Concept of evidence and the quality of evidence-based reasoning in elementary students

Andrea Miralda-Banda[a], Merce Garcia-Mila[a] and Mark Felton[b]

[a] *Department of Cognition, Development and Educational Psychology. University of Barcelona, Barcelona, Spain.*

[b] *Department of Teacher Education. San Jose State University, CA, USA.*

Please address correspondence to:

Andrea Miralda-Banda

andreamiralda@gmail.com

+34933125833

Pg.Vall d'Hebron, 171

08035 Barcelona, Spain.

ORCID: 0000-0002-2733-118X

Concept of evidence and the quality of evidence-based reasoning in elementary students

Abstract

The present study has two goals: to explore elementary students' understanding of evidence and the ways they deploy it to construct arguments, and to examine whether eliciting their concept of evidence during argumentation improves students' evidence-based reasoning. Individual semi-structured interviews were conducted with 4th and 6th graders (N=66) in a public school in Mexico. We found significant differences between groups regarding the concept of evidence, with better performance in the older group. A positive correlation between the concept of evidence and the quality of evidence-based reasoning was found. Also, three performance profiles were observed after eliciting the concept of evidence when grade was excluded as a factor. Results suggest that the concept of evidence plays an essential role in developing argumentative competence in pre-adolescence.

Keywords: concept of evidence; argumentation quality; elementary school; evidence-based reasoning; evidence generation; evidence quality.

1. Introduction

Scientific argumentation involves the coordinated use of theory and evidence to substantiate conclusions about natural phenomena based on shared epistemological norms and practices (Erduran, Ozdem and Park, 2015). It involves advancing claims about why something happens (theory) using relevant information (evidence) to establish how it is that we know (epistemology). In this way, scientific argumentation is akin to scientific explanation, which also involves presenting an

account of natural phenomena using theory and evidence. However, unlike explanations, scientific argumentation presumes that knowledge claims are contestable or tentative and must be established through careful consideration not only of how the theory can *account* for the data, but also how the data *substantiate* the theory (Osborne and Patterson, 2011) within a framework of alternatives (Kuhn, 2010; Rapanta, 2019). Across a range of contexts from clinical diagnosis to chronic care management, doctors must engage in *abductive argument* (Walton, 2001; Rapanta, 2018) as they draw tentative conclusions or argue for courses of action under conditions of uncertainty, based on the best available evidence. In order to adequately process these abductive arguments, however, patients must share a common set of beliefs about good evidence with health care providers in order to arrive at shared decisions (Haynes, 2002). They must possess a *concept of evidence* (Aikenhead, 2004) to make sense of recommendations and weigh them against their own impulses, conventional wisdom, or competing recommendations from other professional and non-professional sources (Upshur and Colak, 2003). These are complex proficiencies that rely on the epistemological understanding that claims are contestable, that evidence must be used to support these claims, and that standards can be set for evaluating the strength of evidence in abductive arguments.

According to Josephson and Josephson (1994) abductive arguments hinge on judgments of likelihood that consider, among other things, (1) how plausible a conclusion is and (2) how it compares to alternative conclusions. In this paper, we use a developmental lens to explore young children's concept of evidence (CE), as it relates to these two considerations, in the interest of understanding its role in abductive reasoning. We believe that gaining a deeper insight into CE may shed light on how

best to educate the public to engage with scientific arguments related to personal and public health as well as a wealth of other socio-scientific issues.

1.1 Concept of Evidence

Walton (2016) states that the notion of evidence has been underexplored in argumentation theory, proposing that standards of what constitutes good evidence should be set, to resolve ambiguities and address different levels of precision in its use as a concept. For Walton (2016), in its broadest sense, evidence refers to any argument presented to support or undermine a claim, whereas addressing its more precise meaning must take the context of a disciplinary field into account. Kuhn and Reiser (2005) define argument-based evidence as scientific data that students gather and combine to construct and defend their claims (i.e. numerical data, background information, observation and facts obtained from reading and discussions). Finally, Macagno (2016) defines evidence as measurements, the authority of experts or qualified sources, or experiments that can be used to support a point of view (p.181). Concept of evidence, which addresses what evidence is and how data can be brought to bear on a question in a given discipline, has emerged from research exploring the practical applications of scientific knowledge in professional settings (Aikenhead, 2004; Gott, Duggan, Roberts and Hussain, 2018). As such, it represents a cornerstone of scientific argumentation and a critical proficiency when it comes to understanding how individuals evaluate abductive arguments when making decisions.

Here, evidence is defined as quantitative or qualitative data used to answer a question, solve a problem, or make a decision (Aikenhead, 2004; McNeill, 2011). As such, it is open to critique in terms of the validity of the methods used to collect the data, the sufficiency of the data to substantiate the claim, the relevance or

generalizability of the data to the question at hand (Gott, Duggan, Roberts and Hussain, 2018).

According to Aikenhead (2004), among practicing professionals, “concepts of evidence are usually applied unconsciously as tacit knowledge” (Aikenhead, 2004, p. 247). That is, professionals often apply CE automatically, based on a set of internalized beliefs about the nature of evidence and standards for its use. However, in contexts like medical conversations, the tacit beliefs and assumptions of the patient must also be taken into account when considering the effectiveness of abductive arguments. Patients may treat professional and non-professional advice equally because they do not possess a CE that allows them to compare the strength and validity of one abductive argument against another. Therefore, a natural question to ask from an educational point of view is: where does concept of evidence come from, and what drives its development over time?

1.2 Concept of evidence in school-aged children

The development of a CE begins with the understanding that individuals can hold incompatible views and that the relative strength of these views can be established with arguments and evidence. Research suggests that even young children possess a theory of mind, understanding that others can hold inaccurate or false beliefs and that multiple incompatible views on a phenomenon can be resolved using evidence (Duschl, Schweingruber and Shouse, 2007; Sandoval, Sodian, Koerber and Wong, 2014). However, children are often prone to the absolutist assumption that despite a multiplicity of views, there is only one correct interpretation of a phenomenon and that other views are simply based on false facts (Chandler, Hallett and Sokol, 2002). While this stance represents an important epistemological

milestone towards comparing competing views, it engenders a misconception that has important implications for scientific argumentation. Students must understand that scientific arguments are often constructed under conditions of uncertainty and refined through an iterative process of scrutiny and revision (Osborne, Erduran and Simon, 2004) based on evaluations of available evidence. In addition, while children may understand that beliefs are linked with evidence, they often have difficulty distinguishing between the two (Kuhn, 2005; Berland and Reiser, 2004; Sandoval and Millwood, 2005; Zimmerman, 2007). Kuhn (1991) has observed that young adolescents (even many adults) have difficulties in distinguishing between genuine evidence and what she calls *pseudoevidence*. Kuhn observed that when asked to substantiate a claim, many students provide a "story-like" description of a possible causal sequence (*pseudoevidence*) without providing information that bears on its validity or verifiability (*genuine evidence*). Similarly, Moore (2013) found that without scaffolding, most young adolescents justify their claims using personal testimony rather than shared knowledge.

A CE also involves knowing the function of evidence in an argument and some studies suggest that, under the right conditions, students are able to use evidence in ways that demonstrate some conceptual understanding. Sodian, Zaitchik, and Carey (1991) worked with first- and second graders to analyze their understanding of the relationship between a given piece of evidence and its consequent claim. They concluded that when the claim was simple and clear, and the evidence conclusive, students can distinguish the two and judge the *relevance* of evidence. However, judgments about the *sufficiency* of evidence may take longer to develop (Berland and Reiser, 2004). Sandoval and Millwood (2005) analyzed the quality of evidence cited to justify claims in high school biology and concluded that, although students cited

evidence to justify their claims, such evidence was generally not sufficient, and they had difficulty linking specific data with specific statements. In addition, Garcia-Mila and colleagues (2016) found that senior high school and third-year university students were comparable in the use of evidence to justify and counterargue their claims, with no differences between the two educational levels in using confirming- and disconfirming evidence.

On the other hand, Chen, Hand, and Park (2016) worked with 5th grade students over 16 weeks and found that students improved in their ability to critically examine arguments from evidence. They found a significant increase in the proportion of utterances where students challenged the quality of evidence, focusing specifically on the sufficiency of evidence and soundness of reasoning. Indeed, several studies show that as students engage in argumentative practices, they gradually develop more sophisticated views on the relationship between claims and evidence (Chen, Hand, and Park, 2016; McNeill, 2011; Kuhn, Zilmer, Crowell and Zavala, 2013). For instance, Zembal-Seul, McNeill and Herschberger (2013) engaged K-5 students in scientific practices where they had to provide explanations for inquiry activities. They concluded that students as young as kindergarteners can apply standards of appropriateness and sufficiency if they have been introduced to evidence use. Studies like these suggest that children can evaluate the relative strengths of arguments based on evidence under highly supportive conditions.

In short, there are clear discrepancies in the literature about children's nascent concepts of evidence. Some of these discrepancies may be due to differences in the types of evidence or tasks used to evaluate competence (Amsel and Brock, 1996; Gott and Roberts, 2008; Koslowski, 1996; Kuhn and Udell, 2003; Mean and Voss, 1996;

Sodian and Bullock 2008). Others may be due to the fact that younger students can perform better under highly supportive conditions. Alternatively, this discrepancy could be interpreted as the existence of a developmental stage in which the concept of evidence is still developing, resulting in its highly inconsistent application to reasoning. It would be in this "intermediate" stage where students show inconsistencies in evidence-based reasoning, where triggering the concept of evidence would improve the quality of evidence used. This stage can also be understood as a nascent metacognitive competence. The development of argumentation requires focusing on the metacognitive aspects involved in any reasoning activity that has cognition as its object (Kuhn, 2000). Argumentation involves metacognitive competence with different types and levels of knowledge and skills: (1) declarative knowledge (i.e., metacognitive knowing) and (2) procedural knowledge (i.e., metastrategic knowing) (Kuhn and Pearsall, 1998). We propose that the two types of knowing may reinforce each other in such a way that by enhancing the concept of evidence (i.e., metacognitive knowing), we enhance the proper use of it in an argument (metastrategic knowing) (Rapanta, et al. 2013).

1.3 Reconciling discrepancies in the extant literature

When we look at the literature on the Nature of Science (NOS), we see there is a gap between explicit knowledge of NOS type rules (such as the definition of theory and definition of evidence) and tacit use of those rules (i.e., justifying claims, coordination between theory and evidence) (Khishfe, and Abd-El-Khalick, 2002). Could the reason for children’s inconsistent performance in justifying their theories be an ill-defined or implicit CE? Could eliciting the definition of evidence during the

argumentative process help clarify the task so they can justify their claims more effectively?

One way to reconcile apparently discrepant findings on early competence in evidence-based reasoning is to propose that performance is enhanced by situational factors and does not become reliable and consistent until students have developed a robust and explicit CE to guide their reasoning. Our hypothesis is that a more sophisticated CE, which explicitly distinguishes between theory and evidence and identifies the role of each one in an elaborated argument, may be the means by which students guide their thinking when arguing independently. The primary goal of the present study was to explore the relationship between preadolescents' argumentation with evidence and their CE, or their understanding of what evidence is and the role it plays in substantiating a claim. We investigate the quality of evidence-based reasoning in 4th and 6th graders and explore whether eliciting an explicit definition of students' CE during argumentation positively impacts the quality of their subsequent reasoning. Specifically, our research questions were:

RQ 1: What is the quality of 4th and 6th graders evidence-based reasoning, and are there differences between the two age-groups?

RQ2: What is the concept of evidence held by 4th and 6th graders, and are there differences between the two age-groups?

RQ3: Is there a relationship between the quality of students' CE and the quality of their evidence-based reasoning?

RQ4: Does eliciting CE during a semi-structured interview improve the quality of students' subsequent evidence-based reasoning?

2. Method

2.1. Design

This study consists of a mixed quasi-experimental design that combines quantitative and qualitative analysis. On one hand, it is and exploratory, emergent study embedded in a larger experimental project, which generated new post hoc hypotheses that we test in the present study. That is, the analysis of the effect of triggering the concept of evidence during the interview on the use of evidence. On the other, it is a classic cross-sectional developmental study with two ages under analysis (4th and 6th graders).

2.2 Participants

Participants were 66 students (36 females, 30 males, distributed in equal proportion across two grades) from a rural elementary school in Mexico: 33 fourth graders (mean age= 9.7, SD=0.7) and 33 sixth graders (mean age= 11.5, SD=0.9). These two age groups were selected because a sizable portion of the literature has situated argumentative developments in young adolescence, around 11 years (see Rapanta, Garcia-Mila and Gilabert, 2013, for a review). We were interested in observing argumentative competence before and after this age in order to analyze the consistency in the use of evidence by students in a developing stage in which the concept and use of evidence are emerging and had been reported students at this age perform inconsistently (Kuhn; 2000) .

2.3 Procedure

An individual semi-structured interview was used to explore students' ability to justify their opinion based on evidence they generated. The interview questions were

adapted from Kuhn (1991) with additional questions introduced to explore students' CE. Interviews were conducted by the first author in a quiet room in the school library. After piloting topics, we chose to ask students about increases in the stray dog population because it was familiar and an authentic concern in their community. It has been argued that the familiarity of argumentative content is most effective at eliciting students' underlying argumentative competence (Von Aufschnaiter, 2008), particularly with socio-scientific rather than formal scientific issues (Osborne et al., 2004).

During the interview, students were asked, "*What causes people to abandon so many dogs?*" followed by, "*How do you know that this is the cause?"* and "*If you wanted to convince someone that this is the reason why people abandon dogs, how would you prove it?"* Their answers were coded as Pre-CE prompt justifications. Next, students were asked to explain their CE by defining the term (*"Do you know what evidence is?"* and *"Could you give me an example, so I can understand it better?")* and explaining its function ("*What do you use evidence for?*" *and "So, evidence is…"*). Afterwards, students were asked, once again, to state their causal explanation and the evidence they would use to support it. These answers were coded as Post-CE prompt justifications. Thus, the CE prompt served both as a measure of the sophistication of their CE and as a scaffold for refining their initial arguments and evidence. The average length of the interview was twenty minutes. The entire interview is presented in the Appendix.

2.4 Instruments

Rubrics were developed to assess the quality of evidence students used to justify their reasoning (Table 1) and the quality of their CE (Table 3). Both coding schemes were developed by the first two authors. For each scheme, an initial set of

codes w developed based on the literature, and then refined iteratively based on student data.

One indicator of the quality of argumentation is the reference to the interlocutor's point of view (Kuhn, 1991). In the present study, this indicator is found in the students' understanding and quality of the answers of the interview questions that address the opponent's claim ("What would you say if someone did not agree with you and said [the opposite]?") Their answers were used to code the function of the argument, whether it is a justification, a counterargument or a rebuttal.

Another indicator of argumentative quality is the use of data as evidence to support the claims directly without interpretation or elaboration of the information (less sophisticated) (Macagno and Walton, 2014) or its use to build the warrant (more sophisticated) (Macagno, Mayweg-Paus, and Kuhn, 2015; Sandoval and Millwood, 2005). In order to assess and compare the use of evidence in non-dialogic and dialogic settings Macagno (2016) focused on two broad criteria for the use of evidence: (1) structure and (2) relevance. The first one consists of a 5-level classification of evidence: Anecdotal, Authority Con, Reference to a source (scientists, government, institutes), Popular Opinion, Laboratory (Authority), and Statistics. These criteria are crossed with the function of evidence: justification (support), conterargumentation, (weaken) rebuttal (support) or neutral.

Regarding the relevance criteria, Walton (2004) states that a dialogic argument is generally considered relevant as long as it is functional (informative) to the goal of such dialogue (Walton, 2004). According to Macagno (2016) and Kuhn and Udell (2003),evidence can be nonjustificatory (not addressing the problem at all) nonfunctional (addressing only tangential aspects of the problem) and functional

arguments (addressing core aspects of the problem).Therefore, evidence is irrelevant when it does not support the claim sufficiently or when it unconnected to previous moves.

2.4.1 Rubric for coding the quality of evidence.

For the present investigation, we define evidence as qualitative or quantitative information used to demonstrate a causal sequence of events or in support of a claim. The quality of the evidence is defined in terms of its adequacy in demonstrating the claim or causal sequence, its relevance to the phenomenon described, its validity and generalizability.

A descriptive five-point ordinal scale was used to analyze the quality of evidence used to justify claims (Table 1). The complexity levels of the rubric were defined using two criteria: the source of information, ranging from personal or anecdotal evidence to external and validated evidence; and its generalizability, ranging from a single piece of evidence to multiple pieces of corroborated evidence.

For the design of the rubric to codify the validity of the evidence, we take as a reference the work of Kuhn (1991) and the coding criteria of Macagno (2016) for (1) the structural use of the evidence and (2) the relevance of such evidence to determine if the evidence is informative. Another indicator incorporated into our rubric was the sophisticated use of evidence, whether it was used simply, to support the claim (or weaken the opposite position) directly or involve a more scientific use of the evidence in interpreting and elaborating the information, building on the warrant. This derived into coding one sophisticate use of evidence (closer to scientific) of better quality that several simple uses of evidence.

[Insert Table 1 approximately here]

To illustrate how the coding scheme was applied, we present examples of evidence at each level of performance in Table 2. These examples are drawn from the interview data.

[Insert Table 2 approximately here]

2.4.2 Rubric for coding the concept of evidence.

The rubric in Table 3 was used to explore students' understanding of CE, which is operationalized here as a definition of evidence and an explanation of its function in argumentation. A five-point ordinal scale was used to code definition of evidence, and a three-point ordinal scale was used to code the function of evidence (concrete and abstract levels).

[Insert Table 3 approximately here]

To illustrate students' definitions of evidence at each level, representative examples selected from interview data are presented in Table 4.

[Insert Table 4 approximately here]

Table 5 provides representative examples of the function of evidence at each level, taken from interview transcripts.

[Insert Table 5 approximately here]

2.5 Data analysis

All interviews were transcribed and coded using the analysis rubrics. For statistical analysis, the maximum level reached by each student was taken as a dependent measure. Inter-rater reliability was calculated for 33% of the interviews across the rubrics and reached 85% agreement. Differences in coding were resolved through discussion. Statistical analysis was conducted sequentially according to research questions. First, Mann-Whitney nonparametric tests of means comparison were conducted to assess differences between the two age-groups in the level of evidence generated to justify a claim (RQ1), the CE (RQ2), and the function of evidence (RQ2). Second, Spearman correlations were used to explore relationship between the level of evidence students produced, and their understanding of the CE and its function both in general and by grade (RQ3). Finally, Chi-square analyses were conducted to explore the influence of eliciting students' CE during interviews (RQ4). Qualitative data are provided to complement and illustrate each of the quantitative findings.

3. Results

3.1 Justification quality and evidence generation

3.1.1 Differences in the quality of evidence-based reasoning pre- and post-CE prompt by grade

In order to answer our first RQ, which involved comparing 4th- and 6th-graders' evidence-based reasoning performance, students were prompted to provide evidence to justify their claims before and after defining evidence, and differences in both measures were compared by grade. Mann-Whitney U nonparametric tests showed no

significant differences between groups regarding the level of evidence used to justify claims either before or after the CE prompt (Table 6). However, we did find significant differences between groups for students' CE and its function, with the older group performing better on both variables (Table 6).

[Insert Table 6 approximately here]

Although no significant differences were found between groups in the level of evidence either before (*p*=.375; see table 6) or after (p=.765) the CE prompt, sixth graders did show more fluctuations in performance (see Figures 1 and 2). As we can see, 51% of the of 6th graders (17/33) and 70% of 4th graders (23/33) provided experiential evidence (simple and elaborated) to justify their claims; 27% (9/33) of 6th graders and 20% (7/33) of 4th graders used evidence from external sources of information, such as media, books, authorities or experts; and 21% (7/33) of 6th graders failed to provide any evidence at all to back up their claim, providing only supporting reasons as justification. Supporting reasons were additional statements that elaborated on an explanation without offering further evidence to support the claim. In contrast, only 9% (3/33) of 4th-graders did not provide evidence. It is worth mentioning that although seven of the sixth graders did not provide evidence in most of the cases, they did provide several reasons to support their claims in order to make the cause of abandonment more plausible. In fact, sixth graders tended to provide supporting reasons more frequently than fourth graders even in those cases where the evidence was coded at higher levels. In contrast, fourth graders avoided using supporting reasons and instead backed up their claims with basic evidence, most often from their own experiences.

With respect to the levels of evidence used after students presented their CE, 18% (6/33) of fourth graders and 12% (4/33) of sixth graders made an argument without any evidence (level 0); almost 30% of each group used Level 1 or 2 evidence (9 and 8 of 33, respectively); and around 25% of each age group used Level 3 evidence. None of the students used level 4 evidence in any of the groups either before or after the CE prompt. Level 4 evidence represents an elaborated justification backing their opinion with supporting reasons and evidence from multiple valid sources. The distribution of cases for evidence level Pre- and Post CE prompt are shown in Figures 1 and 2 respectively.

[Insert Figure 1 approximately here]

[Insert Figure 2 approximately here]

No statistical differences in level of evidence were found when Pre- and Post-CE prompt performance was analyzed. Nevertheless, some slight changes in performance were observed. In evidence level Post-CE prompt, sixth graders showed a discrete improvement. The percentage of students who were not able to provide evidence got reduced from 21 to 7%. Correspondingly, the percentage of students in level 2 of evidence increased from 15 to almost 30% Post-CE prompt, coming mostly from Level 1. These changes could be related with the repeated measurement process and the elicitation of CE.

3.1.2 Evidence to support the reason versus to prove the causal sequence of abandonment

In relation with the quality of elementary students' evidence-based reasoning (RQ1), in a more detailed analysis of the data, we found that many of the students'

justifications provided evidence to support the feasibility of their reasoning rather than to support the likelihood of their reasoning was correct. Rather than generate evidence to support their cause-and-effect argument, they simply presented evidence that the presumed cause had occurred.

The following examples show how both age groups tend to justify "the reason" for abandoning dogs ("because they bite"): "I could show you the bite marks to prove that dogs bite", or in a more elaborated way "I could show you the bite marks, like torn clothing or a witness who saw the dog bite him." In other cases, students did not refer to evidence and focused instead on the plausibility of the reason. For example, a sixth-grade student argued: "people abandon dogs because they transmit rabies and scabies, and if a dog bites a kid, the illness could spread…[and] cause an outbreak." In these examples, students' efforts clearly target their reasoning and ignore the need to substantiate the theory with evidence.

A Mann-Whitney test of means comparison was conducted to analyze differences in students' use of evidence. Significant differences were found ($U=412.50$, $p=.038$; 4^{th} grade Mean=1.79, SD= .41, 6^{th} grade Mean=1.55 SD=.50). Most 4^{th}-graders (79% of the group, 26/33 students) tended to provide evidence to demonstrate the causal sequence, while 6^{th} graders' evidence types were distributed more evenly between categories: 54% (18/33) directed to illustrate the causal sequence and 46% (15/33) to stay their reason as valid. Figure 3 shows the distribution of frequencies.

[Insert Figure 3 approximately here]

3.2 Concept of Evidence

3.2.1 Differences in the definition of evidence by grade

To address our RQ2 about the concept of evidence held by 4th- and 6th-graders, the levels of comprehension the concept and function of evidence, were assessed by group and compared by grade. A Mann-Whitney U test showed significant differences between groups on the variable *CE* with the oldest group showing better performance (4th grade *mean*= 0.64; 6th grade *mean*=1.42; *U*=303.0 *p*=.001). Figure 4 shows the performance of each group across the five levels. As we can see, more than a half of the 4th-graders (17/33) fell in the lowest category for defining evidence (level 0), 36 % of this group (12/33) provided an incomplete definition and were not able to give an example to explain their CE (level 1); around 10% (3/33) gave an incomplete definition of evidence with a concrete example (level 2); and only one student provided a sufficient definition of evidence with a concrete example (level 3). None of the students in 4th grade reached Level 4 in CE.

For 6th-graders, students' performance was distributed across all levels: 18% (6/33) did not provide a definition of what evidence is; 39 % (13/33) provided an incomplete definition without an example; 30% (10/33) gave an incomplete definition with a concrete example of evidence; and only 12% offered a sufficient and complete definition of evidence (4/33, 2 in level 3 and 2 in level 4) on the rubric.

[Insert Figure 4 approximately here]

The most common answers students provided to illustrate their CE were related to criminal investigations, often derived from their contact with criminal investigation programs on television. This connection mainly occurs at levels 2 and 3, where

students provided a concrete definition of evidence. In these cases, students showed difficulty distinguishing between the physical evidence they referred to in their examples (like fingerprints, blood or footprints), and the definition of evidence (see Table 4).

In more complex and abstract definitions (level 4), police investigations do not appear, perhaps because students have developed a CE beyond the examples of their daily lives. Questions about the function of evidence helped to explore students' comprehension of CE in greater depth. The results are presented below.

3.2.2 Function of evidence

As with *CE*, significant differences were found between grades in explanations of the function of evidence using a Mann-Whitney U test, with 6th graders (*Mean=1.06)* outperforming 4th graders (*Mean*= 0.52, U= 320.5, *p*=.002). Figure 5 presents the distribution of student responses regarding the function of evidence by grade (see Table 3 for the rubric). In the 4th grade, 52% (17/33) of the students failed to assign any function to evidence (level 0), 45% (15/33) provided a concrete explanation of its function, and only one student provided an abstract explanation of its function.

On the other hand, 52 % (17/33) of the 6th graders assigned a concrete explanation of the function of evidence (level 1) and 27% (9/33) provided an abstract and generalizable function (level 2). However, 21% (7/33) also failed to offer any explanation of its function (see Figure 5).

[Insert Figure 5 approximately here]

As with their definitions of evidence, students' explanations of its function usually involved specific, concrete examples, again generally associated with criminal investigations from television programs. Abstract definitions of the function of evidence were less context dependent, better differentiated from the examples, and therefore closer to a formal definition.

3.3 Relationship between level of evidence and concept of evidence

To explore the relationship between the quality of the Concept of evidence held by the students and the quality of their evidence-based reasoning (RQ3), a Spearman's correlation test was run to explore the association between *CE* and the level of evidence used to justify claims both before and after CE prompts (see Table 7). A positive correlation between the CE and the level of evidence in Post-CE prompt justifications for both grades was found ($p<.001$). In addition, a modest correlation was found between the *CE* and the level of evidence in Pre-CE prompt justifications ($p<.001$).

A positive correlation was also observed between explanation for the function of evidence and the level of evidence cited Post-CE prompt ($p=.014$), but no significant correlation was found between these two variables for Pre-CE prompt (see Table 7). These results could indicate a possible positive effect of eliciting CE on the level of evidence.

[Insert Table 7 approximately here]

3.3.1 Profiles of justification performance in relation to concept of evidence

Once a positive correlation was found between CE and evidence based-reasoning, we decided to explore this relation more deeply by grouping students' performance in three different profiles*: improved*, *maintained* and *decreased performance,* in order to answer our RQ4 about the effect of CE elicitation during the interviews and its potentiality to improve students' evidence-based reasoning (RQ4). The Chi-square analysis of the distribution of cases yielded significant differences when grade level was removed as a factor (see Table 8). Students who performed well in the CE definition also showed a higher level of evidence to justify claims that was maintained in the Pre and Post assessment (10/66). On the other hand, students with low or very low performance on the CE definition but who had some previous knowledge about CE, showed better performance on level of evidence for Post-CE justifications (10/66). Finally, those students who did not have a developed CE (13/66), performed worse on level of evidence for Post-CE justifications

[Insert Table 8 approximately here]

Of the three profiles, students at levels 1 and 2 of CE showed the most improvement. This group showed an interesting positive effect of the elicitation of CE during the argumentative process. Despite having only a basic knowledge about evidence and its function, they provided higher quality evidence after defining evidence and explaining its purpose.

4. Discussion and conclusions

The main goals of the present study were to explore elementary students' understanding of what evidence is and the role it plays in substantiating a claim, and to examine whether eliciting students' CE during argumentation had an impact on the quality of the evidence they generated to justify their claims. For this purpose, the quality of the evidence students provided before and after giving a definition of the concept and function of the evidence was assessed.

4.1 Concept of evidence in elementary school

Students in both grades performed better in explaining the function of evidence than defining the concept. Their CE focused primarily on concrete examples of physical evidence used in criminal investigations, suggesting they had internalized some understanding of evidence from informal contexts but had not developed declarative knowledge of the function of evidence in an argument. These results lead us to conclude that students seem to have tacit knowledge of how evidence can be used to confirm the truth of knowable events, but lack the explicit, generalizable understanding of its function in strengthening a claim in light of alternatives. Despite some students citing evidence from informal contexts, it is worth noting that more than half of the fourth graders (17/33) and one-fifth of the sixth graders (7/33) were unable to explain the function of evidence even in a concrete example. Results were similar with respect to students' definitions of evidence. Most students provided concrete examples of evidence (e.g., fingerprints, blood or footprints) without producing a generalized definition. These findings illustrate how students struggle to draw a distinction between the physical evidence referred to in their examples and the meaning of evidence, and in fact, they struggle with abstract definitions on the whole.

Only in cases where students could produce abstract definitions, did we see general examples of evidence replace examples from criminal investigations, possibly because these students had appropriated a CE and could apply it to different contexts. Only two students out of the 27% of sixth graders that understand the abstract function of evidence were able to provide a complete abstract definition of evidence, and only 3 students of the whole sample were able to provide a sufficient concrete definition. Another conclusion derived from the results, is that although the oldest group performed better in explaining the function of evidence, most elementary students show a naive CE. This could help explain their difficulty in finding evidence to structure a well-grounded justification of their claims.

Barchfeld and Sodian (2009) point out that explicit understanding of the quality of evidence is hardly found even in adults and is acquired over time across educational experiences. They suggest the need for further research to explore whether a nascent understanding of the validity of evidence can be taught in young students. We agree that further research is needed not only to explore students' understanding but also to improve it by engaging students in argumentative practices that explicitly foster their CE.

4.2 Quality of evidence-based reasoning in elementary school

Regarding the quality of evidence-based reasoning, students in both age groups showed only rudimentary performance in generating evidence, and they had difficulty relating claims with relevant evidence. These findings are consistent with studies that show that students struggle with constructing well-grounded arguments and differentiating between evidence and theory (Kuhn, 1991; Osborne et al, 2004; Berland and Reiser, 2004; Sandoval and Millwood, 2005; Fulton y Poetler, 2013). Our

results are also consistent with the literature regarding students' tendency to justify using more personal evidence than shared knowledge, particularly without scaffolding (Moore, 2013).

Moreover, a considerable percentage of students, particularly in 6th grade (more than 20% of the group) did not provide any evidence to back up their claims before they were asked to define evidence. It is interesting that not only were there no positive developmental trends, but also a higher number of 6th grade students failed to provide any evidence at all to support their opinion. This finding is consistent with the National Research Council report about wide variations in the sophistication of reasoning strategies across individuals of the same age (Duschl et al., 2007). Nevertheless, although many students did not offer evidence to support their claims, they did focus on explaining the plausibility of their claims. Older students tended to provide more supporting reasons to back their claims, explaining the reasonableness of their assertion more than substantiating the causal sequence with evidence. They did not seem to be aware of the fact that making one's claims more plausible does not make them more likely than other claims. Indeed, they consistently addressed the plausibility of their abductive arguments without addressing the strength of evidence against alternatives. The lack of evidence generation could be related to students' difficulties in locating valid evidence to support their claims, which would be consistent with Brem and Rips' (2000) observation that people resort to explanations when they do not have available evidence or doubt its validity.

Alternatively, this problem in evidence use may be because they had difficulty distinguishing between theory and evidence. It is worth mentioning that even those who cited high-level sources of information, such as surveys or DNA tests, were

unable to link their evidence with causal claims. That is, the level of complexity of the information cited as evidence and the ability to differentiate between theory and evidence do not seem to be directly related. Just because students can (or cannot) provide better evidence when asked to support a claim does not mean they can distinguish between making a claim more plausible and strengthening it against competing claims. Put another way, students fail to distinguish the two considerations outlined by Josephus and Josephus (1994) when weighing abductive arguments: (1) how plausible a conclusion is and (2) how it compares to alternative conclusions. As a result, they mistake the task of arguing for a plausible conclusion with arguing that it is the strongest conclusion.

4.3 Impact of eliciting concept of evidence on the quality of justifications

When performance patterns of pre-post CE prompt were analyzed, we found that the percentage of sixth graders who only provided reasons to justify their claims went down, while the percentage of 4th graders who provided no-evidence all went up. CE may have played a role in both of these findings. In the case of sixth graders, eliciting the definition and function of evidence may have prompted them to produce better evidence. In contrast, the fourth graders who showed a less developed CE, were more likely to respond 'I don't know' when asked how they might justify their claims.

Thus, even though no significant differences were found between groups in the level of evidence either before or after they defined their CE, those students that had a more complex understanding of what evidence is were more likely to properly justify their opinion when prompted. The fact that this relationship was stronger for the justification made after defining evidence may suggest that eliciting the CE during the

argumentation process can help students who have previous knowledge of evidence to improve their performance. This interpretation of our findings is further substantiated by our Chi-square analysis of the three profiles of performance, which showed that students with incipient knowledge of the meaning and function of evidence were more likely to benefit from the elicitation of CE during argumentation.

Our findings suggest that, although students may lack a coherent epistemological framework for using evidence to construct an argument, they can produce higher quality evidence in support of their claims when prompted. We conclude that this gap between what they can do and what they actually do, points to the importance of helping students not only distinguish theory from evidence, but also clearly articulate the role of evidence in advancing theory.

4.4 Promoting argumentation through the use of evidence

The present study investigates elementary students' concept of evidence (CE), a basic component of their epistemological understanding, which involves their definition of evidence and the role it plays in providing grounds for a claim. Our results suggest that an explicit CE is linked to the quality of evidence students use to justify their reasoning, and that triggering the CE prompts improvements in their performance. Despite all expectations reflected in school curriculum, research in argumentation has shown the complexity of teaching argumentative process and the struggles that elementary, secondary and even higher education students have in evaluating and producing well-grounded arguments (Barchfeld and Sodian, 2009; Brown et al., 2010; Kuhn, 1991). These difficulties may be partially explained by the lack of clarity about what evidence is and how to use it to generate knowledge. We propose that one way to improve students' justifications and help them distinguish

theory from evidence could be explicitly eliciting and addressing students' CE during the argumentative process as a way of scaffolding their reasoning.

Our results show that explicitly asking for students' CE definition during argumentation could improve students' awareness of evidence quality criteria and its function in justification, resulting in better quality of their arguments, especially for those students who have at least an incipient understanding of evidence. These findings highlight the relevance of considering students' preconceptions about evidence in pedagogical interventions. Indeed, students' poor understanding of the CE, and the fact that most of their evidence draws from popular culture, highlight the need to make the meaning and use of evidence more explicit in school.

4.5 Limitations of the study and further research.

A limitation of the present study is the lack of a control group that would have allowed us to confirm that the statistical differences described in our results were caused by the elicitation of the concept of evidence. This design flaw resulted from the fact that the study emerged as a set of questions from a larger study, in which CE was not identified as an independent variable. As a result, we have provided an exhaustive qualitative analysis to better understand the statistical findings more deeply. We felt the effort worthwhile, particularly in light of the paucity of studies dedicated to CE, particularly in young adolescents.

We think that our results highlight the importance of students' concept of evidence to design pedagogical interventions. The elicitation of the concept of evidence seems to increase students' awareness of their knowledge about evidence (or lack thereof), and this awareness may contribute to more consistent and explicit argumentative performance (Kuhn and Udell, 2003, Kuhn and Pearsall, 1998).

Acknowledgments.

This study was supported by the Mexican Consejo Nacional de Ciencia y Tecnología (CONACYT) and the Spanish Ministerio de Economía y Competitividad (EDU2013-47593-C2-2-P; DGICYT 2018_RTI2018-097289-B-I00 and the PRX18/00039).

References:

Aikenhead, G.S. (2004). Science-Based Occupations and the Science Curriculum: Concepts of Evidence. Science Education, 89(2), 242-275. doi:10.1002/sce.20046

Amsel, E., Brock, S. (1996). The development of evidence evaluation skills. Cognitive Development, 11(4), 523–550. doi:10.1016/S0885-2014(96)90016-7

Barchfeld, P., Sodian, B. (2009). Differentiating Theories from Evidence: The Development of Argument Evaluation Abilities in Adolescence and Early Adulthood. Informal Logic, 29(4), 396-416. doi:10.22329/il.v29i4.2906

Berland, L. K, Reiser, B. (2004). Students constructing and defending evidence-based scientific explanations. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Dallas, TX.

Brem, S. K., Rips, L. J. (2000). Explanation and evidence in informal argument. Cognitive science, 24(4), 573-604. doi:10.1016/S0364-0213(00)00033-1

Brown, Furtak, Timms, Nagashima, Wilson (2010). The Evidence-Based Reasoning (EBR) Framework: Assessing scientific reasoning, Educational Assessment, 15(3-4), 123-141. doi:10.1080/10627197.2010.530551

Chandler, M. J., Hallett, D., Sokol, B. W. (2002). Competing claims about competing knowledge claims. In B. K. Hofer P. R. Pintrich (Eds.), Personal epistemology: The

psychology of beliefs about knowledge and knowing (pp. 145-168). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. doi:10.1353/csd.2005.0021

Chen, Y-C., Hand, B., Park, S. (2016). Examining Elementary Students' Development of Oral and Written Argumentation Practices Through Argument-Based Inquiry. Science and Education, 25, 277–320. doi:10.1007/s11191-016-9811-0.

Duschl, R. A., Schweingruber, H. A., Shouse, A. W. (2007). Taking science to school Learning and teaching science in grades K–8. Washington, DC: National Academic Press. doi:0.17226/11625

Erduran, S., Ozdem, Y., Park, J. Y. (2015). Research trends on argumentation in science education: A journal content analysis from 1998–2014. International Journal of STEM Education, 2(1), 5.

Fulton L., Poetler, E. (2013). Developing a scientific argument. Modeling and practice help students build skills in oral and written discourse. Science and Children, Summer 2013, 30-35.

Garcia-Mila, M., Pérez-Echeverría, M.P., Postigo, Y., Martí, E., Villarroel, C., and Gabucio, F. (2016). Nuclear power plants? Yes or no? Thank you! The argumentative use of tables and graphs. Infancia y Aprendizaje, 39 (1), 187-218. doi:10.1080/02103702.2015.1111605

Gott, R., Duggan, S. and Roberts, R. and Hussain, A. (2018). Research into Understanding Scientific Evidence. http://www.dur.ac.uk/rosalyn.roberts/Evidence/cofev.htm (accessed 1th June 2019).

Gott, R., Roberts, R. (2008). Concepts of Evidence and their role in open-ended practical investigations and scientific literacy; background to published papers. United Kingdom: Durham University.

Haynes, R. B. (2002). What kind of evidence is it that Evidence-Based Medicine advocates want health care providers and consumers to pay attention to?. BMC health services research, 2(1), 3.

Josephson, J.R. Josephson, S.G. (1994). Abductive Inference: Computation. Philosophy. Technology, New York, NY: Cambridge University Press.

Kuhn, D. (1991). The skills of argument. Cambridge, MA: Harvard University Press. doi:10.1017/CBO9780511571350

Kuhn, D. Pearsall, S. (1998). Relations between metastrategic knowledge and strategic performance. Cognitive Development 13 (2), 227-247. doi.org/10.1016/S0885-2014(98)90040-5

Kuhn, D. (2000). Metacognitive development. Current Directions in Psychological Science, 9(5), 178-181. http://dx.doi.org/10.1111/1467-8721.00088

Kuhn, D. Udell, W. (2003) The Development of Argument Skills. Child Development, 74 (5), 1245-1260.

Kuhn, D. (2005). Education for Thinking. USA: Harvard University Press.

Kuhn, L., y Reiser, B. (Abril de 2005). Students constructing and defending evidence-based scientific explanations. En annual meeting of the National Association for Research in Science Teaching. Dallas, EUA.

Kuhn, D. (2010) Teaching and Learning Science as Argument. Science Education, 94 (5), 810-824. doi:10.1002/sce.20395

Kuhn, D., Zilmer, N., Crowell, A., Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. Cognition and Instruction, 31(4), 456-496. doi:10.1080/07370008.2013.830618

Khishfe, R., and Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. Journal of Research in Science Teaching, 39, 551-578. doi:10.1002/tea.10036

Macagno, F., Walton, D. (2014). Emotive language in argumentation. New York: Cambridge University Press.

Mayweg-Paus, E., Macagno, F., Kuhn, D. (2015). Developing argumentation strategies in electronic dialogs: Is modeling effective? Discourse Processes1–18. http://dx.doi.org/10.1080/0163853x.2015.1040323.

Macagno, F. (2016). Argument relevance and structure. Assessing and developing student' uses of evidence. International Journal of Educational Research, 79, 180-194.

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. Journal of Research in Science Teaching, 48(7), 793–823. doi:10.1002/tea.20430

Means, M. Voss, J. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. Cognition and Instruction, 14(2), 139–178.

Moore, W. (2013). The Use of Evidence in Young Adolescents' Argumentation. Doctoral Dissertation. USA: Columbia University.

National Research Council. (2007). Taking Science to School: Learning and Teaching Science in Grades K-8. Committee on Science Learning, Kindergarten Through Eighth Grade. Richard A. Duschl, Heidi A. Schweingruber, and Andrew W. Shouse, Editors. Board on Science Education, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. doi:10.17226/11625.

Osborne, J. Paterson, A. (2011). Scientific Argument and Explanation: A Necessary Distinction?. Science Education, 45(4), 627-638. **https://doi.org/10.1002/sce.20438**

Osborne, J., Erduran, S. Simon, S. (2004). Enhancing the Quality of Argumentation in School Science. Journal of Research in Science Teaching, 41 (10) 994–1020. doi:10.1002/tea.20035.

Rapanta, C. (2019). Argumentation strategies in the classoom. Vernon Press, Wilmington, DE.

Rapanta, C. (2018). Teaching as Abductive Reasoning: The Role of Argumentation. Informal Logic, 38(2), 293-311.

Rapanta, C., Garcia-Mila, M., Gilabert, S. (2013). What Is Meant by Argumentative Competence? An Integrative Review of Methods of Analysis and Assessment in Education. Review of Educational Research, 83(4), 483–520. doi:10.3102/0034654313487606

Sandoval, W.A. Millwood K.A. (2005). The Quality of Students' Use of Evidence in Written Scientific Explanations. Cognition and Instruction, 23(1), 23–55. doi:10.1207/s1532690xci2301_2

Sandoval, W.A., Sodian, B., Koerber, S. Wong, J. (2014). Developing Children's Early Competencies to Engage with Science. Educational Psychologist, 49(2), 139-152. doi:10.1080/00461520.2014.917589

Sodian, B., Bullock, M. (2008). Scientific reasoning—Where are we now? Cognitive Development, 23(4), 431–434. doi: 10.1016/j.cogdev.2008.09.003

Sodian, B., Zaitchik, D. Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. Child Development, 62, 753-766. doi:10.2307/1131175

Upshur, R. E., Colak, E. (2003). Argumentation and evidence. Theoretical Medicine and Bioethics, 24(4), 283-299.

Walton, D. (2001). Abductive, presumptive and plausible arguments. Informal Logic, 21(2).

Walton, D. (2004). Relevance in argumentation. Amsterdam-Philadelphia: John Benjamins Publishing Company.

Walton, D. (2016). Argument Evaluation and Evidence. Switzerland: Springer.

Von Aufschnaiter, C., Erduran, S., Osborne, J., Simon, S. (2008). Arguing to Learn and Learning to Argue: Case Studies of How Students' Argumentation Relates to Their Scientific Knowledge. Journal of Research in Science Teaching, 45 (1), 101–131. doi:10.1002/tea.20213.

Zembal-Saul, C., McNeill, K., Hershberger, K. (2013). What's your evidence? Engaging K-5 students in constructing explanations in science. Pearson Education Inc.: New Jersey, NJ.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. Developmental Review, 27, 172–223.

Appendix

Justification before elicitation

- Why do you think is people abandon dogs?
- How do you know that is the cause?
- If you want to convince someone that this is the cause people abandon dogs, how would you demonstrate it?
- Can anyone else prove that this is the cause of dogs' abandonment? Who? How?

Evidence concept and function elicitation

- Do you know what is an evidence?
- What does evidence serve for?
- Could you give me an example, so I can understand it better?
- Thus, an evidence is…

Justification after elicitation

- Then, if I tell you: "What evidence could you provide to demonstrate that this is the cause of dogs' abandonment" What evidence would you give me?
- Can you think in another evidence?
- Can anyone else give me another evidence(s) to demonstrate that this is the cause of dogs' abandonment? Who? What evidence?

Table1. Rubric for coding the quality of evidence

Level	**0**	**1**	**2**	**3**	**4**
Evidence provided	Does not provide any evidence to justify an opinion	Provides single, simple piece of evidence based on personal experience	Provides multiple or elaborated evidence based personal experience	Provides a single piece of evidence based on external sources of information that are considered valid, including others' experience *in situ (internet, television, books, information provided by experts or authorities)*	Provides more than one piece of evidence based on external sources of information that are considered valid *(media information, information provided by authorities, statistics, experiments*)

Table 2. Examples of each level for quality of evidence

Level of evidence	***Answer to the question: How could you prove that dogs are abandoned for that reason?***
0	*'I don't know" / "I had the idea"/ "It came to mind'*
1	*'Because if I approach the dog it could bite me'/ 'Because my neighbor always abandons his dogs'*
2	*'Because I've seen a lot of people who have Bull Terriers, Pit Bulls or those kinds of dogs…[that] bite other people and they sometimes abandon them or kill them ...'*
3	*'Like ... in a school, doing as a survey which reason is more ... if they abandon because they are bad for health or for something else"/ 'we can ask to the doctor if it is possible to be infected with an illness by a dog to confirm'*
4	*'We could interview people who have abandoned dogs and have oral testimonies [...] we could also do a survey about the main cause people abandon dogs`*

Table 3. Rubric for coding the definition and function of evidence

Level	**0**	**1**	**2**	**3**	**4**
Definition of Evidence	No definition or unrelated definition to evidence meaning.	Incomplete definition and no example	Incomplete definition with an example (concrete)	Sufficient concrete definition with an example (concrete)	Sufficient definition with or without an example (abstract)
Level	0	1		2	
Function of evidence	Lack of understanding of the function of the evidence	Concrete understanding of the function of evidence, attached to an example.		Abstract comprehension of the function of evidence	

Table 4. Examples of CE at each level

Level	Concept of Evidence
0	*'Is like a dome of medical stuff'*
1	*'Pictures that somebody take of someone to tick him off for something'*
2	*'In a manner of speaking, if somebody steals, but didn't wear gloves and when he touches things, he leaves his fingerprints and they can track him'*
3	*´(It)is like...now we are doing exams, and the teacher think we didn´t study, she said: "You aren't studying!" and we say "yes!", so our teacher told us "show me an evidence", and if we get good grades, we are already demonstrating it...that's an evidence.´*
4	*'Evidence is when we have something to clarify things, it is like an argument, but evidence is, for example: I have evidences that I enjoy studying, I have diplomas and that is evidence…'* *'Evidence is what you bring, about an episode of your life or a fact, in documents, pictures or videos…'*

Table 5. Examples of the function of evidence at each level

Level	Function of evidence
0	Does not understand the function *'So they can see how evidence is made, so they can make it'*
1	Provides a concrete function *'Evidence serves to track who did it'*
2	Provides an abstract function *'It serves us to demonstrate what we are saying'*

Table 6. *Mann-Whitney U test*s. Means comparisons for evidence level, concept and function by grade

Variables	Grade	Mean	SD	*p*
Evidence level Pre-CE prompt	4	1.70	0.91	.375
	6	1.48	1.12	
Evidence level Post-CE prompt	4	1.58	1.08	.765
	6	1.67	1.03	
Concept of evidence	4	0.64	0.78	.001
	6	1.42	1.06	
Function of evidence	4	0.52	0.56	.002
	6	1.06	0.70	

Table 7. Correlation between CE and evidence level before and after evidence definition. **p<.01

Spearman´s Rho N=66	**Justification** Pre-CE (Evidence level)	**Justification** Post-CE (Evidence level)
Concept of evidence	.300*	.531**
Function of evidence	.125	.483**

Table 8. Profiles regarding the relationship between CE and level of evidence.

	Evidence-based reasoning post CE elicitation			
Level of CE	Decrease	Maintain	Improve	Total
0	13	5	5	23
1-2	5	6	10	21
3-4	2	10	4	16
Total	20	21	19	60
***Chi Square p* =.008**				

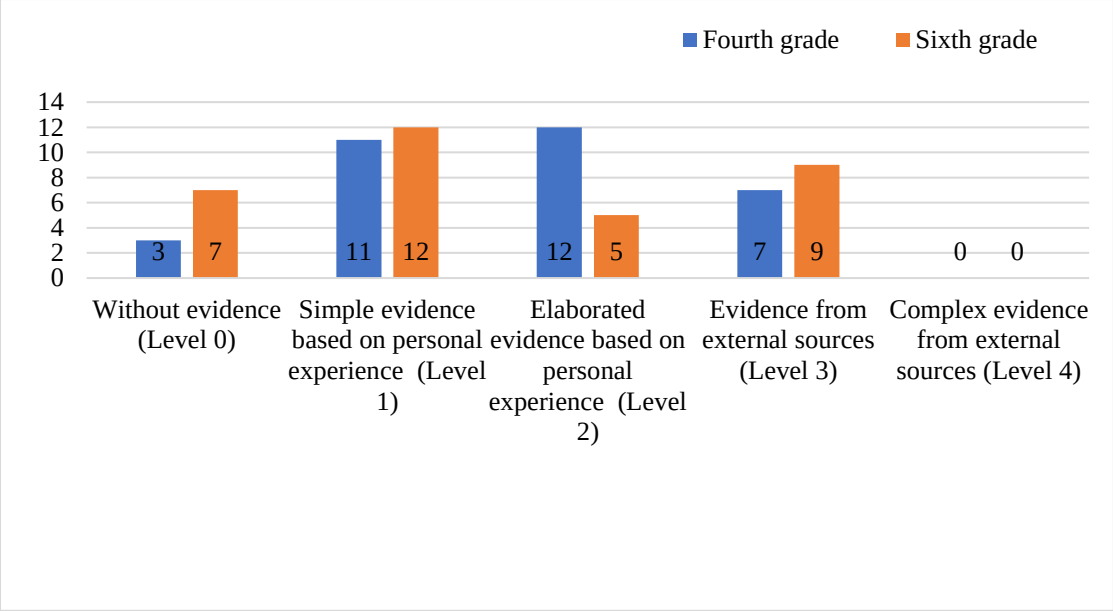

Figure 1. Distribution of Frequencies according to Levels of Evidence Pre-definition of CE by grade.

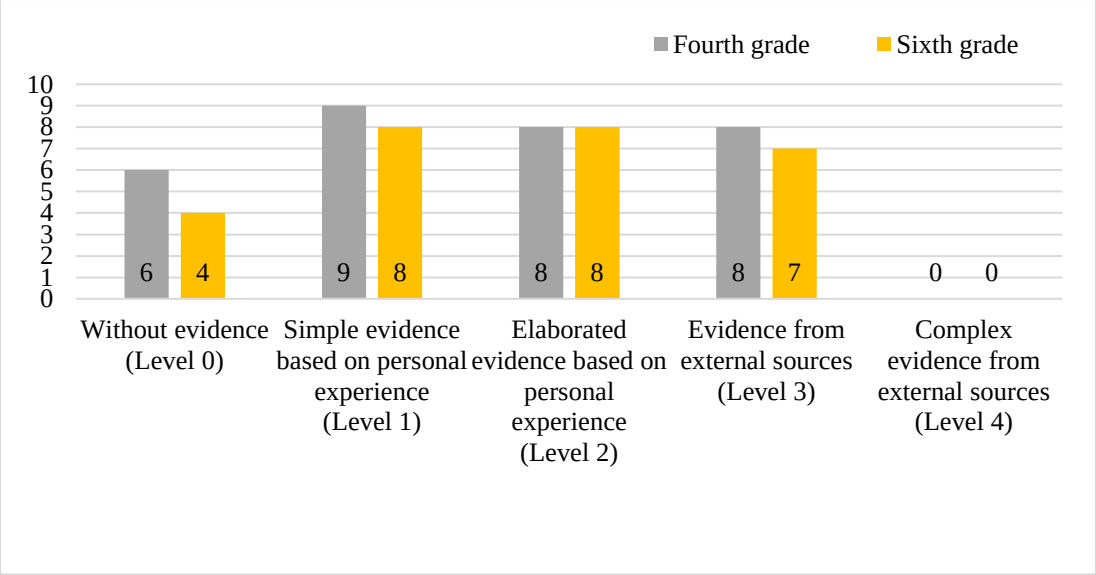

Figure 2. Distribution of Frequencies according to Levels of Evidence Post-definition of CE by grade.

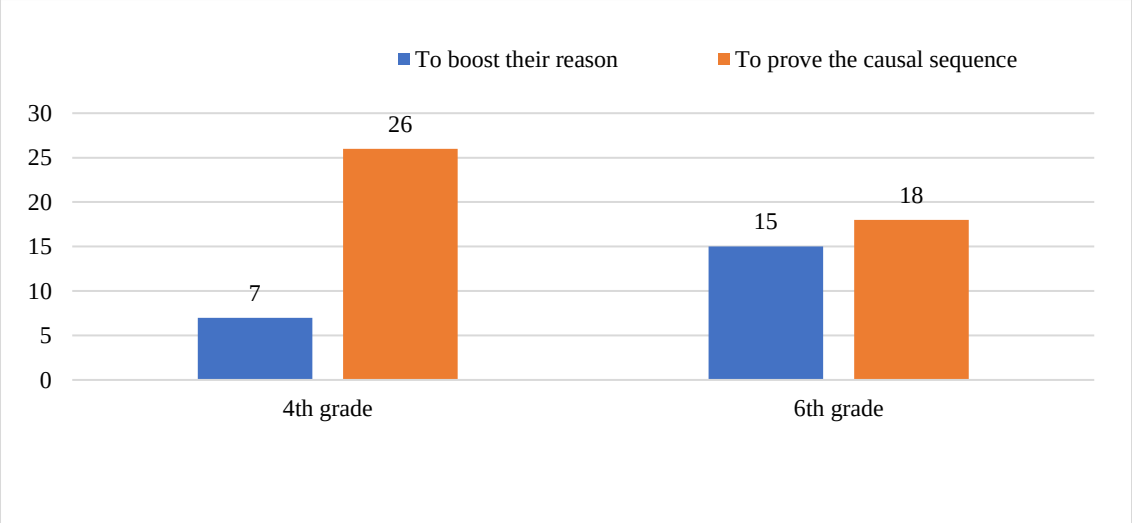

Figure 3. Distribution of Frequencies as Trends in providing evidence to boost the reason or to prove the causal sequence, by grade.

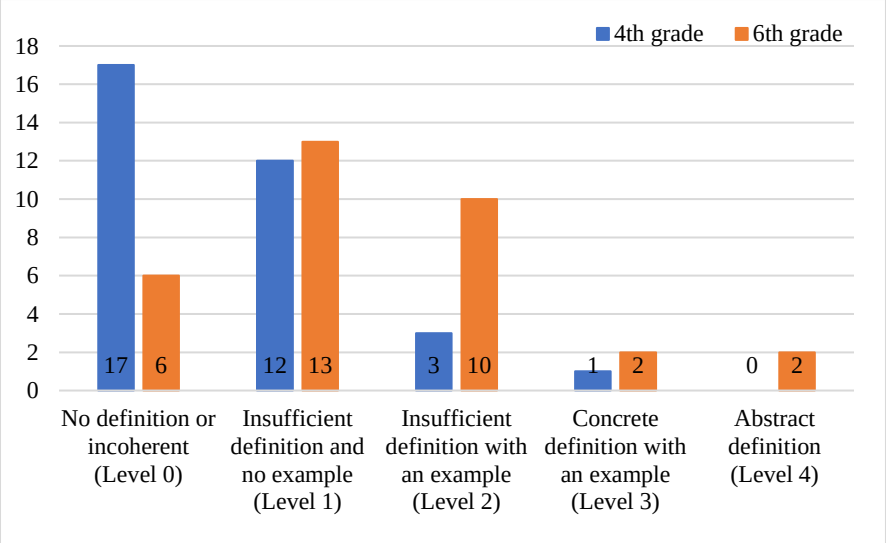

Figure 4. Distribution of Frequencies for the level of concept of evidence by grade

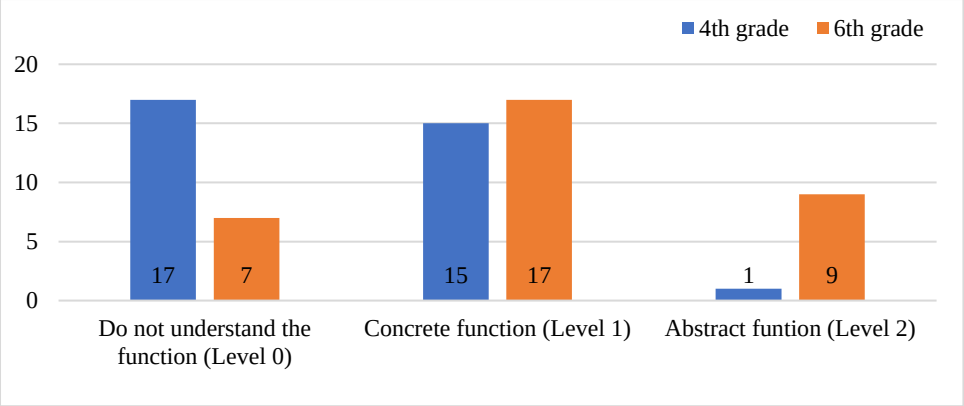

Figure 5. Distribution of Frequencies for the comprehension of the function of evidence by grade