Elementary Education Online, 2018; 17(3): pp. 1476-1495
İlköğretim Online, 2018; 17(3): s.1476-1495. [Online]: http://ilkogretim-online.org.tr
doi 10.17051/ilkonline.2018.466372

A Review of the Assessment Tools for the Student-Led Cognitive Outcomes/Contributions in the Sense of Inquiry-based Teaching

Öğrenen-temelli Bilişsel Çıktıların/Katkıların Araştırma-Sorgulama Temelli Öğretim Bağlamında Değerlendirilmesine Yönelik Araçların Derlemesi

Yılmaz SOYSAL, *İstanbul Aydin University, Faculty of Education, yilmazsoysal@aydin.edu.tr*

Abstract. The purpose of this review is to summarise an array of tools both for science teachers and particularly for science teacher educators to reconsider student-led cognitive outcomes that are initiated and maintained within the in-class science inquiry activities. For this purpose, first, the essential characteristics of the inquiry-based teaching are described and Bloomian taxonomies in assessing the student-led outcomes and student-led intellectual contributions to the classroom discourse is interrogated. Based on the multifaceted and social-interactive characteristics of the inquiry-based teaching, four assessment tools are displayed, justified and exemplified for the pedagogical purposes of the science teaching and learning. The assessment tools are gathered from six different groups of scholars' research efforts. As a whole, the tools are able to assess the quantitative-qualitative student-led outcomes, the students' capacities of operating inquiry skills and practices, the students' abilities to attain evidence-based reasoning and the students' capabilities to generate varying degrees of argumentation. Concrete and fictional instances and potential in-class uses of the offered tools are clarified for the science educators and science teachers.

Keywords: Inquiry-based teaching, cognitive outcomes, intellectual contributions

Öz. Bu derlemenin amacı, fen eğitimcileri ve fen öğretmenleri için fene dayalı araştırma-sorgulama süreçlerinde başlatılan ve sürdürülen öğrenen-temelli bilişsel katkıların değerlendirilebilmesini sağlayacak araçları özetlemektir. Bu amaçla öncelikle, araştırma-sorgulama temelli fen öğretiminin temel özellikleri tanımlanmış ve öğrenen-temelli bilişsel çıktı ve katkıların değerlendirilmesi için sunulan Bloomcu taksonomiler sorgulanmıştır. Araştırma-sorgulama temelli öğretimin çok yönlü ve sosyal-etkileşimli yapısı temel alınarak dört farklı değerlendirme aracı sunulmuş, gerekçelendirilmiş ve fen öğrenmenin ve öğretmenin pedagojik amaçları bağlamında örneklendirilmiştir. Değerlendirme araçları, altı farklı araştırma grubunun araştırma-temelli çabaları sonucunda geliştirilmiştir. Sonuç olarak, değerlendirme araçları, öğrenenlerin nicel-yönelimli ve nitel-yönelimli bilişsel katkılarını ve çıktılarını değerlendirme, öğrenenlerin araştırma-sorgulama beceri ve pratiklerini gerçekleştirebilme kapasitelerini değerlendirme, öğrenenlerin delil-temelli akıl yürütme becerilerinin değerlendirilmesi ve öğrenenlerin argüman kurma ve çürütme yönüne kapasitelerinin değerlendirilmesi yönünde etkili bir biçimde kullanılabilir. Değerlendirme araçları ile gerçekleştirilen örnek değerlendirmeler ve sınıf-içi kullanım potansiyelleri fen eğitimcileri ve fen öğretmenleri için somutlaştırılmıştır.

Anahtar Sözcükler: Araştırma-sorgulama temelli öğretim, bilişsel çıktılar, entelektüel katkılar

Received: 13.03.2018 Accepted: 02.07.2018 Published: 31.07.2018

INTRODUCTION AND PROBLEM STATEMENT

The purpose of this review is to provide an array of tools for science teachers and for the science teacher educators to reconsider the student-led cognitive outcomes in the sense of science inquiry activities. If the inquiry-based teaching has been well proved in augmenting the student-led outcomes in the science classroom (e.g., Akkus, Gunel, & Hand, 2007; Crawford 2000; Furtak et al., 2010; Gunel, 2006), *the measurement and evaluation of the student-led outcomes* should be attained in a sense that is considerably appropriated for the inquiry-oriented student-led cognitive contributions. To be clear, a learning outcome clarifies the specific statements displaying what students will know, value or be able to do by the end of, for instance, inquiry-based processes (Biggs, 1992; 1993). The student-led cognitive contributions to the inquiry sessions may be the assessable end-products or written forms of student-led perspectives or reflections. As the inquiry-based teaching and learning environments are rather sophisticated and multifaceted, to be assessable and measurable, the analytical or holistic cognitive contributions of students must specify things that can be observed (Collis & Davey, 1986; Panizzon, 2002), that are public, common and shared (Mercer, 2004; 2010), and not activities or states that are only internal to students' own minds (Driscoll & Wood 2007).

In educational research and particularly in the science education, Bloomian taxonomies (e.g., Bloom, Englehart, Furst, Hill, & Krathwohl, 1956) in assessing student-led outcomes have been become salient. One of the close colleagues and contemporaries (David R. Krathwohl) of the Benjamin Bloom defined four overarching features of the taxonomy:

1. common language about learning goals to facilitate communication across persons, subject matter, and grade levels;
2. basis for determining for a particular course or curriculum the specific meaning of broad educational goals, such as those found in the currently prevalent national, state, and local standards;
3. means for determining the congruence of educational objectives, activities, and assessments in a unit, course, or curriculum;
4. panorama of the range of educational possibilities against which the limited breadth and depth of any particular educational course or curriculum could be contrasted (Krathwohl, 2002, p. 212).

In the Bloomian classification system, there are six categories describing the student-led cognitive outcomes as sequenced consecutively: knowledge, comprehension, application, analysis, synthesis, and evaluation (Bloom et al. 1956). Bloom and his colleagues (1956) asserted that their classification system can provide substantial tools for operational definitions of the thinking processes of the students in a hierarchical, cumulative and sequential manner within a wide-ranging spectrum (e.g., from a domain to the other, or from a subject or grade level to the other).

Several critics have been raised regarding the Bloomian taxonomy in assessing student-led outcomes, however (Sugrue, 2002), particularly in the context of the science inquiry. Even though the Bloomian taxonomy have been pervasive in the educational field of inquiry over 50 years (e.g. eventually being translated into 22 languages), it has not still been rigorously researched to produce materialistic evidences of its application quality (Sugrue, 2002). As Sugrue (2002) acclaimed that the concrete distinctions have not made between either of the two lowest stages (knowledge or comprehension) or between the four highest stages (application, analysis, synthesis, and evaluation).

The Bloomian taxonomy has also been, therefore, criticised pertaining its narrower dimensionality. To explicate, knowledge and comprehension stages are considered as lower-order thinking skills. Other four stages as application, analysis, synthesis, and evaluation are conceived as the higher-order thinking skills. This splits the taxonomy into two binary dimensions by excluding the gradual cumulation of the student-led cognitive contributions.

In response to several criticisms made, Anderson, Krathwohl, Airasian, Cruikshank, Mayer, Pintrich, Raths and Wittrock (2001) revised the original taxonomy and added a novel dimension entitled as "the knowledge dimension" involving "factual, conceptual, procedural, metacognitive" knowledge. Moreover, the cognitive process dimension was subjected to the changes as "create" stage was added (Anderson et al. 2001). Thus, in the recent form of the taxonomy, there is a move from one dimension (the cognitive process dimension) to two dimensions (including the knowledge dimension).

The so-called unique taxonomy has been still suffering from assessing inquiry-based operations such as making rough inductions, deductions and generating arguments, however. Making rough inductions, deductions and engaging in negotiations and argumentations are inherent to the inquiry-oriented science activities. However, as mentioned, these skills of reasoning may not be truly and thoroughly assessed only by taking the Bloomian taxonomies into account (Ennis, 2002; 2004; 2011; Facione, 1990). In this sense, particularly Robert Ennis raised his concerns regarding the instructional potentials and credibility of the Bloomian taxonomy in assessing the student-led outcomes. For instance, Bloomian taxonomies have limitations when it comes to assessing critical thinking skills (Ennis, 2002; 2004; 2011) as one of the desired aspects of the student-led outcomes. In addition, Ennis (2011) asserted that educators must exceed beyond the Bloom's taxonomy to re-conceive particular abilities and dispositional characteristics of the presumable critical thinkers as the students who are engaged in the science inquiry. As a whole, there should be additional alternatives in assessing student-led cognitive outcomes as addressed, justified and exemplified in the current study.

Second, in the context of assessing student-led cognitive outcomes, there has been an ongoing paradigm war. To explicate, cognitive outcomes has been explored through either *process-product paradigm* or *sociolinguistic paradigm* (Brophy & Good, 1986). Proponents of process-product paradigm consider cognitive outcomes as a function of teacher behaviours (Carlsen, 1991). Regularly, "this is done by constructing a taxonomy of teacher behaviours, counting (or experimentally manipulating and counting) these teacher behaviours over one or more lessons, and then correlating cumulative counts with individual student or pooled-class outcome measures." (Carlsen, 1991, p. 157-158). Thus, within the context of the process-product paradigm, cognitive outcomes (e.g., what students will know, value or be able to do by the end of the inquiry-based alignments) are assessed through nation-wide examinations and their context, contents and embedded curricular objects are predetermined by considering Bloomian-like taxonomies.

Furthermore, sociolinguistic paradigm considers *student-led cognitive contributions* in a different sense. Sociolinguistic research paradigm regards classrooms' contextual entities through uses of diverse methodological approaches such as *conversation analysis* and *discourse analysis* (Mercer, 2004; 2010). To explicate, an assessment of the student-led cognitive outcomes can be attained by taking the in-class teacher-student interactions into account. In a temporal sense, the analytical, fine-grained and in-depth analysis of the discourses created in the science inquiry can be more illustrative to what extent and how the students are able to intellectually contribute to the classroom discursive exchanges and interactions. Thus, process-product paradigm proponents carry out an assessment of the cognitive outcomes *out of the classroom as an end-product* (Mercer, 2004; 2010). Moreover, the proponents of the sociolinguistic paradigm conduct an assessment of the student-led intellectual contributions *in the classroom as a temporal, emerged, shared and created product* (Mercer, 2004; 2010).

As Carlsen (1991) asserted, although two distinctive paradigms concern themselves with divergent issues regarding the assessment of the students' cognitive outcomes and contributions, they may inform each other. With this rationale, in the current study, both paradigms' methodological lenses are used to capture a holistic understanding pertaining the assessment of the student-led cognitive outcomes (process-product paradigm) and intellectual contributions (sociolinguistic paradigm) emerged during science inquiry sessions. To put it differently, *out-of-classroom* (product) and *in-the-classroom* (temporal process) applications of the proposed assessment tools are provided for external reader to lead them to make more concrete, valid and holistic assessments of the student-led cognitive outcomes *after the inquiry activities* (process-product paradigm) and student-led intellectual contributions to the classroom discourse *within the inquiry sessions* (sociolinguistic paradigm).

Significance of the Study

This review incorporates several aspects ensuring its contributions to the science education. At the outset, this review presents alternative tools for assessing cognitive outcomes/contributions. The Bloomian taxonomy has been a domineering tool, however, this study shows that there may be other respectively more functional and instrumental tools to make assessments of cognitive outcomes. Moreover, this study aimed at making the instructional uses of the proposed tools visible to the external readers by providing concrete or fictional examples. This was needed for concretising each abstracted tool that would be more transparent for the external reader. Beyond, as mentioned earlier, there has been a main tension between the process-product paradigm and sociolinguistic paradigm. This study displays concrete signs and traces how two paradigms' points of views on the cognitive assessment can be merged with together for the sake of more informative assessments of student-led outcomes in the sense of science teaching and learning.

Describing Teaching by Inquiry and Contextualised Student-led Outcomes

In this section, essential characteristics of teaching by inquiry are documented to justify the necessary uses of the alternative assessment tools for the cognitive outcomes and intellectual contributions that are anticipated to be emerged during the science inquiry sessions.

Current national science education reforms stated that teachers should create inquiry-based learning environments by supporting student-led inquiries and interacting with students in the presence of authentic question generation (National Research Council, [NRC], 2012). Science education reform efforts (ACARA, 2013; NGSS Lead States, 2013) have also indicated, among others, two essential parts of science teaching: conducting inquiries in the science classroom and generate science arguments. These two practices are associated with scientific practices referring to the scientists' work and ways of students' learning in science classrooms (NRC, 2012). The common works of students and scientists were set in *Framework for K-12 Science Education* (NRC, 2012) and incorporate undertaking investigations by collecting, analysing and interpreting data and establishing evidence-based arguments around the inquiries. There are essential features of inquiry classes illuminating which elements of cognitive outcomes should be featured and assessed. The indispensable features are listed and justified in terms of authentic and case-based assessments of cognitive outcomes

(1) As acknowledged, the student-led voices should also be dominative and contributing for creation of common and shared knowledge (Mercer, 2010) in the sense of teaching through inquiry. Thus, both in a quantitative and qualitative sense, the proportions or frequencies of the student-led voices must be increased, varied and proliferated (Martin & Hand, 2009; McNeill & Pimentel, 2010; Pimentel & McNeill, 2013). Quantitative aspect of cognitive outcomes refers to *countable amounts* of students' verbal contributions to the inquiry activity (Lefstein, 2008). In an inquiry activity, students may provide short responses that those are pitched at recall; lower-order stages (Chin, 2006; 2007). These short responses can be "Yes", "No", "I agree with you." The short responses of students may take a few seconds to utter (Lefstein, 2008).

Furthermore, not only the quantities, but also the qualities of the student-led voices should be improved through the authentic inquiry-oriented in-class sessions (Brown, et al. 2010a, 2010b; Furtak et al., 2010; Hardy, et al. 2010; Shemwell & Furtak, 2010). In this sense, students may be able to consider more than one aspect of the concept under negotiation and make relations between being considered aspects for more extended responses, in turn, attain *generalizations.* Thus, the quantitatively-oriented and qualitatively-oriented cognitive contributions of students should be defined within a measurable or countable manner to reveal the impacts of inquiry-based activities on cognitive outcomes.

(2) There are many science concepts that are experienced by the students during the inquiry activities. The concepts of scientific inquiry can be best conceptualised, comprehended and transferred to the external contexts by students when they are engaged in operating skills of inquiry (Benedict-Chambers et al. 2017). During a typical inquiry activity, students may operate both hands-on (e.g., theory-laden observation, measurement, comparison, investigation design) and minds-on (explanation, induction, deduction, argumentation) skills (Grimberg & Hand, 2009). For instance, the students may be guided to gather, analyse and interpret data when answering their research questions (Cavagnetto, 2010; Cavagnetto, Hand, & Norton-Meier, 2010; Cavagnetto & Hand, 2012). In a similar vein, students have opportunities to experience several cognitive pathways. They may make observations, measurements and comparisons. They may provide exemplifications of the phenomena under discussion. They may compose analogies and made clarifications to communicate with others. They establish cause-effect relations to propose scientific explanations. They may construct evidence-based explanations derived from their own inquiry. They may make judgements on others' claims, arguments, observation reports and the credibility of the data sources. They also achieve inductive and deductive inferences from the discourse. They also offer newer investigation designs to revise their peers' experimental thinking.

As a whole, when learning science concepts and scientific practices are considered in an isolated manner, science contents can be acquired through rote learning or memorisation (Benedict-Chambers et al. 2017). Thus, it is an imperative for inquiry sessions in which teachers and students collectively use scientific practices (hands-on) to develop meanings (minds-on) of phenomena. In conclusion, it would be progressive to recommend an assessment tool to make theoretical and practical cognitive outcomes of students to inquiry processes transparent for science teacher and science teacher educator.

(3) When the students collect and analyse data in an inquiry activity, it is also expected that the students must create their own evidences derived from gathered data (Crawford, 2000; Cavagnetto & Hand, 2012). It is where the reasoning qualities of the student-led utterances come in. In the science education literature, there has been an ongoing controversy in adjusting the interrelations between the *data, evidence* and *reasoning* that is attached with the purposes of the current study in recommending an assessment tool for predicting cognitive outcomes.

As aforesaid, in the inquiry activities, the students should be promoted to pose their own research questions, then, collect, analyse and interpret data. Thus, generic inquiry implementations may be carried out in an argument construction sequence as *"questions > claims > evidences"* (Cavagnetto & Hand, 2012). In other words, in the activities, the students pose researchable questions, then, propose their claims and create their evidences for their hypothetical pre-claims (Crawford, 2000). There are of course alternative triplet cycles (e.g., *claim > evidence > reasoning*) that put a clear *isolation* between evidence and reasoning (McNeill, 2009; McNeill & Krajcik, 2008; McNeill, Lizotte, Krajcik, & Marx, 2006).

In the NRC's (2007) *Ready Set Science*, claim, evidence and reasoning are explained as: *Claim:* What happened, and why did it happen? *Evidence:* What information or data support the claim? *Reasoning:* What justification shows why the data count as evidence to support the claim? (p. 133). This triplet (claim > evidence > reasoning) have supported and applied by many scholars (McNeill, 2009; McNeill & Krajcik, 2008; McNeill, Lizotte, Krajcik, & Marx, 2006; McNeill & Pimentel, 2010; Pimentel & McNeill, 2013). It is clear in the NRC's (2007) documentations that reasoning occurs only at a defined point of the science inquiry processes. This implies that after collecting appropriate and sufficient evidences, students may be engaged in reasoning processes (McNeill, Lizotte, Krajcik, & Marx, 2006). However, reasoning must be throughout as a critical aspect of entire processes of inquiry activities. As a result, reasoning might be undervalued to some extent if the evidence and reasoning are isolated (Cavagnetto & Hand, 2012).

In aforesaid context, the controversy is that how it is possible to merge with reasoning and evidence. The answer finds itself in the relationship between data and evidence in terms of this study. In this study, data is accepted as the student-led observations and experimental-based recordings. In the inquiry implementation, what is seen and recorded is commonly accepted as being the data obtained from student-led inquiries. However, there should be a question to be asked that how the students would use the data? In a common sense, to our knowledge, *the data does not speak*. In other words, there must be "a data analysis procedure on the part of the students' to produce original evidences for the posed claims (Cavagnetto, 2010; Cavagnetto & Hand, 2012). Thus, there should be a transformation of data into evidence requiring specific types of reasoning (Cavagnetto & Hand, 2012). This shift is required a *cognitive work on the part of the students* that is termed as reasoning quality in this study. In other words, as Cavagnetto and Hand (2012) states:

> "A student has to analyse and synthesize the data points into some coherent series. There are critical decisions that need to be made such as what to keep, what to discard, and how well the data points are connected. That is, data does not speak and so the learner has to apply some critical thinking and reasoning to be able to make decisions to produce the required evidence he/she needs to make an argument." (p. 46).

In this sense, within the scope of this study, both *data and reasoning* is conceived as *evidence*. Thus, it has been an imperative to assess the student-led data collection, analysis and interpretation processes in the context of producing their own evidences in supporting or falsifying prior claims regarding the questions posed.

(4) Once the students produce their evidence-based claims by pondering on gathered data, it would be time to negotiate the validity and reliability of the generated evidences as in the form of arguments. In inquiry-based implementations, it is more possible and potential to observe that students produce counter-arguments, rebuttals, alternative explanations against to their classmates' claims (Cavagnetto, 2010; Cavagnetto, Hand, & Norton-Meier, 2010; Cavagnetto & Hand, 2012). In this context, teachers should monitor and collide with alternative findings (arguments) and contradictory explanations (arguments). In a specific sense, it is necessary to prompt students for researching into alternative or contrasting research questions to augment the scope of the negotiations of meanings (Soysal & Radmard, 2017; 2018). It is the routine of productive inquiry-based implementations (Cavagnetto & Hand, 2012). To support, as Cavagnetto and Hand (2012) summarized "when procedures are uniform for all students, where data are similar and where claims match expected outcomes, then the reporting of results and conclusions often lacks opportunities for deeper student learning about the topic or for developing scientific reasoning skills." (p. 48). In other words, when science teachers guide students to alternative research questions that are varied in terms of being examined research variables, there would be more counter-arguments and alternative experimental inferences that may be smashed during whole group discussions enhancing the coordination of the data and claim. As a whole, it would be invaluable to assess the student-led arguments' qualities that can be acknowledged as possible cognitive outcomes.

In the summary, four essential aspects of in-class inquiry-based implementations should be taken into consideration both in triggering and maintaining the teaching through inquiry and in assessing the potential cognitive outcomes or intellectual contributions. Thus, four different, but interrelated assessment tools are introduced and exemplified in the sense of above-located characterisations of inquiry-based implementations. In the next section, at the outset, methodological foundations of the study will be presented.

Methodology: Systematic Selection of the Related Studies

In this section, two specific features of the methodology of the current study will be presented and justified. The two prominent specifications are *conceptual framework* in including and/or excluding the studies that were subjected to a systematic review and *procedural framework* incorporating technical processes in collecting, analysing and interpreting the selected studies.

Conceptual framework: For a systematic review or locating the studies in favour of hypothetically-based assertions attained in this study, the basic criterion was to clarify "eligibility". Eligibility refers to the *theory-laden* or *intervention-based appropriateness* of the selected studies that are thought to be included in a study or which studies will be excluded from the systematic review (Abrami, Cohen & d'Apollonia, 1988). For many systematic and purposeful reviews, the most important question that a researcher should ask herself or himself is to which studies are more potential or eligible in including to the pool of the studies (Gliner, Morgan & Harmon, 2003; Lin, Lin ve Tsai, 2014; Suri & Clarke, 2009). One of the surrounding eligibility criterion can be deduced from *operational definitions* of concept(s) under examination (Abrami, Cohen and d'Apollonia 1988).

In this study, four featured concepts or *themes* had framed the researcher's mind to select or exclude a study during retrieving processes. These themes are operationally defined within above section and can be listed as "*countable amounts of cognition*", "*cognitive pathways*"; "*argument quality*" and "*reasoning typologies*". The inclusion of a study was truly decisioned whether the study gets in touch with any of the predetermined themes. To put it differently, four themes have been accepted as the fundamental characteristics of an inquiry-based in-class process as justified above. As a rational, therefore, *a fine-grained griddle* was composed to filter the proper studies from the irrelevant or unrelated ones.

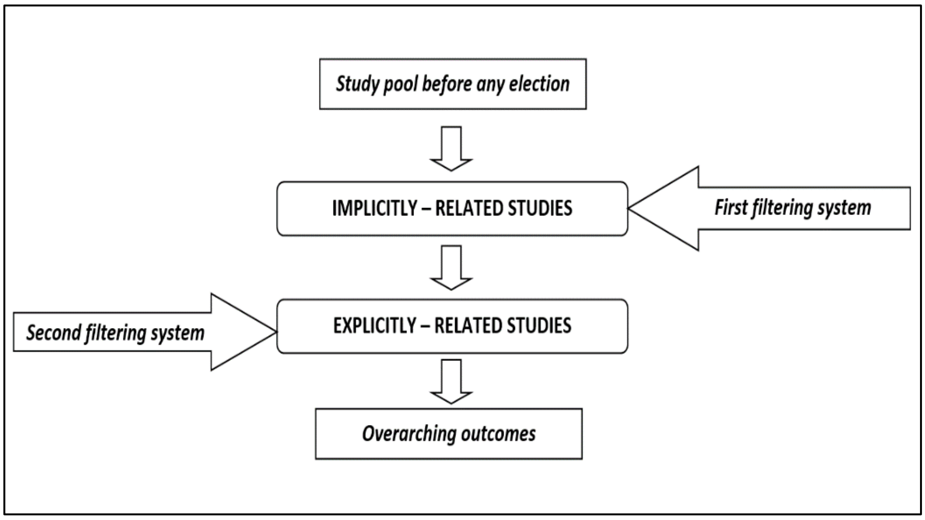


Figure 1. Phases for conceptualising the related studies

In this context, the pooled researches for the current study can be sorted in two classes as "explicitly-related studies" and "implicitly-related studies" (see also Figure 1). Implicitly-related studies provide the conceptual, factual and epistemological tools to comprehend and examine the ways in which students are able to contribute to classroom discourse during inquiry sessions (e.g., Ennis 2002; 2004; Facione 1990). Explicitly-related studies are the selected research to inquiry into in order to determine which types of concrete tools can be used and applied for assessing student-led cognitive contributions and outcomes emerged in science inquiry (e.g., Furtak et al., 2010; Shemwell & Furtak, 2010). In a sense, implicitly-related studies were the predictors and the indicators of the main studies as explicitly-related studies. In "References" section implicitly-related studies are marked by (*) and explicitly-related studies are flagged by (**). Totally, 36 studies ($n_{indirect}$ = 17; n_{direct} =21) were incorporated in this study to establish a broader picture in a thick forest. In a sense, implicitly-related studies guided the researcher for looking into more convenient sources of the tools of identifying student-led cognitive outcomes or contributions. In other words, the conceptualisations embedded in the implicitly-related studies were functionalised as the initial filtering systems for selecting or ignoring the studies (see also Figure 1). Explicitly-related studies therefore the outcomes of the in-depth analysis of more concrete and essential studies devoted to a direct exploration of the cognitive outcomes or intellectual contributions of students (see also Figure 1).

Procedural framework: This review was involved studies exploring assessment of student-led outcomes in an implicit or explicit manner. Selected studies were mostly comprised a fine-grained analysis of the ways of assessing the cognitive outcomes. Specific procedures were operated for obtaining the most relevant studies. In searching of related studies, computerized data bases and functional digital operators (e.g., ERIC; Boolean Operator) were used to filter out the appropriate studies.

The search was conducted in 2017 through considering specific keywords: "cognition", "attainment", "achievement", "assessment", "evaluation", "inquiry", "outcomes", "intellectual", "contribution", "tool" or other synonyms or related terms were used in a combined, systematic and pragmatist manner. Primary and secondary references were limited to Academic Journals and Extended Reports. The author accounted for the diversity regarding types of the determined Journals to grasp the different scholar-led voices regarding the phenomenon under consideration.

For a systematic sampling of the current research on the assessment of cognitive outcomes, the author strictly took two aspects of the selected studies into account. At first, selected studies should be devoted to improvement the theory of science education. Secondly, the studies were particularly selected by checking a criterion whether they explored any sets of tools for assessing the cognitive outcomes in an explicit manner.

It was also a matter of selection whether the pooled studies incorporated diversifying participants as students who were varying in terms of academic grades such as secondary science classrooms or middle school level. Finally, techniques of analysis of the cognitive outcomes taken by the pooled studies were another criterion. To explicate, some studies extracted cognitive outcomes by analysing episodes in an interpretivist sense and other studies operated (lag) sequential analysis techniques to attain a systematic observation through coding-counting.

The systematic determination of the studies serviced two purposes. Firstly, there was a better sampling of the related studies that were considerably representative as the selected works reflected both past and current streaming of the research on the assessment of the cognitive outcomes. Secondly, the systematic approach was useful in re-categorising the detected assessment tools around newly invited theoretical frames, thus, incorporated a pragmatist approach in determining and analysing an intensifying research area.

Descriptions of the Measurement Tools in the Context of Teaching by Inquiry

In this section, four assessment tools are described and exemplified. The assessment tools are entitled as "The SOLO Taxonomy", "Cognitive Pathways", "Reasoning Typologies" and "Argument Structures". Each assessment tool is devoted to a particular aspect of assessing the cognitive outcomes and contributions.

The SOLO Taxonomy and Student-led Cognitive Contributions

The Structure of Observed Learning Outcomes (SOLO) represents the levels of progressively complex understanding of students through *five general stages* that are intended to be relevant to all subjects of all disciplines (Biggs & Collis, 1982). In the SOLO, cognitive contributions of students are conceived as an increasing order in number and complexity of connections. Indeed, the SOLO is a hierarchical model and suitable for assessing learning outcomes of different subjects, levels and diverse lengths of assignments (Biggs & Collis, 1982; Chan et al., 2001).

Table 1. The SOLO Taxonomy

Levels of learning stages	Levels of understanding	Descriptions
Stage of Ignorance	Pre-structural	The task is engaged, but the learner is distracted or misled by irrelevant aspects or information; nothing meaningful has been learned.
Stages of surface learning	Uni-structural	The learner focuses on the relevant domain and picks up one aspect to work with; one specific thing has been learned.
	Multi-Structural	The learner picks up more and more relevant or correct features, but does not integrate them; several relevant, independent and meaningful aspects have been learned.

Stages of deeper learning	Relational	The learner now integrates the parts with each other, so that the whole has a coherent structure and meaning; aspects learned are integrated into a structure.
	Extended abstract	The learner now generalizes the structure to take in new and more abstract features, representing a higher mode of operation; aspects learned are generalized to a new domain

The SOLO incorporates three levels of learning stages: *level of ignorance, levels of surface learning,* and *levels of deeper learning* (Table 1). The levels of learning stages are characterized by five levels of understanding: *pre-structural, uni-structural, multi-structural, relational, extended abstract*. Pre-structural stage is indeed outside of the taxonomy as it refers to stage of ignorance (Biggs & Collis, 1982). In pre-structural level, the students present an irrelevant or false utterance [e.g., "*The weight is the mass*."].

In uni-structural and multi-structural stages, there are true cognitive contributions of students. In these two levels, only quantity of cognitive contributions are more of an issue. To explain, students may cognitively contribute by displaying many aspects of negotiated science concepts. For uni-structural stage, only one aspect of the topic is provided ["*Weight is the load*."]. For multi-structural stage, students may declare several aspects of the phenomenon ["*Weight means multiplying the mass with gravity, I mean 10. That equals the weight. For example, if this is 50 gr; we can learn its weight-newton by multiplying it with 10. But if it is gr, we should first convert it to kg; then multiply it with 10 and learn its newton. We should convert first*."]. However, this does not mean an interconnected or related style of the representations of facts. In other words, even though students provide many aspects of the phenomenon, they are liable to talk about the unconnected or isolated parts of it. There is only a *quantitative* increase in the cognitive contributions of students, but, there is not a qualitative shift.

In the last two levels of the SOLO taxonomy, cognitive contributions are assessed not only from a quantitative sense but also qualities of the contributions are regarded. Last two stages (i.e., relational, extended abstract) assess the abstraction levels of cognitive contributions. For instance, in the relational stage, students provide interconnected ideas. Students make connections between their ideas by presenting them in a cohesive manner instead of an isolated sense [*So, we continue to face difficulties. Since there is no fulcrum, our work doesn't get easier like a simple machine! (talks about his/her own experiments)*]. To compare, in the multi-structural level, the students present, for instance, five distinctive pieces of information about force concept. Then, students attain an abstraction by attaching three distinctive parts of the previously presented five aspects into one coherent line of reasoning. There is therefore a qualitative shift in students' cognitive contributions. Because, students are now able to present a coherent utterance by collapsing and reducing several ideas into a *generalized* one.

For the extended abstract, at the outset, students may make several genuine relations between the proposed ideas. However, this is not adequate for an extended abstract. Actually, students have to attain *inductive reasoning* requiring enlarged generalizations. It refers that students should move beyond the related conceptions by constructing generalized statements for transcendental contexts ["*There can't be a lever without a fulcrum! All levers have a fulcrum. Otherwise how can we lift! We can't lift anything without a fulcrum*."]. Put it differently, Potter and Kustra (2012) indicated the way of moving from relational ideas to extended abstracts as "practice with synthesis and evaluation will help students develop greater understanding of relationships between ideas and the reasons things are done a certain way, etc., and as they are forced to use this knowledge in increasingly unfamiliar, varied, situations, their ability to generalize and adapt will grow." (p. 14).

Several researchers confirmed the SOLO's comprehensiveness and objectiveness in terms of assessing, for instance, cognitive contributions (Chick, 1998; Lake, 1999). The SOLO has also been extensively used in many disciplines in assessing cognitive attainments of the students such as in biology, mathematics and language (Chick, 1998; Lake, 1999). Chan et al. (2001) revealed a positive correlation between the SOLO scores and writing styles (language), learning strategies, learning motivations, gained grades, prior academic competences. It signifies that if a student has a better, for instance, writing style or learning strategy, it can be estimated by the SOLO. In the context of science teaching and learning, particularly Chin (2006; 2007) offered scholars to apply the SOLO in assessing cognitive contributions. Based on the recommendation of Chin (2006; 2007), van Booven (2015) showed the usability and objectiveness of the SOLO in assessing cognitive contributions emerged in inquiry-based activities.

Cognitive Pathways of Student-led Cognitive Contributions

Grimberg and Hand (2009) developed a qualitative analysis method to look through cognitive operations and devised an assessment tool (Appendix-1: Cognitive Pathways). For the purposes of an inquiry-based implementations, this analytical assessment tool (Grimberg & Hand, 2009) is guiding and informative, yet insufficient. As detailed later in this section, there are more enlarged cognitive pathways of the students emerged in their talks while engaging in the inquiry. Thus, other cognitive pathways should be added to the catalogue.

The integration or re-synthesis is the original contribution of the present study to the related theoretical frameworks. The developed codes (analytical cognitive pathways) applied in the study of Grimberg and Hand (2009) are substantially matched with the *critical thinking skills* defined by *Robert Ennis* (Ennis, 2002; 2004; 2011; Appendix-2) and described by *Peter A. Facione* (Facione, 1990's Delphi Report; see also Appendix-3). This therefore leaded the author to re-embrace critical thinking skills (CTS) in the sense of teaching by inquiry to make the scope of the catalogue more comprehensible for the purposes extended inquiry-based implementations.

CTS, stated in two jurisdictions' executive documentations (Ennis, 2011; Facione, 1990's Delphi Report), were embedded in the initial catalogue of Grimberg and Hand (2009). Ennis (2011) originally defined 12 characteristics of CTS and collapsed them into five higher-order categories (Appendix-2). These categories are *Basic Clarification* (e.g., ask and answer clarification and/or challenge questions); *Decision-Making* (e.g., judge the credibility of a source); *Inference* (e.g., make material inferences as rough induction); *Advanced Clarification* (e.g., making operational definitions to clarify the meaning); *Supposition and Integration* (e.g., self-examination; self-correction). Ennis's (2011) revised catalogue of CTS is the result of an ongoing research.

In a similar vein, Facione (1990) revealed the sub-skills and main skills of CTS within a Delphi study. These are *Interpretation* (e.g., categorization, decoding significance, clarifying meaning); *Analysis* (e.g., examining ideas, identifying arguments, analyzing arguments); *Evaluation* (e.g., assessing claims, assessing arguments); *Inference* (e.g., querying evidence, conjecturing alternatives, drawing conclusions); *Explanation* (e.g., stating results, justifying procedures, presenting arguments); *Self-regulation* (e.g., self-examination, self-correction). The author, therefore, deduced that three portrayals of cognitive pathways (Ennis, 2011; Facione, 1990; Grimberg & Hand, 2009) could be combined for a more pragmatist and systematic examination of cognitive contributions. Grimberg and Hand's (2009) assessment tool provides inquiry-related cognitive pathways as in the form of hands-on and minds-on practices. Ennis' (2011) and Facione's (1990) assessment tools allow to expand the embedded aspects of Grimberg and Hand's (2009) assessment tool.

Table 2. Cognitive Pathways of Student-led Responses

Cognitive Pathways	***Analytical Codes***	***Indicators***
Perception	Observation	Stating/sharing experienced-based, practical-based or individual-based simple observations, personal experiences, or data that result from students' observations
	Exemplifying	Introducing relevant examples, instances, samples, trials of events, concepts of content under discussion; sharing/stating variables effecting other variables
	Measurement	Reference to any quantitative aspect of the data, stating proportions, making simple calculations by using observed or gathered quantitative data

	Compare	Reference to common/different characteristics of two or more pieces of data or objects; comparing two events/situations/sayings/measurements
Conception	Analogy	Mapping elements from a source domain (well-understood situation) into a target domain (non-familiar situation)
	Basic clarification (low level interpretation)	Questions that stimulate clarification supporting other operations; pieces of knowledge and information that stimulate clarification supporting other explanations/sayings/arguments/opinions/positions/ideas, etc.; decoding significance; simply clarifying meaning within discourse
	Advanced clarification (high level interpretation)	Define/redefine terms, concepts, definitions by using appropriate criteria, examples, instances, samples; making operational definitions to clarify the meaning, arguments, sayings, claims, assertions, etc.
	Cause/effect	Identification of a cause and its effect
Abstraction	Explanation	Offering unproved inference without data/evidence; interpreting/stating/inferring from data-based results, interpreting/stating evidence-based results; justifying procedures (e.g. experimental); justifying experimental/observational data
	Judgement by…	Assessing claims, assessing arguments; judge observation reports; judge the credibility of a source; examining ideas; judge deduction; judge definitions
	Inference-I: induction	Reasoning that links few examples to general premises; make material and generalized inferences (roughly "induction")
	Inference-II: deduction	Reasoning that links general premises to a specific; deduce
	Investigation design	Planning new experiments; stating new ways of investigating/representing/solving a problem/situation; stating new ways of collecting, analyzing and interpreting data
Dispositions	Supposition	Consider and reason from premises, reasons, assumptions, positions, and other propositions with which they disagree or about which they are in doubt, without letting the disagreement or doubt interfere with their thinking
	Integration	Integrate the dispositions and other abilities in making and defending a decision
	Self-regulation	Self-examination; self-correction

As an important note, Grimberg and Hand (2009) also advanced three higher-order categories to sort out cognitive contributions: *perception, conception, abstraction.* The perception category incorporates lower-order cognitive processes as observation ["*I saw that the heavier object and lighter object hit the ground at the same time*"], measurement ["*There are about fifty-five millimetres distance between the objects*".], compare ["*Their shapes are different*"] and exemplifying ["*For instance, lever is a simple machine*"]. In this stage, students' cognitive contributions are conceived at the level of sensations.

In the conception level, students may be able to conceptualize and particularly clarify their meanings through their vocabularies. Therefore, *basic clarification* ["*There is light in the room. We can see each other, otherwise we couldn't.*"] and *advanced clarifications* ["*The light both reverberating and dispersing; and at the same time some of it is being absorbed.*"] are the most featured elements of this stage.

Lastly, in abstraction stage, students are able to make evidence-based explanations, attain inductive reasoning [*"We fixed weights of 100 gr to both of them, and when we released them like this (as there is space between them) they stayed like this, there was no change. We tried by increasing the weight little by little (showing the weights in his/her hand). I mean we increased their energy, potential energy. In no way it showed a difference. We measured one by one. We concluded that potential energy cannot be transferred on reels"]* and deductive reasoning, or to offer more technical *investigation designs* for others' experimental procedures *["No, in fact there is gravity,* *but at the same level (s/he means because the forces or weights are equal) so we don't feel it… I mean there is not much difference of height between them, may be because of that they didn't move… Maybe it would change if we could try for infinitely times. If I had the chance to try, I would try like that."*]. In other words, students exceed scopes of conceptions being negotiated during inquiry-based implementation by making evidence-based scientific explanations. In addition, as barrowed from particularly Ennis (2011), another category was added as critical thinking dispositions that can be seen in the bottom of Table 2; *supposition, integration* and *self-regulation* ["*I shifted my idea, because later I thought that it desolates when we put salt in water, and according to my friends…*"]. These are necessary as surrounding critical thinking dispositions that scaffold the occurrences of CTS (Ennis, 1996).

Reasoning Quality of the Student-led Cognitive Contributions

In a typical inquiry-based implementation, it is imperative for students to reason about gathered data to transform them into evidences. Only collected data do not speak up anything about natural phenomena as students should study on the data to generate evidences. When students ponder about data to empower their arguments, they change the data into evidences. Aforesaid, *data plus reasoning* ensures *evidence*. In an in-class inquiry, students should compose researchable questions, make assertions for their questions and generate evidences for their claims. As a rational, a coding catalogue should be incorporated abovementioned skills of inquiry-based implementations.

In a specific issue of *Educational Assessment Journal*, Brown and her colleagues (Brown et al., 2010a, 2010b; Furtak et al., 2010; Hardy et al., 2010; Shemwell & Furtak, 2010) made an array of publications. They presented research-based applications of an ongoing research and *Evidence-Based Reasoning Video Framework (EBR Video Framework)* was one of the most important production of these efforts (Furtak et al., 2010). They offered an alternative way of thinking about *reasoning* particularly for classroom discourses within the inquiry class. They also proposed a tool to map out the classroom talk (Brown et al., 2010a, 2010b). This mapping out system elucidates reasoning phenomenon within a *data plus reasoning is equal to evidence* argument that is relevant for the instructional context and epistemological posture of inquiry-based activities (Brown et al., 2010a, 2010b; Furtak et al., 2010; Hardy et al., 2010; Shemwell & Furtak, 2010).

To compose EBR Video Framework (Furtak et al., 2010), data was collected from the elementary and middle school students' whole-class discussions during inquiry-based teaching about concept of sinking and floating. Sinking-floating is a conceptually challenging science phenomenon for the students. Furtak et al. (2010) indicated the functions of the EBR Video Framework as it "captures teachers' and students' co-constructed reasoning about science phenomena and the quality of the backing for those claims. The EBR-Discourse framework conceptualizes reasoning along a continuum, where the most sophisticated science discourse is conceptualized to consist of claims about science phenomena that are supported by a generalized statement about relationships between properties (a rule). In addition to this statement of a rule, (empirical) backing such as reference to observations (data) or summaries of that data (evidence) may be used to support the claim. The least sophisticated reasoning is considered to consist of a single claim or claims without any form of support." (p. 182).

The EBR Video Framework comprises three sets of codes. The first is the *quality of reasoning* that accounts for "the extent to which claims are backed up with data, evidence, and rules" (Furtak, et al., 2010, p. 182) by students. The other two categories are *teacher's contribution* and *conceptual level*. Only first set of the codes are dedicated to cognitive contributions labelled as *unsupported, phenomenological, relational,* and *rule-based* (Table 3).

Table 3. Reasoning Typologies of Student-led Responses

Quality of Reasoning	***Label***	***Description***
Unsupported	No reasoning	Elements of reasoning present, but no

		processes of reasoning; pseudo, circular, or tautological reasoning
Phenomenological	Data-based reasoning	Data applied to a claim: Partial reasoning structures rely on data or evidence only. Those structures that reference only data or specific phenomena (phenomenological reasoning) as backing for a claim rely on single observations by students
Relational	Evidence-based reasoning	Evidence applied to a claim, including analysis of data in the form of comparisons between properties or summaries of data
Rule-based	Inductive or deductive rule-based reasoning	1. Deductive reasoning (top-down), applying a rule to make a claim with respect to a new premise 2. Inductive reasoning from data to rule 3. Applying a rule with new evidence (exemplifying with analogy) 4. Complete reasoning structure (whole framework)

Unsupported reasoning specifies that there is no reasoning in student-led utterances. This clarifies the least sophisticated cognitive contribution. Unsupported reasoning includes single claim(s) without any backing. Additionally, claims with no reasoning may also include circular and tautological statements [“*I do not agree with you, because you’re wrong*”]. Partial reasoning structures become visible in *phenomenological reasoning*. In phenomenological reasoning, students present some sets of data, but they have *no* idea (reasoning) about why proposed data support their claim(s) [“*That is because of pressure. It happens when you hold one side of it. If you take your hand from that side of the pipe, it will disappear too (meaning the slope).*”]. In other words, Furtak et al. (2010) explained *data-based* or phenomenological reasoning structures that “rely on data or evidence only. Those structures that reference only data or specific phenomena (*phenomenological reasoning*) as backing for a claim rely on single observations by students (e.g., the rock sinks because I saw it sink) or single properties (e.g., the rock sinks because it’s heavy).” (p. 185).

Relational reasoning is rather complex compare to former categories. In relational reasoning, students’ claims are backed up by evidence(s). The students are now able to apply evidence(s) to claim(s). They are also able to reason about why evidence(s) support or are against their hypothetical claim(s) [“*Because you are stretching your arm and the distance is expanding that way. We are implementing more force, it causes more pressure than the other one; as a result, your arm aches.*”]. In the top level, *rule-based reasoning* requires students to make inductive and/or deductive reasoning by collecting a substantial amount of data and transforming them into evidences allowing generalizations that may be supported by scientific principles. Hereby, students are able to apply a scientific principle or a rule that was previously found in their data to other fields [“*I think in this way you are losing power but gaining from the distance. Because (bending the arm) we use less force, but as we see we move our arm less. However, when we stretch our arm we use more force and move our arm more. This arm always works like this. I mean, it is the same as levers. When it is close to that thing (meaning the fulcrum), you can’t lift up your friend.*”].

Argument Structures of the Student-led Cognitive Contributions

During inquiry-based implementations, students are involved in investigative activities in which they generate their own arguments through data collection, analysis and interpretation processes. In these processes, students may criticize and judge others’ claims, questions and evidences. They may respond their peers through their own claims, warrants, backings, rebuttals and attain this in specific subjects as qualifiers.

Thus, an alternative analysis of cognitive contributions can be conducted by detecting argument structures of students. This analysis can be achieved through Argument Structure catalogue (Table 4). For this purpose, a relevant analytical framework can be Toulmin’s (1958) argument model (Toulmin Argument Pattern; *TAP;* Erduran et al., 2004; Simon et al., 2006). TAP has been used as a basis in characterizing argument structure, for instance, in the sense of classroom talk. It also has been treated as a systematic scheme to derive research-based assessment tools for analyzing argument structures of students that can be qualified as a cognitive contribution to contents taken in inquiry classes (Jimenez-Alexiandre & Erduran, 2008; Erduran et al. 2004; Simon et al. 2006; Soysal, 2012).

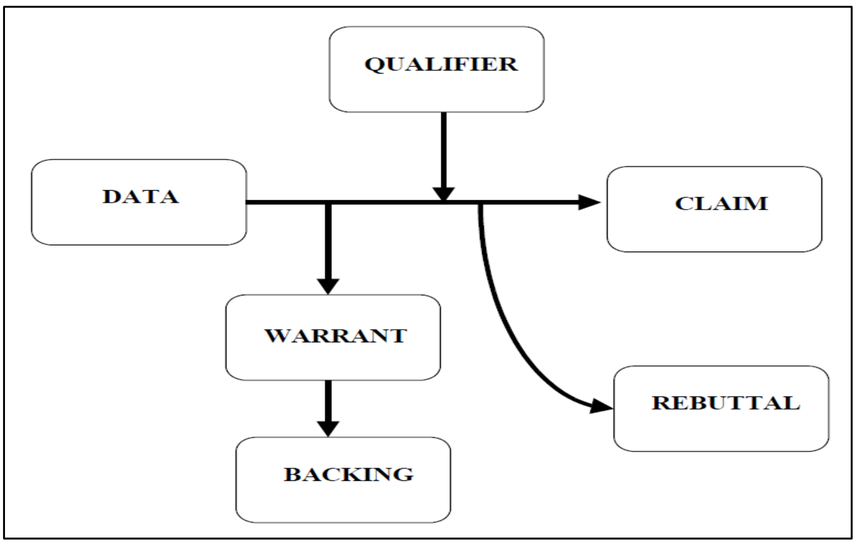
TAP originally incorporates six components of a completed argument (Toulmin, 1958) displayed in Figure 2. These are claim, data, warrant, qualifier, backing and rebuttal. A claim is an assertion about what exists, or the values people hold. Based on the TAP, claims remark a statement about a specific outcome. It can be a prediction of a student regarding what will be happening in the future (e.g. "*In an inclined plane, heavier masses will reach the bottom first*"). A claim can also be an observation of what was happened in the past (e.g., "*In our inclined plane, the heavier one reached the bottom first*"). A claim can also be a conclusion of what happens in the present (e.g., "*The heavier always reach the first in an inclined plane*").

Within TAP, data is described as statements that are used as evidence to support claims. Data component normally incorporates students' recorded observations/measurements. In inquiry-based implementations, data may come from other sources, for instance, everyday experiences of students. Data can be conceived as it "would be generally observed in a science classroom, the premise often identifies an object and a relevant feature or property" (Brown et al. 2010; p. 132).

Toulmin (1958) identified warrants as statements that explicate relationships and coordination between data and claim. Warrants may be matched with evidences. A warrant attaches to claim that is already backed up by reference to a contextualized relationship between two properties, a property and a consequence of that property, or a specific finding (Brown et al., 2010a; 2010b; Furtak et al., 2010; Hardy et al., 2010; Shemwell & Furtak, 2010).

A qualifier purports conditions in which claim can be more feasible (Toulmin, 1958). Qualifiers are the statements illuminating the proper characteristics of the condition in which the claim is posed. A qualifier "is the given information from whence the claim is derived upon. Includes: object, state of an object, general expression ("subject of reasoning"), point of reference" (Furtak et al., 2010, p. 184).

Backings are the underlying assumptions for data-claim coordination that may be presented explicitly or implicitly within an argument (Toulmin, 1958). A backing can be taken as a secondary warrant reinforcing main justification of the claim (Erduran et al. 2004; Simon et al. 2006). Rebuttals are expounded as statements consisting contradictions to data, warrant, backing or qualifier of an argument. In inquiry-based implementations, students are given opportunities to generate counter evidences against to others' arguments. Students are also engaged in the judgement processes of others' data, warrant, backing or qualifier.



*Modified from Toulmin (1958); *The Uses of Argument*

Figure 2. Toulmin Argument Pattern*

Based on the TAP, a hierarchical argument structure assessment criterion was composed (Table 4). By considering the complexity of an argument, four levels of the argument structures were established in the context of this study.

Table 4. Argument Structures of the Student-based Responses

Levels of argument structure	***Descriptions***	***Diagrammatic representations***
LEVEL-1	Only consisting a claim component (C)	C-
LEVEL-2	Consisting of backing (B) or data (D) in addition to a given claim	C-D C-B
LEVEL-3	Consisting of at the least a warrant/justification (W) for a given claim	C-W C-D-W C-D-W-B
LEVEL-4	Consisting of rebuttal(s) (R) against to others claim, warrant, and backing.	C-D:R C-W:R C-D-W:R C-D-W-B:R

Level-1 argument structure incorporates only claim component that is not justified or warranted. Level-1 argument structure is the simplistic one among other levels. In this level, students may propose their subjective or unsupported ideas [“*A sledgehammer makes things more difficult; but we use it.*”].

Level-2 argument structure may include data or backing in addition to claim, but, it is still lacking warrants or claim-based justifications. In this level of argument, students may be able to provide some observational data for their claims; however, they still do not coordinate their claims with data. Students are not able to explicate why data supports their claims in specific contexts [“*For example a sledgehammer. It is a heavy thing, and we have to use more power lift and hit with it.*”].

Level-3 argument structures include at the least a warrant or justification. In this level of argument, students can associate claims with data through displaying a cognitive work. [“*In my opinion, yes. With the bat (racket) we can hit the ball to the further. I can’t do that with bare hands. And this proves that it is a simple machine*.”]

Level-4 argument structure incorporates rebuttals to others’ claim, data, backing, justification or warrant as the most complex argument structure [“*We can imagine something like this. We can’t cut a bone with a small knife; we use something bigger to cut into pieces (the bone-in meat). But for the second situation, we need more power. It may make things difficult at the beginning, however it enables us to complete our task.*”].

In this manner, a rebuttal may be expressed in utterances that can be fragmented into other components of an argument ([Claim+Data+Warrrant]: Rebuttal; CDW: R). In the context of this study, it is acknowledged as a theory-laden sense (Erduran, Simon & Osborne, 2004; Toulmin, 1958) that rebuttals must also be clarified as higher-level arguments including other additional components of an argument instead of just saying that “I do not agree with you”. A similar analytical component analysis was exhibited by Simon, Erduran and Osborne (2006).

CONCLUSIONS AND RECOMMENDATIONS

Six different research groups’ ongoing research endeavours (*The SOLO taxonomy*: [(i)John Biggs and his colleagues (e.g., Biggs & Collis, 1982)]; *Cognitive Pathways*: [(ii)Robert Ennis, (iii)Peter A. Facione; (iv)Bruna I. Grimberg and Brian Hand (e.g., Ennis, 2011; Facione, 1990; Grimberg & Hand, 2009)]; *Reasoning Typologies*: [(v)Erin Marie Furtak and her colleagues (e.g., Furtak, Hardy, Beinbrech, Shavelson & Shemwell, 2010)]; *Argument Structures*: [(vi)Shirley Simon and her colleagues (e.g., Simon, Erduran & Osborne, 2006)]) were reviewed in the previous sections to represent a representation of the assessment tools for cognitive outcomes at the end of the inquiry processes and cognitive contributions to discursive interactions over the course of the inquiry-based processes. Based on the collective efforts of the scholars in this field, two suggestions are proposed for further research aims of science teacher educators.

First, the proposed assessment tools can be effectively applied for in-class interactions' and exchanges' verbatim transcripts' analysis. Concrete examples of the in-class uses of the proposed assessment tools were provided as the above-located excerpts. As mentioned earlier, this side of the analysis signals the in-the-classroom analysis procedures incorporating the principles of the sociolinguistic paradigm. Furthermore, for a more holistic and combined manner, the proposed assessment tools can be applied to student-led assignments, written reflections, selected inquiry-based works (e.g. questions, experimental designs from the students' product files), graphical representations or other multi-modal clarifications, longer essays on the inquiry-based activities or for the assessment of the open-ended, conceptual exam questions. To our knowledge, these can be considered as the end-products of the inquiry processes within the principles of the process-product paradigm.

Secondly, for establishing more concrete validity and reliability mechanisms for the assessment procedures, the offered assessments tools can be applied in a combined manner. To support, most of the tools incorporate perception, conception and abstraction stages of cognitive outcomes. The perception stage attaches to "data collection procedures", the conception stage associates with "data analysis-interpretation procedures", and abstraction stage relates to the "generalisation" displaying a wholistic representation of an inquiry-based implementation. In a common sense, most of the offered tools (e.g., the SOLO taxonomy, Cognitive Pathways, Reasoning Quality) requires generalised arguments (Argument Structures) that can be beyond the phenomenon and its related aspects that are gained during the inquiry processes. As a whole, end-scores on an assessment tool (the SOLO Taxonomy) can predict, confirm and validate the alternative end-scores that are created by other offered tools (Cognitive Pathways, Reasoning Quality).

REFERENCES

Abrami, P. C., Cohen, P. A., & d'Apollonia, S. (1988). Implementation problems in meta-analysis. *Review of Educational Research, 58*, 151-179.

*Anderson, L., Krathwohl, R., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M. (Eds.) (2001). *Taxonomy for learning, teaching and assessing: A revision of bloom's taxonomy*. New York, NY: Longman.

*Akkus, R., Gunel, M., & Hand, B. (2007). Comparing an inquiry-based approach known as the Science Writing Heuristic to traditional science teaching practices: Are there differences? *International Journal of Science Education, 29*, 1745-1765.

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2013). Australian curriculum: Science. Retrieved from https://www.australiancurriculum.edu.au/

Benedict-Chambers, A., Kademian, S. M., Davis, E. A., & Palincsar, A. S. (2017). Guiding students towards sensemaking: teacher questions focused on integrating scientific practices with science content. *International Journal of Science Education, 39*(15), 1977-2001.

**Biggs, j. B. & Collis, K. F. (1982). *Evaluating the Quality of Learning: the SOLO taxonomy* (New York, Academic Press).

**Biggs, J. B., (1992). *Modes of learning, forms of knowing, and ways of schooling*. In Demetriou, A., Shayer, M., and Efklides, A. (Eds.), Neo-Piagetian Theories of Cognitive Development. London: Routledge. (pp. 31-51).

**Biggs, J.B. (1993). What do inventories of students` learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology, 63*, 3-19.

*Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (Eds.). (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook 1: cognitive domain. New York: David McKay Co. Inc.

Brophy, J., & Good, T. L. (1986). *Teacher behavior and student achievement*. In M. C. Wittrock (Ed.), Handbook of research on teaching (pp. 328-375). New York: Macmillan.

**Brown, N. J. S., Furtak, E. M., Timms, M., Nagashima, S. O., & Wilson, M. (2010a). The Evidence-Based Reasoning Framework: Assessing Scientific Reasoning. *Educational Assessment, 15*, 123-141.

**Brown, N. J. S., Nagashima, S. O., Fu, A., Timms, M. & Wilson, M. (2010b). A Framework for Analyzing Scientific Reasoning in Assessments. *Educational Assessment, 15*, 142-174.

Carlsen, W.S. (1991). Questioning in classrooms: A sociolinguistic perspective. *Educational Research, 61*, 157-178.

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K-12 science contexts. *Review of Educational Research, 80*(3), 336-371.

*Cavagnetto, A., Hand, B. M., & Norton-Meier, L. (2010). The Nature of Elementary Student Science Discourse in the Context of the Science Writing Heuristic Approach. *International Journal of Science Education, 32*(4), 427-449.

Cavagnetto, A., & Hand, B. M., (2012). *The Importance of Embedding Argument Within Science Classrooms.* In M.S. Khine (ed.), Perspectives on Scientific Argumentation, Springer Science+Business Media B.V. 2012 (pp. 39-53).

**Chan, C. C., Hong, J. H. & Chan, M. Y. C. (2001). *Applying the Structure of the Observed Learning Outcomes (SOLO) taxonomy on student's learning outcomes: a comparative review*. (Unpublished manuscript, Hong Kong, Hong Kong Polytechnic University) .

**Chick, H. (1998). Cognition in the formal modes: research mathematics and the SOLO taxonomy. *Mathematics Education Research Journal, 10*(2), 4-26.

*Chin, C. (2006). Classroom interaction in science: Teacher questioning and feedback to students' responses. *International Journal of Science Education, 28*, 1315-1346.

**Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching, 44*(6), 815-843.

**Collis, K.F. and Davey, H.A. (1986). A technique for evaluating skills in high school science. *Journal of Research in Science Teaching*, *23*, 651-663.

*Crawford, B.A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching, 37*, 916-937.

*Driscoll, A. and Wood, S. (2007). *Developing Outcomes-based Assessment for Learner-centred Education: A Faculty Introduction*. Sterling, Virginia: Stylus.

**Ennis, Robert H. (2002). *Goals for a critical thinking curriculum and its assessment*. In Arthur L. Costa (Ed.), Developing minds (3rd Edition). Alexandria, VA: ASCD. (pp. 44-46.)

**Ennis, Robert H. (2004). Applying soundness standards to qualified reasoning. *Informal Logic, 24*(1), 23-39.

**Ennis, R. (2011). Critical Thinking: Reflection and Perspective, Part I. *Inquiry: Critical Thinking Across the Disciplines, 26*(1), 4-18.

**Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science education*, 88(6), 915-933.

**Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*. "The Delphi Report," Committee on Pre-College Philosophy. (ERIC Doc. No. ED 315 423).

**Furtak, E. M., Hardy, I., Beinbrech, C., Shavelson, R. J. & Shemwell, J. T. (2010). A Framework for Analyzing Evidence-Based Reasoning in Science Classroom Discourse, *Educational Assessment, (15)*, 3-4, 175-196.

Gliner, J. A., Morgan, G. A., & Harmon, R. J. (2003). Metaanalysis: Formulation and interpretation. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 1376–1379.

**Grimberg, B. I. & Hand, B. (2009) Cognitive Pathways: Analysis of students' written texts for science understanding. *International Journal of Science Education, 31*(4), 503-521.

*Gunel, M. (2006). *Investigating the impact of teachers' implementation practices on academic achievement in science during a long-term professional development program on the Science Writing Heuristic*. Unpublished PhD thesis, Iowa State University, Iowa.

**Hardy, I. Kloetzer, B., Moeller, K., & Sodian, B. (2010). The Analysis of Classroom Discourse: Elementary School Science Curricula Advancing Reasoning with Evidence. *Educational Assessment, 15*, 197-221.

Jiménez-Aleixandre, M. P., & Erduran, S. (2008). *Argumentation in science education: An overview*. In S. Erduran, & M. P. Jiménez-Aleixandre (Eds.), Argumentation in science education: Perspectives from classroom-based research (pp. 3-27). Dordrecht: Springer.

*Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview, *Theory into Practice, 41*(4), 212-218.

**Lake, D. (1999). Helping students to go SOLO: teaching critical numeracy in the biological sciences. *Journal of Biological Education, 33*(4), 191-198.

Lefstein, A. (2008). Changing classroom practice through the English national literacy strategy: A micro-interactional perspective. *American Educational Research Journal, 45*, 701-737.

Lin, T.C., Lin, T.J. & Tsai, C.C. (2014). Research trends in science education from 2008 to 2012: A systematic content analysis of publications in selected journals. *International Journal of Science Education*, *36*(8), 1346-1372.

*Martin, A. M., & Hand, B. (2009). Factors affecting the implementation of argument in the elementary science classroom. A longitudinal case study. *Research in Science Education, 39*, 17-38.

*McNeill, K., Lizotte, D., Krajcik, J., & Marx, R. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences, 15*(2), 153-191.

*McNeill, K. L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching, 45*(1), 53-78.

*McNeill, K. L. (2009). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education, 93*(2; 2), 233-268.

*McNeill, K. L., & Pimentel, D. S. (2010). Scientific Discourse in Three Urban Classrooms: The Role of the Teacher in Engaging High School Students in Argumentation. Science Education, 94, 203-229.

Mercer, N. (2004). Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistic, 1*(2), 137-168.

Mercer, N. (2010). The analysis of classroom talk: Methods and methodologies. *British Journal of Educational Psychology, 80*, 1-14.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: The National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC. Retrieved from http://www.nap.edu/catalog/13165/a-framework-for-k-12-science-education-practices-crosscutting-concept

NGSS Lead States. (2013). Next generation science standards: For states, by states. Retrieved from:http://www.nap.edu/catalog/18290/next-generation-science-standards-for-states-by-states

**Panizzon, D. (2002). Using a cognitive structural model to provide new insights into students` understanding of diffusion. *International Journal of Science Education, 25*(12), 1427-1450.

Pimentel, D. S. & McNeill, K. L. (2013). Conducting talk in science classrooms: Investigating instructional moves and teachers' beliefs. *Science Education, 97*(3), 367-394.

*Potter & Kustra (2012). *Course Design for Constructive Alignment*. Centre for Teaching and Learning, University of Windsor (accessed from uwindsor.ca/ctl/system/files/PRIMER-on-Learning-Outcomes.pdf on 15.01.2017)

**Shemwell, J. T., & Furtak, E. R. (2010). Science Classroom Discussion as Scientific Argumentation: A Study of Conceptually Rich (and Poor) Student Talk. *Educational Assessment, 15*, 222-250.

*Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education, 27*, 137-162.

Soysal, Y. (2012). Sosyobilimsel *Argümantasyon Kalitesine Alan Bilgisi Düzeyinin Etkisi: Genetiği Değiştirilmiş Organizmalar*. Abant İzzet Baysal Üniversitesi Eğitim Bilimleri Enstitüsü, Yüksek Lisans Tezi, Bolu.

Soysal, Y., & Radmard, S. (2017). Sosyal Oluşturmacı Öğretimin Öğretmen Adaylarının Öğrenme ve Öğretmeye Yönelik İnançlarına ve Sınıf içi Uygulamalarına Etkisinin İncelenmesi. *İlköğretim Online, 16*(4), 1505-1531, DOI: 10.17051/ilkonline.2017.342972.

Soysal, Y., & Radmard, S. (2018). Social negotiations of meanings and changes in the beliefs of prospective teachers: A vygostkian perspective, *Educational Studies, 44*(1), 57-80, DOI: 10.1080/03055698.2017.1345676.

*Sugrue, B. (2002). Problems with Bloom's taxonomy. Retrieved on January 20, 2017, from: https://eppicinc.files.wordpress.com/2011/08/sugrue_bloom_critique_perfxprs.pdf

Suri, H. & Clarke, D. (2009). Advancements in research systhesis methods: From a methodologically inclusive perspective. *Review of Educational Research*, *79*(1), 395-430.

**Toulmin, S. (1958). *The uses of argument* (Updated edition ed.). Cambridge: Cambridge University Press.

van D. Booven (2015). Revisiting the authoritative–dialogic tension in inquiry-based elementary science teacher questioning. *International Journal of Science Education, 37*(8), 1182-1201.

APPENDIXES

APPENDIX-1: COGNITIVE PATHWAYS		
COGNITIVE CATEGORIES	**DEFINITIONS**	**LEVELS**
Observation	Data that result from students' observations	Perception
Measurement	Reference to any quantitative aspect of the data	Perception
Compare	Reference to common/different characteristics of two or more pieces of data or objects	Perception
Analogy	Mapping elements from a source domain (well-understood situation) into a target domain (non-familiar situation)	Conception
Clarifications	Questions or knowledge that stimulate clarification supporting other operations	Conception
Claim	Unproved inference or explanation	Conception
Cause/effect	Identification of a cause and its effect	Conception
Induction/generalization	Reasoning that links few examples to general premises	Abstraction
Deduction	Reasoning that links general premises to a specific	Abstraction
Investigation design	Planning new experiments	Abstraction
Argumentation	Negotiation of meaning with others	Abstraction

APPENDIX-2: CRITICAL THINKING SKILLS (by Robert Ennis)	
SKILLS	***CHARACTERISTICS***
Basic clarification	Focus on a question
	Analyse arguments
	Ask and answer clarification and/or challenge questions
Decision-making	Judge the credibility of a source
	Observe, and judge observation reports.
Inference	Deduce, and judge deduction
	Make material inferences (roughly "induction")
	Make and judge value judgments
Advanced clarification	Define terms and judge definitions, using appropriate criteria
	Attribute unstated assumptions (an ability that belongs under both basic clarification and inference)
Supposition and integration	Consider and reason from premises, reasons, assumptions, positions, and other propositions with which they disagree or about which they are in doubt, without letting the disagreement or doubt interfere with their thinking
	Integrate the dispositions and other abilities in making and defending a decision

APPENDIX-3: CRITICAL THINKING SKILLS (by Peter A. Facione)	
SKILLS	***SUB-SKILLS***
Interpretation	Categorisation, decoding significance, clarifying meaning
Analysis	Examining ideas, identifying arguments, analysing arguments
Evaluation	Assessing claims, assessing arguments
Inference	Querying evidence, conjecturing alternatives, drawing conclusions
Explanation	Stating results, justifying procedures, presenting arguments
Self-regulation	Self-examination, self-correction