claims may be different from the evidence in natural sciences. However, the background for a majority of science education researchers is scientific, belonging to natural sciences, and this influences our frameworks, how we think about or how we understand research, even when the focus of this research is on education, not on physics, chemistry or biology.

There is consensus on the relevance of methodological issues for science education research. We, as a field, need to produce research that is rigorous. And a reason of a practical order is that, in the process of publishing, major reasons for rejection are poor research design and flaws in methods. But, although methodology matters, science education research has tended to have a stronger focus either on theoretical approaches or on empirical findings about learning and teaching (Jiménez-Aleixandre, 2019). As a consequence, methodological issues specific of the science education field have been given less attention. A great deal of studies about science education draw from domain-general research methods. Certainly, we have much to learn from them, but they may be less suited to address some domain-specific issues. A recent volume about theoretical and methodological issues in sociocultural research in science and engineering education, edited by Gregory Kelly and Judith Green (2019a) seeks to fill this gap, addressing methodological challenges related to these specific fields.

A central issue in methodological debates, according to Heap (1995) “is the concept of science under which research is carried out and in terms of which is justified” (Heap, 1995; p. 271); this points towards the distinction between natural sciences and social and human sciences, concerned with inquiry into human actions and experiences, as science education research. Heap (1995) argues that in qualitative educational research claims are not necessarily the standard type of empirical claims. For social and cultural inquiry employing observational methods he discusses two types of non-empirical claims, logical and normative. Logical claims are defined in terms of the necessary conceptual relations in a proposition, for instance characterizing “appropriate participation” in a lesson as “actively interpreting the meaning being signaled”. Normative claims are those related to contingent relations, what ought to be, or what should be related to what. For Heap an example of normative claim is the influence of the teacher’s purpose for a lesson in the type of classroom discussion. This claim could be tested empirically or normatively, by considering that purpose and discussion are conceptually and normatively interlinked. Thus Heap introduces the notion of “rationally accountable, domain-oriented inquiry” (p. 271) in order to guide qualitative studies.

The meaning of “rationally accountable” is developed in Kelly and Green (2019b) introductory chapter to their edited volume. They characterize research studies framed in a sociocultural perspective as “warranted accounts”: “Our goal (...) is to make visible ways of developing and explaining processes involved in designing, conducting, and constructing research leading to *warranted accounts* of educational phenomena (Heap, 1995).” (Kelly & Green, 2019b, p. 1, my emphasis). This volume brings a unique approach to methodological reviews about science education, by focusing on reflections about the process of research itself. Each chapter discusses, beyond the final form of research questions, rubrics or tools for analysis, the decisions made during the process. Thus, what Kelly and Green seek to make visible is how, for what purposes or with what outcomes, science is socially, discursively, and conceptually constructed in educational settings. The meaning of “warranted accounts” may be related to Mitchell’s (1984) “telling cases”, detailed presentations of ethnographic data that make visible previously obscured or unknown information. Likewise, Erickson (2012) states that in qualitative research “The emphasis is on discovering *kinds* of things that make a difference in social life; hence, an emphasis is placed on *qualitas* rather than on *quantitas*.” (Erickson, 2012, p. 1451; author’s emphasis).

It should be noted that qualitative and quantitative approaches should not be viewed as discrete categories or dichotomies, but rather as the ends on a continuum (Creswell, 2014), with mixed methods in the middle. The differences are in the emphasis. For Erickson (2012) the priority of emphasis on quality “does not mean that information about frequency is irrelevant to qualitative inquiry” (Erickson, 2012, p. 1451). Large-scale quantitative studies are appropriate to document products of learning or the “current state”, seeking generalizability. However, as Kelly and Green (2019b) suggest, they may lack the sensitivity to examine how access to knowledge and learning are constructed through social interaction. Therefore, they may fail to explain why or how that current state was constructed.

Qualitative research seeks to explore and understand the meanings individuals or groups ascribe to social or human problems and issues (Creswell, 2014). In education and science education, qualitative studies emphasize the socially constructed nature of events; they seek to analyse processes of knowledge construction. Erickson (2012) claims that it is the responsibility of qualitative researchers to go beyond what the actors understand explicitly, identifying the meanings that are outside their awareness. However, as Kelly and Green (2019b) point out, some qualitative studies in science classrooms do not make use of systematic and theoretically grounded approaches, and they suggest the importance of utilizing large-scale video data sets.

The focus of this paper is on specific research methods; although our approach belongs to qualitative and mixed methods studies they may be of application, in many cases to quantitative studies. The two studies from our research group analysed in this paper were chosen because its quality is evidenced by their publication, in the first case in a journal indexed in the Journal Citation Report of the Social Sciences Citation Index (JCR of SSCI); in the second in chapter by a high impact publisher, Springer:

– Monteiro, S. F., & Jiménez-Aleixandre, M. P. (2016). The practice of using evidence in kindergarten: The role of purposeful observation. *Journal of Research in Science Teaching*, 53(8), 1232–1258.

The paper is part of a longitudinal study examining kindergarteners' engagement in scientific practices during three years, from 3–4 to 5–6 years of age. Its focus is on generating and using evidence to support claims, and on learning about snails' biology during a 5-month project on snails in Early Childhood Education. The *Journal of Research in Science Teaching* (JRST) is one of the three science education journals with highest impact factor, over 3; currently 3.210; the other two being *Studies in Science Education* (3.455) and *Science Education* (3.035). This paper was also awarded the *Research worth reading* distinction in 2017, by a joint committee NARST / NSTA (National Association of Research in Science Teaching/ National Science Teachers Association), for its potential "to be most relevant and valuable for practitioners".

– Jiménez-Aleixandre, M. P. & Brocos, P. (2017). Processes of negotiation in socio-scientific argumentation about vegetarianism in teacher education. In F. Arcidiacono & A. Bova (Eds.) *Interpersonal argumentation in educational and professional contexts* (pp. 117–139). Dordrecht: Springer.

This chapter is part of the thesis of Pablo Brocos, about the development of criteria in argumentation about sustainable and healthy diets, comparing two case studies in teacher education and in high school, with a focus on non-structural aspects of argumentation. The chapter examines the processes of negotiation about the choice of diet to be agreed by the group (vegetarian, vegan, omnivorous), and about the evidence and justification employed to support it. Springer is one of the major international publishers and all their books go through a peer-review process.

## **2. The Process of Refining Research Questions**

According to the French philosopher Gaston Bachelard (1938) all knowledge consists of answers to questions. Research reports also consist of answers to questions, therefore good questions are at the heart of inquiry (Erickson, 2012). The research questions (RQs) should not be viewed as fixed and immovable; sometimes they need to be refined, revised or even changed. Erickson claims that in a qualitative study changing RQs “means that the post hoc analysis is working properly – discovering subtleties and contingencies that could not have been foreseen” (Erickson, 2012, p. 1462), and he further suggests that this is the goal of participant observational fieldwork, in order to discover what could not be anticipated.

So a first step when undertaking a research is to ask robust research questions that may be answered empirically, and not simply be answered yes or no.

We may compare in Monteiro and Jiménez-Aleixandre's (2016) paper the first and last formulations of the RQs; initially they were:

“(1) To examine the meanings for what constitutes evidence constructed during the project. (2) To examine how children use evidence to evaluate and revise knowledge, in other words the features of kindergarten children's engagement in the practice of using evidence.”

Both were worded as objectives, and a first change was to transform them in questions. The first one was too general; following the referees' suggestions it was developed into two specific questions, in reference to the development of data into evidence, an original contribution of the study; the second was also refined and broken down into two, including in the new RQ2 about the ways of gathering empirical evidence. Thus the new RQs in the published paper are:

RQ1 What meanings do kindergarteners construct for what constitutes evidence? How are those meanings reflected in the development of data into evidence?

RQ2 Which ways of gathering empirical evidence are jointly constructed by children and teacher during the project?

RQ3 How do children use evidence to revise their understandings?

In Jiménez-Aleixandre and Brocos (2017) the draft formulations of the research objectives were:

“(1) To examine the processes of negotiation about the choice or option to be agreed by the group (vegetarian, vegan, omnivorous)

(2) To examine the evidence and justification to be employed to support the option;

(3) To examine the quality of the resulting argument and the strategies for building and defending it, i.e., issues pertaining to metaknowledge.”

These were in the format of objectives, and it was necessary to formulate them as questions. When we began to write the chapter, it became clear that it was lacking a clear focus, addressing too many issues. Objective 1 focused on the option (diet) and objective 2 in argument components as evidence and justifications; however from the analysis it became clear, first that there were other objects negotiated (*negotia*), such as task goals, meaning of notions and terms, or strategies; second that there were non-structural dimensions and elements to be considered, such as the relative weight of evidence and values. Thus both objectives were collapsed in the new RQ1. Objective 3 addressed two different issues, difficult to explore deeply in 10 000 words, so the new sharpened focus (now RQ2) was on strategies, negotiation levels, an original contribution from the study, and negotiation paths:

Overarching question: to explore the processes of negotiation involved in building a shared argument in a decision-making context about vegetarian or non-vegetarian diets.

RQ1 Which dimensions have greater weight in the negotiation process and in the final decision?

RQ2 Which patterns, in terms of strategies and negotiation levels, reveal the negotiation paths in four small groups?

In both cases the first versions of the RQs exemplify common problems in formulating questions, being too general, lacking a sharp focus or failing to capture the specificities of the issue studied.

### **3. Modifying Research Designs and Data Collection**

The issue of design research in order to gather data that constitute quality evidence for the question examined needs to be coherent with the RQ, to ascertain that it fits the problem. Erickson (2012) characterizes it as where to be as a researcher, with whom, and how. Spindler and Spindler (1992), discussing ethnographic studies in education, suggest that observation should be direct, prolonged and *in situ*. Direct because, besides using recording devices, the researcher should be in the setting where the action is occurring; prolonged and *in situ* because “the *validity* of ethnographic observation is based in observation *in situ* lasting

long enough to permit the ethnographer to see things happening not once, but repeatedly.” (Spindler & Spindler, 1992, p. 65, authors’ emphasis). The design involves identifying relevant data sources to answer the RQs, to warrant the claims that one would like to be able to make (Erickson, 2012).

The full study in Early Childhood Education (ECE) had the purpose of identifying progression in the engagement in scientific practices along time, hence it was designed as a longitudinal study, lasting three school years –the full duration of ECE. The study reported in the paper, with 25 children (5–6 years of age) and their teacher, involved participant observation during the 5-months project. Data collection included video recording of 9 sessions (5 hours), field notes and children’s drawings (N=190). Designing a range of data collection was needed to provide evidence of participants’ performances, in other words actions and processes, thus for each RQ:

For RQ1 data about the process of transformation of raw data into evidence: the sources are the pupil’s conversations during the nine sessions, and their representations in the 190 drawings. Additionally it was necessary to develop criteria for what is “evidence” at this age.

For RQ2 data about specific ways of how children gather empirical, or first hand evidence: the sources are the video-recordings and the field notes.

For RQ3 about how children used evidence to revise and change their ideas: the sources are drawings about the same topic along the project, their revisions of the previous drawings, prompted by the teacher, and the recorded classroom discourse.

The modifications during the three school years related for instance to data collection, for in year 1 the teachers did not understand that we needed to be there in all the sessions related to the project, which sometimes were not scheduled. In the second and third year the researcher (first author) attended almost all sessions.

The full study on argumentation about sustainable diets is framed in design-based research (DBR) (Barab & Squire, 2004). It comprises three cycles: two carried out in teacher education –first case study– and the third, a full school year project, in a high school –second case study. The results of each cycle were used to modify the design of the next iterations. The study reported in the chapter, the first cycle, involved 85 pre-service primary teachers enrolled in the science education course taught by the first author, which involved six argumentation tasks, from which the last is discussed in the chapter, and three tasks about nutrition. They worked in small groups (N=20) and they were asked to construct an argument about which diet would be more sustainable and healthy. Conflict was embedded in the task design, for information pointed both to benefits and drawbacks of every option.

For both research questions the data sources are students’ written products (individual pre-test, portfolios, final essays; group arguments) and video recording of small groups from which all participants agreed to it. The chapter focuses on the written arguments of the 20 small groups and the oral debates of four groups, one from each seminar session.

The modifications, according to DBR, affected to the second and third cycle. For instance, the results of the first iteration, discussed in the chapter, point to assertive strategies in some groups, deciding the option before evaluating the evidence. Another result was that values had relevance in the arguments, but they were addressed implicitly. The design of the task was modified including scaffolding for the evaluation of evidence and for explicit consideration of values.

#### **4. Constructing specific analytical tools to analyse dimensions of argumentation**

The elaboration of specific tools in order to analyse different dimensions of argumentation is a crucial step. Due to space constraints only analytical tools related to one question of each study are discussed, and we refer to the publications for more details.

In the ECE study, for RQ1 about the development of data into evidence, data reduction proceeded through discourse analysis (Gee, 2005) and the use of the constant comparative method (Glaser & Strauss, 1967). The codes emerged from the interaction between the literature and data in successive iterations, and the construction of tools, extensively discussed in the paper, cannot be reproduced. We will focus on how we established two levels and two sets of coding categories in the analysis: first, identifying the 276 argument components in the conversations, and distinguishing between claims (125), evidence (57) and justifications (9). It was necessary to create a new category that we called *raw data* (85), and defined it as a description unrelated to a claim or question. This is related to the finding that not all raw data develop into evidence. For the purposes of a study in the ECE context, we consider evidence data whose discursive role is to support a claim. The second level of analysis focused only on the 57 evidence statements, with the purpose of identifying levels of *epistemic judgment*, which would warrant the paper claims about how data are transformed into evidence. Within these 57 evidence statements two levels were identified, L1 (N=36), closer to data, as in “(snails have) heart, brain”, which supports the claim “they are almost like us”; and L2 (N21), involving evaluative judgments, such as to identify patterns, to connect data to claims, to compare or to evaluate alternatives.

In the vegetarianism dilemma study for part of RQ2, about *negotiation levels*, a notion and coding elaborated by the authors, which represents the degree of interaction, the engagement with alternative ideas or positions, drawing from Bakhtin's (1986) idea of taking into account more than one viewpoint, and engaging with the interlocutors' statements. The codes emerged from the interaction between the literature and the study data. Two broad categories were identified in the data, depending whether alternative *negotia* (options, evidence, meanings, etc.) were or were not considered. The lower levels correspond to episodes in which: L0, *negotia* were not discussed and the agreements were implicit, and L1, *negotia* were explicitly accepted, but not compared with alternatives. The higher levels correspond to episodes in which: L2 alternative *negotia* were discussed but negotiation was unresolved; and L3 alternative *negotia* were discussed until reaching agreement.

## **5 Presenting results and conclusions grounded in evidence**

A cause of concern in writing research reports is the coherence across the manuscript among research questions, data sources, reported findings and conclusions. Results and conclusions must be supported in data, grounded in evidence. The steps from data collection to data reduction to identifying patterns in data are complex; they require experience and guidance from experts. We discuss part of the results from only one RQ of each study.

In the ECE study, results for RQ3 about how children used evidence to revise their understandings were derived, first from a thematic analysis of the 21 conceptual topics about snails emerging in the data; second, 8 topics were identified recurring through two, three or five sessions; third, from these eight, the question of mouthparts, discussed in five sessions, was selected. For, as Spindler and Spindler (1992) point out, it is necessary to see things happening repeatedly. The study documents the revision of children's ideas about snails' mouthparts during the five months, from drawing anthropomorphic “teeth” and “tongue” to several cycles of revision based on three types of evidence, gathered respectively through purposeful observation, for instance of marks in food, experiments, and secondary sources such as watching videos or searching the Internet, resulting in a proposal about the radula's mechanism. The teacher's scaffolding, prompting reflection, thinking back about their

observations and revision of previous models, was crucial. An original theoretical contribution of the paper is the notion of *purposeful observation*: one that is prolonged, systematic, with a clear focus, discussed and used to test claims and compare initial ideas with later ones

In the teacher education study, results for part of RQ2 about patterns of strategies identified in the negotiation, the comparison of the process in the four small groups across five dimensions: 1) the departure point of the negotiation (Baker, 1994) from initial agreement to explicit opposed alternatives; 2) the negotiation *strategy*, either *assertive*, with a choice was first decided, and then justified with the data, or *exploratory*, with participants engaging in the evaluation of evidence before deciding on an option; 3) the *consensus* reached or the absence of it; 4) *conflict management*; 5) Which dimension had greater *weight* in the final decision. An issue emerging from the comparison is the influence of the interactions, or how did the appropriation of contributions from others influence the negotiation path.

The issue of discussion cannot be addressed for lack of space; but it is important to recall that findings are descriptive and analytical, while the discussion is evaluative and interpretative.

### **By way of conclusion: Evidence about Argumentation Processes**

We suggest that in studies about practices, for instance about scientific practices as argumentation or modelling, and about processes, long-term projects, as the ones reported in the papers, have specific affordances. In particular, for classroom studies about what students do, for instance about practices or processes, appropriate data refer to performances of actions and to discourse, rather than responses to questionnaires. In other words, research evidence should be about participants' performances, rather than about learning products.

The discussion about methodological steps is framed in concerns about what counts as valid research in science education. The analysis of research is illustrated through a reflection on the processes, and in particular on the changes, involved in two studies about argumentation in different ages and contexts. The reflection on the decisions highlights false starts, as initial research questions that needed revisions; in these two studies involving new questions much more specific, and breaking down themes. It shows that qualitative or mixed methods approaches can yield quality evidence to support claims. For both studies the suggestions of the anonymous referees and editors of the *Journal of Research in Science Teaching* and of the book editors, Francesco Arcidiacono and Antonio Bova, were of great help in sharpening the focus, directing us towards making data analysis more explicit and, in general, improving the quality of the publication. This points out to the role of referees – sometimes considered by beginning researchers as feared enemies – in supporting the community in publishing quality studies.

### **Acknowledgements:**

Work supported by the Spanish Ministerio de Economía y Competitividad (MINECO), contract grant number: EDU2015–66643-C2–2-P. Thanks to my doctoral students and co-authors of the studies analysed, Sabela F. Monteiro and Pablo Brocos.

## **REFERENCES**

- Bachelard, G. (1938). *La formation de l'esprit scientifique*. Paris: Vrin.
- Baker, M. (1994). A model for negotiation in teaching-learning dialogues. *Journal of Artificial Intelligence in Education*, 5(2), 199–254.



- Bakhtin, M. M. (1986). *Speech genre and other late essays*. Austin: University of Texas Press.
- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The journal of the learning sciences*, 13(1), 1–14.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches*. Thousand Oaks, CA: Sage [1st Edition, 2003].
- Erickson, F. (2012). Qualitative research methods for science education. In B. Fraser, K. Tobin & C. McRobbie (Eds.) *Second international handbook for Science Education*. Volume 2 (pp. 1451-1469). Dordrecht: Springer.
- Gee, J. P. (2005). *An introduction to discourse analysis: Theory and methods*. London: Routledge.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: strategies for qualitative research*. London: Weidenfeld and Nicholson.
- Heap, J. L. (1995). The status of claims in “qualitative” educational research. *Curriculum Inquiry*, 25(3), 271-292.
- Jiménez-Aleixandre, M. P. (2019). Foreword: Meeting methodological challenges. In G. J. Kelly & J. Green (Eds.) *Theory and methods for sociocultural research in science and engineering education* (pp. xiv–xviii). New York: Routledge.
- Jiménez- Aleixandre, M. P. & Brocos, P. (2017). Processes of negotiation in socio-scientific argumentation about vegetarianism in teacher education. In F. Arcidiacono & A. Bova (Eds.) *Interpersonal argumentation in educational and professional contexts* (pp. 117–139). Dordrecht: Springer.
- Kelly, G. J. & Green, J. (Eds.) (2019a). *Theory and methods for sociocultural research in science and engineering education*. New York: Routledge.
- Kelly, G. J. & Green, J. (2019b). Framing issues of theory and methods for the study of science and engineering education. In G. J. Kelly & J. Green (Eds.) *Theory and methods for sociocultural research in science and engineering education* (pp. 1–28). New York: Routledge.
- Mitchell, J. (1984). Typicality and the case study. In R. F. Ellen (Ed.) *Ethnographic research: A guide to general conduct*. (pp. 238-241). Orlando, FL: Academic Press.
- Monteira, S. F., & Jiménez-Aleixandre, M. P. (2016). The practice of using evidence in kindergarten: The role of purposeful observation. *Journal of Research in Science Teaching*, 53(8), 1232–1258. DOI10.1002/tea.21259.
- Spindler, G., & Spindler, L. (1992). Cultural process and ethnography: An anthropological perspective. In LeCompte, M.D., Millroy, W.L. & Preissle, J. (Eds) *The Handbook of Qualitative Research in Education* (pp. 53–92). San Diego: Academic Press.