PEL Pensamiento Educativo. Revista de Investigación Educacional Latinoamericana
2017, 54(2), 1-18

Pontificia Universidad Católica de Chile
Facultad de Educación

A Theoretical and Methodological Approach to Evaluating Metacognition in School Contexts

Aproximación teórico-metodológica a la evaluación de la metacognición en contextos escolares.

Consuelo Iturrieta, Loreto Jara & Gisela Watson

Fundación Educación 2020

Abstract

The organization called Educación 2020 has engaged in the challenge of implementing and validating the program of Tutoring Networks in Chile, which is a method for teaching and learning that originated in Mexico in the 90s with the aim of generating change in the national policy in public education. The main challenge of this method is to develop skills for the 21stcentury, specifically related to metacognition. The review of scientific literature shows that metacognitive skills allow students to have better learning results due to increased understanding of their learning processes and the use of strategies that enable them to reach their goals. Regarding the objective to evaluate the implementation of the program, Educación 2020 seeks to measure the learning results of the students and the possible changes in the development of metacognitive skills. The proposal is an instrument of low costs that would be easy to implement before and after the start of the program. The application will be done by the teacher in a classroom with students between the ages of 11 and 16 that are participating in the program of Tutoring Networks. As soon as the instrument is validated, it will be publically available for use in the educational system.

Keywords: metacognition, metacognitive knowledge, metacognitive skills, instrument for evaluation

Post to:

Loreto Jara Males
ljara@educacion2020.cl
Área de Política e Innovación, Fundación Educación 2020
República #580, Santiago, Chile.

ISSN:0719-0409 DDI:203.262, Santiago, Chile
doi: 10.7764/PEL.55.1.2018.2

Resumen

Fundación Educación 2020 se ha propuesto el desafío de implementar y validar el programa Redes de Tutoría en Chile, que consiste en un método de enseñanza y aprendizaje que nace en México en la década de los 90 y busca generar cambios en su política pública educacional. El desafío principal de este método es el desarrollo de habilidades para el siglo XXI, particularmente la metacognición. La revisión de evidencias científicas respalda que la metacognición permitiría al estudiante la obtención de mejores logros, gracias al mayor conocimiento de su proceso de aprendizaje y de utilización de estrategias que le permitan alcanzar las metas propuestas. Con el objetivo de evaluar la implementación del programa, Educación 2020 busca medir los resultados en los estudiantes y los posibles cambios en el desarrollo de dicha habilidad. Se propone un instrumento de evaluación de fácil aplicación y bajo costo, con medición pre y post intervención. La aplicación la realiza el docente en aula y está dirigida a estudiantes entre 11 y 16 años, que participan en el programa de Redes de Tutoría. En la medida que se valide el instrumento, se espera que sea de utilidad pública para los actores del sistema escolar.

Palabras clave: metacognición, conocimiento metacognitivo, habilidades metacognitivas, instrumento de evaluación.ser vivo

Introduction

The Chilean school system has defined that the aim of education is to cultivate values, provide knowledge, and enable children and young people to generate the skills necessary to fully develop in the spiritual, intellectual, affective, physical, and artistic domains (Ley General de Educación No.20370, 2009). Attaining such a highly complex aim requires major efforts from education centers throughout the country. Although progress has been made in the implementation of a set of structural changes in the school system, standardized tests show that improvements in learning outcomes have been slow, with major problems in the distribution of educational achievement associated with student SES (Agencia de la Calidad de la Educación, 2016).

The results of the 2016 SIMCE test show no significant progress in the Reading scores of 4th grade students, although a slight upward trend can be observed in Mathematics. In 10th grade, results are more alarming: Reading scores have dropped by 7 points over the last decade; also, while Mathematics scores have increased by 10 points during the same period, they have not significantly improved during the last 5 years (Agencia de la Calidad de la Educación, 2016).

These results become more serious when considering the large SES-related gaps present, even though public spending represents 60.1% of the total education budget in Chile (Mineduc, 2015): in 6th grade, the difference between high- and low-SES students can reach 72 points in Mathematics and 51 in Reading Comprehension (Agencia de la Calidad de la Educación, 2016).

In line with these observations, the Education Quality Agency [Agencia de la Calidad de la Educación] has stated that the 2012 PISA tests show that 38% of Chilean students fail to meet the minimum competence level required for problem-solving (Agencia de la Calidad de la Educación, 2014). The Agency's report also highlights the level of inequity observed in Chile: "the higher the socioeconomic and cultural level of a Chilean student, the better his or her problem-solving competences will be" (Agencia de la Calidad de la Educación, 2014, p.6).

The results of the 2015 PISA test reveal a similarly unfavorable situation for Chile: annual average scores did not increase from 2012 to 2015, neither in Science nor in Mathematics; however, Reading scores did increase slightly (3 points per year) (Banco Interamericano del Desarrollo [BID], 2016a). Approximately 35%

of students perform poorly in Science, 49% in Mathematics, and 28% in Reading; on average, 48% of Chilean 15-year-olds lack the basic competences measured by PISA tests (BID, 2016b).

Students' opinions are in line with these findings and reveal a wish to change the traditional teaching methods in order to improve their learning processes. In "Tell Them What You Want to Learn", a survey conducted by UNESCO in 2016, over 60 thousand students stated that they wanted to learn differently (UNESCO-OREALC, 2017). The survey showed that…

> (…) they want to learn at school, with lessons and teachers; they want to learn in other spaces beyond the traditional ones; they want to learn in other ways. There is a demand for a broader notion of learning, which requires not only a good school, but also other educational spaces and other learning methods. Students are asking for education to extend beyond basic skills, incorporating integrated, complex, and dynamic knowledge for life-long learning… (UNESCO-OREALC, March 14, 2017).

Also, consistent with UNESCO data, in a study conducted by the Latin American Network of Civil Society Organizations for Education [Red Latinoamericana de Organizaciones de la Sociedad Civil por la Educación, REDUCA] (2017) that involved workshops for children in Brazil, Chile, Colombia, and Mexico, students expressed their rejection of teaching methods based on listening and re-writing, since they wished to participate more actively in their learning process. Students stated that, for them, the ideal teacher "must teach in an active way –that is, his/her methods must be dynamic and fluid– which should result in the use of a variety of methodologies during the teaching-learning process, with students learning in an enjoyable, context-sensitive way" (REDUCA, 2017, p.38).

The same study concludes that teachers should be equipped with the theoretical and methodological knowledge necessary to approach the topics covered in their subjects from a variety of perspectives and using multiple tasks. This should enable them to constantly reformulate their methodologies and develop novel strategies that can replace traditional pedagogical practices (REDUCA, 2017).

Considering this information, Educación 2020, as an institution that seeks to further quality and equity in education through the advancement of public policies and the encouragement of direct citizen action, has implemented initiatives aimed at improving school education based on the following core idea: recent and ongoing educational reforms have targeted the conditions of the education system (infrastructure, resources, regulations) but have failed to have a significant impact on the learning outcomes of children and young people. Therefore, a paradigm shift is needed in the Chilean school system in order to establish the learning-teaching process as the core and fundamental purpose of the transformation, considering that:

> … educational systems must be transformed into systems for lifelong learning, with an approach centered on students and their learning and mediated by teachers and other participants, including classmates, all of which entails an expansion of the role of teachers. (UNICEF, 2015, p.42).

In this context, institutions have taken up the challenge of studying, examining, and implementing educational experiences based on innovative models that inspire the transformation of the Chilean educational system. The models selected should have an impact on the core of teaching and result in the development of learning skills suited to the 21st century, a change that involves rethinking what is taught, how it is taught, and how learning is assessed (Rincón-Gallardo, 2014). In order to generate meaningful learning as a result of these transformations, the vast literature on educational innovation shows that the changes implemented should encourage self-regulated learning, be context-sensitive, involve collaborative models, and highlight the emotions and motivations of those who participate in the process (Blanco & Messina, 2000; OECD, 2013).

Tutoring Networks are one of the innovative models promoted by Educación 2020. This educational approach, which originated in Mexico, is aimed at developing autonomous learning in order to awaken people's desire to learn and their motivation to teach (Consejo Nacional de Fomento Educativo, 2016; López, 2016). The educational axiom of this approach is that "good learning happens when the learner's interests match the teacher's skills" (Secretaría de Educación Pública, 2012, p.18; Rincón Gallardo, 2011). Its implementation in learning communities comprises the following steps: first, the preparation of a topic (learning challenges related to topics of interest); second, the tutoring relationship, in which topics are covered via a personalized dialog characterized by the use of existing resources and knowledge, the orientation of learning through guiding questions, and the identification of 'how people learn what they learn'. The third step takes place at the end of the tutoring relationship and involves a "public demonstration", a display in which the tutee explains to his/her peers what he/she has learned. The process ends when, on another occasion, the tutee tutors a classmate. Thus, the classroom gradually becomes a learning community in which teachers and students have the chance to learn and teach. A network is generated when the topics covered are reproduced and when other actors from the same school and other schools of the region join the process (López, 2016).

Thus, as described by Farrell, Manion, and Rincón Gallardo (2017), those who play the role of tutees choose their study topics from a catalog offered by those who play the role of tutors. Both the tutor and tutee roles are mainly played by students, but teachers, paraprofessional educational personnel, and/or members of the educational community can also participate. Tutees, supported by their tutors, follow their own path and rhythm to study each topic: "Tutoring Networks stress the development of independent learning skills; therefore, covering the whole curriculum is not as important as ensuring that students will improve their ability to learn on their own in each new research project" (Farrell, Manion, & Rincón Gallardo, 2017, p.25).

Teachers, focusing on individual students, dyads, or small groups, transform their role in the classroom without interrupting their teaching efforts. According to the same study, teachers quickly learn how to adopt a radically different teaching approach, which results in excellent outcomes for students. Teachers receive intensive training on how to apply the methodology, which generally involves short sessions held in their schools with an advisor (Farrell, Manion, & Rincón Gallardo, 2017).

The experience of implementing Tutoring Networks in Mexico (which has been studied by Richard Elmore, Santiago Rincón-Gallardo, and other leading researchers in multiple contexts) demonstrates that the learning outcomes of children and young people improve significantly, affecting various aspects of students' learning and of their educational communities: improvements are observed in school climate and internal school efficiency indicators, in socioemotional skills (motivation, security, self-esteem, respect, patient, tolerance), and in the development of complex skills linked to metacognition (Secretaría de Educación Pública, 2012; López, 2016).

Given that the latter aspect —metacognition— can be essential for facilitating better learning outcomes, it has been incorporated as an element to be covered in the implementation of Tutoring Networks in Chile. As of December 2017, Educación 2020 has implemented this model in 34 schools located in three regions of Chile, with projects currently being at various stages of completion. A baseline assessment was administered to each student taking part in the Tutoring Networks project, working on the assumption that these skills will improve as a result of their training as tutors.

It should be noted that, despite the limited period during which these networks have been active in Chile, significant improvements can be observed in small student samples (in specific Language and Mathematics tests), which are still being evaluated and awaiting publication. Some studies have linked metacognition development to better competences in the areas of reading comprehension and problem-solving (Aragón & Caicedo, 2009; Iriarte, 2011; Ruiz, 2002).

This article presents the theoretical and methodological construct that underlies the metacognitive test administered as part of the implementation of the Tutoring Networks in Chile during 2015, 2016, and 2017. It comprises three sections: theoretical-conceptual framework, methodological framework, and discussion and conclusions. In the latter section, the authors outline future challenges regarding the instrument and its potential applications.

Although no analysis is presented of the results yielded thus far by the test, this article and the process that it describes are aimed at enriching reflections about the generation of learning-teaching and evaluation processes in which trust, the ability to identify one's resources, and horizontal dialog with others facilitate the provision of quality education in a context of equity.

Background and conceptual framework

In order to make the instrument more easily comprehensible, background information is presented in three sections: evaluation in school contexts, the metacognition construct, and the evaluation of said construct.

Evaluation in school contexts. The importance of evaluation in school contexts has been discussed by several authors from a variety of perspectives (e.g. Arratia, Flotts, & Rodríguez, 2012; Monereo, 2003; Ravela, 2013; Santos, 1993; Villalón, 2002), with definitions essentially agreeing that it is a reflective activity based on the assessment of information about student learning (Zepeda, 2008). Teachers need valid and reliable information about student learning in order to make suitable teaching decisions that orient their academic progress (Arratia, Flotts, & Rodríguez, 2012). As Zepeda (2008, p.246) points out, "evaluation and learning interact in the teaching progress: evaluation provides students with chances to learn more and better, while learning incorporates evaluation strategies; thus, the process allows students and teachers to develop their metacognition, reflection, and self-regulation skills".

Although evaluation in the classroom is a practice that takes place in all school subjects and often has a summative aim, in order to implement the new ways of learning that students are requesting, it is essential to expand the scope of evaluation, which requires transforming the evaluation culture of schools.

According to Arratia, Flotts, and Rodríguez (2012), a school's evaluation culture can be placed within a continuum between two poles: "near one pole are those schools that emphasize the normative and/or certificatory role of evaluation over its teaching role (…); near the opposite pole are those where the teaching role of evaluation predominates" (p.102). Regarding said teaching function, these authors mention a learning-oriented type of evaluation that focuses on the progress made by each student and is based on the assumption that all students can improve.

Evaluation must fulfill that pedagogical role so that value judgments about the information being measured can be enriched by the joint dialog, discussion, and reflections of those involved in the task being evaluated (Santos, 1993, p.2).

One approach that makes it possible to fulfill the pedagogical role of evaluation is the measurement of the learning strategies employed during the knowledge construction process, which for Celman (1998) "is an area full of educational potential that can have a huge impact on the transformation of said process" (p.12). Specifically, researchers nowadays stress the importance of developing metacognitive skills or "learning to learn", a set of knowledge that society requires in order to educate individuals who are more autonomous in their use of cognitive tools (Celman, 1998).

The above is corroborated by Darling-Hammond and McCloskey's (Programa de Promoción de la Reforma Educativa en América Latina y el Caribe [PREAL], 2009) comparison of the evaluation tools that have been shown to be successful across the globe:

> The European and Asian countries that have considerably improved student learning have explicitly privileged syllabuses and evaluation plans focused on skills for: finding and organizing information to solve problems, conducting research, analyzing and summarizing data, applying the contents learned to new situations, self-monitoring and improving one's performance, engaging in multiple forms of communication, working as part of a team, and learning autonomously (PREAL, 2009, p.1).

The metacognition construct. The variables with the highest impact on learning outcomes, according to a meta-analysis conducted by Wang, Haertel, and Walberg (1990), are those that link students' commitment or engagement to the material to be learned. Specifically, metacognition was found to be the best predictor in this study. As defined by the authors, this variable was considered to be present in all studies on planning, overall monitoring (of the effectiveness of the actions performed during the process and the results obtained), self-regulation use, self-control strategies, and strategies for facilitating the generalization of concepts (Wang et al., 1990).

Nowadays, several definitions of the construct exist, although most tend to be based on that published by Flavell some decades ago: knowledge and regulation of one's cognitive activities during the learning process (Flavell, 1979, in Hennessey, 1999; Veenman, Van Hout-Wolters, & Afflerbach, 2006; Lai, 2011; Jaramillo & Osses, 2012).

In the same vein, Kuhn and Dean (2004) define it as a *construct* similar to awareness and management of one's thinking, or "thinking about one's thinking" (p.270), which may be related to the concept of executive control or executive functions in cognitive psychology. For other authors, metacognition also involves the participation of consciousness as a regulatory mechanism (Klimenko & Alvares, 2009).

Veenman et al. (2006) warn of the theoretical risk that this notion of metacognition entails, since it is linked to the "homunculus problem" or Comte's paradox: "one cannot be divided into two, while one part thinks and the other watches itself think" (p.5). So, these authors assert that metacognition is based on cognition and that they are not independent domains; for instance, it is not possible to plan an activity without proper command of cognitive strategies such as problem-solving or generating a sequence of steps.

Considering this objection, it is possible to link the exercise of metacognition with the person who learns, based on the learner profile proposed by Mateos (2001):

> ... a competent learner is one who uses his/her metacognitive knowledge to self-regulate his/her learning effectively; at the same time, the regulation of his/her learning can enable him/her to acquire new knowledge linked to the task and to his/her own resources as a learner (p.20).

This statement suggests a reciprocal relationship between cognition and metacognition, which makes it possible to propose a hypothetical evaluation instrument that took both constructs into account.

The development of metacognition in human beings can be understood from a Vygotskian perspective and considering the social nature of learning. In line with this view, Tutoring networks always require at least one tutor and one tutee. Whitebread's team continued to examine the issue in more detail and pointed out that "learning to be an effective learner, according to this approach, is a process of acculturation and internalization whereby the child stops being regulated by a third party and starts self-regulating" (Whitebread, Bingham,

Grau, Pino-Pasternak, & Sangster, 2007, p.435). This is precisely the aim of the innovative Tutoring Networks methodology: through the tutoring relationship, the tutor should support the tutee's learning in a way that provides him/her with "scaffolding" (Whitebread et al., 2007) while also playing a "mediating" role, defined as a process in which the tutor guides and leads the tutee's learning process "in a reflective and critical manner in order to encourage him/her to construct meaningful learning, induce the conscious understanding of his/her own learning processes, and anticipate the transference of the contents learned to new situations" (Ruiz, 2002, p.6).

In other words, this learning is focused on metacognitive skills and knowledge, so that students manage to be aware of their learning process, acknowledge the magnitude of the challenge that it poses, determine what resources they have available, plan the activity, test multiple strategies and observe their effectiveness, check the results obtained, and identify what they have learned.

According to Veenman et al. (2006), this distinction between metacognitive knowledge and skills is commonly found in the literature. Similarly, another operationalization labels the two components "knowledge about cognition" and "regulation of metacognition" (Schraw, Crippen, & Hartley, 2006).

Metacognitive knowledge refers to an individual's knowledge of the personal, task, and strategy variables that affect his/her cognitive performance (Schraw et al., 2006; Flavell, 1979, in Veenman et al. 2006; Flavell, 1987, in Whitebread et al., 2007). Personal variables refer to the knowledge acquired by the individual and his/her beliefs about his/her nature as a cognizant organism; task variables refer to the knowledge that the individual acquires about the nature of the task and how to solve it successfully; and strategy knowledge consists in being familiar with the procedure designed in order to achieve a goal (Ruiz, 2002). Klimenko and Alvares (2009, p.18) introduce an example to understand these three types of knowledge: "a student, when tackling a problem, realizes that it belongs to a topic unknown to him (knowledge of a personal characteristic), that the way in which it is presented makes it difficult to understand (knowledge of a characteristic of the task), and that making a graph will help him understand it better (knowledge of a strategy)".

Metacognitive or cognition regulation skills, for their part, refer to a person's procedural knowledge that enables him/her to regulate his/her problem-solving and learning activities (Brown & DeLoache, 2005, in Veenman et al., 2006). Three domains are widely reported in the literature: planning, follow-up, and evaluation (Schraw & Moshman, 1995, in Schraw et al. 2006). Planning involves selecting appropriate strategies and allocating resources; follow-up or monitoring comprises skills for controlling one's own learning; and evaluation refers to the assessment of the processes employed and the products of the knowledge acquired (Schraw et al. 2006).

In brief, as summarized by Jaramillo and Osses (2012, p.119), "it is possible to differentiate two interrelated metacognitive components: one declarative and another procedural, both of which are important for learning". The instrument proposed in this document employs the distinction presented: it is aimed at measuring procedures (planning, follow-up, and evaluation) through the subject's declarative knowledge.

In addition, it has been observed that evidence about the development of metacognition in human beings has evolved over time, mainly with respect to the moment when metacognitive skills "appear" (Lai, 2011) and the impact of this event on classroom work. Regarding the emergence of metacognition, according to Whitebread et al. (2007), early childhood metacognitive skills had been underestimated due to the evaluation methods used. In their research, these authors report the emergence of metacognitive and self-regulatory behaviors between three and five years of age.

Regarding their impact on lessons, it has been observed that interventions have gradually shifted from the direct teaching of metacognitive skills to the creation of social environments that make metacognition possible or easier to apply (Whitebread et al., 2007). For Osses and Jaramillo (2008), the teaching of metacognitive strategies in the classroom can be classified according to two criteria: students' degree of awareness of the strategies used and the degree of autonomy given to them. The first criterion is focused on choosing strategies that induce students to test and confirm their observations; they should plan, supervise, and evaluate the effectiveness of the strategies that they use. The second criterion proposes that teachers act as models and guides of students' cognitive and metacognitive activities, "gradually leading them to achieve increasing competence levels, while at the same time reducing the support provided until students eventually control the process themselves" (Osses & Jaramillo, 2008, p.195).

Upon the basis of the premise that metacognition is fundamental for learning, it should be noted that this concept is not easy to understand. First, more research is needed, and there is a proliferation of associated concepts that complicate this field of study (Veenman et al., 2006); second, barriers exist in school contexts, as teachers are not familiar with methods for teaching and evaluating metacognition (Lai, 2011).

"Self-regulation" is another concept linked to metacognition that has already been used and is hinted at in Mateos (2001). Veenman et al. (2006) explain that some authors regard self-regulation as part of the metacognition construct, while others consider that self-regulation is broader and encompasses metacognition. Considering the sociocognitive perspective analyzed by Veenman et al. (2006), the main difference between these frameworks of reference is that self regulation encompasses motivational and socioemotional processes. Zimmerman and Schunk (2011) add that self-regulation includes not only metacognition and affects (or motivational processes), but also behaviors oriented toward the attainment of learning goals. In this vein, Torrano and González (2004) analyze the model advanced by Pintrich in the year 2000, which divides self regulation into multiple phases and areas, with the latter including motivational, affective, cognitive, behavioral, and contextual aspects.

There seems to be no consensus in this regard, a situation that cannot be resolved in this article. However, the use of the concept must be clarified in order to include it in the methodology. In this proposed classroom evaluation instrument, metacognition should be understood as a concept encompassing self-regulation and —therefore— cognition-related processes: motivation/emotion, behavioral elements, and contextual aspects. This decision was made based on the hypothesis that the metacognition concept generates more openness in teachers (who play a relevant role in the administration and analysis of the instrument) and that self-regulation has been mostly associated with knowledge structures that concern the "self-control of disruptive behaviors" in the classroom.

Metacognition and evaluation. Methodologies focused on developing students' metacognitive skills rather than curricular contents "will help strengthen their strategic knowledge about said contents" (Monereo, 2003, p. 75). Several strategies for developing students' metacognitive capacity have been proposed, including cognitive and metacognitive modeling, metacognitive questioning, metacognitive discussions, and cooperative teaching, among others (Klimenko & Alvares, 2009, p.24).

In order to evaluate students' metacognitive capacity and the methodologies used (understood as the strategies and devices used to foster its development; in this case, the tutoring relationship), the model proposed by Anderson, Krathwohl, Airasian, Cruiskshank, Mayer, Pintrich, et al. (2001), and described by Arratia, Flotts, and Rodríguez (2012), can be examined. This model consists in a bimodal matrix that classifies aims according to what is being evaluated:

	COGNITIVE PROCESSES					
KNOWLEDGE	Remembering	Understanding	Applying	Analyzing	Evaluating	Creating
Factual knowledge						
Conceptual knowledge						
Procedural knowledge						
Metacognitive knowledge						

Figure 1. Bimodal matrix for aim classification (Anderson et al., 2001)

Designing an instrument for assessing metacognitive knowledge through cognitive processes is the challenge that Educación 2020 has undertaken as part of the implementation of the Tutoring Networks.

Still, even today, Monereo's hypothesis (2003) remains valid: few initiatives have been aimed at developing instruments for the assessment of strategic knowledge or metacognition.

According to Lai (2011), commonly used methods include artificial laboratory tasks, self-report instruments, thinking-out-loud approaches aimed at making students' thought process visible, and approaches based on observing the learner.

Said self-report instruments include the Metacognitive Awareness Inventory, created by Schraw and Dennison (1994) to identify metacognitive skills in adults. This inventory has gained widespread acceptance: Turkish- and Spanish-language versions have been adapted and validated, it has been administered in paper and online formats in several countries, and a version for teachers and another for children have been adapted (Huerta, Vesgas, & Galindo, 2014).

Some years later, one of the authors of the MAI and his team reported how difficult it is to measure metacognition due to its non-conscious or non-explicit components. This occurs either because of the automation of certain processes (mainly observed when a subject has mastered a task) or as a result of the difficulties involved in reporting or conveying one's learning process to others (Schraw et al., 2006).

Whitebread et al. (2009) propose an instrument for evaluating metacognition and self-regulation in children aged 3 - 5 that consists in the observation and coding of video recordings (to identify metacognition indicators) and the construction of a checklist known as CHILD 3-5 (Children's Independent Learning Development), which can be used by teachers in the classroom (Whitebread et al., 2009). These researchers note that evaluation instruments of this type are more effective than self-report instruments, which are especially problematic in early childhood due to methodological issues derived from the verbal skills that they require (Whitebread et al., 2007).

In Chile, Jaramillo and Osses (2012) constructed and validated a Likert-type instrument to measure metacognition in terms of knowledge, metacognitive experiences, and cognitive self-regulation. This self-report questionnaire is targeted at 7th and 8th grade students who attend municipal schools in the Araucanía region.

The instrument examined in this article evaluates respondents' use of metacognitive skills (metacognitive knowledge) and the metacognitive skills observable in the way they tackle the tasks presented. Due to the characteristics of its administration —explained in more detail in the next section— the instrument also considers the mediator's role in metacognition (Veenman et al., 2006), since this individual not only observes, but also accompanies the student.

According to the analysis performed by Lai (2011) in her theoretical review, assessing metacognition poses a challenge for several reasons: (1) metacognition is a complex construct, (2) it is not directly observable, (3)

it can be confused with verbal ability and working memory, and (4) existing measures tend to be limited and decontextualized. Taking into account these characteristics and conditions, the evaluation instrument described and analyzed in this article provides a new way of negotiating these challenges from the perspective of schools and within the classroom.

Methodological framework of the Educación 2020 metacognitive evaluation instrument

> Two things are important: to make teaching visible for students and to make learning visible for teachers. Educational outcomes should improve as students become teachers and teachers become students (Hattie, 2009, p.22).

This notion, put forth by John Hattie in his book Visible Learning, is compatible with the Tutoring Networks model and its focus on the metacognitive domain of learning.

It is well known that an excessive curricular focus on conceptual elements prevents teachers from reflecting on how students learn, while also limiting opportunities for students to be aware of how they are learning.

In essence, Tutoring Networks allows students to become gradually familiar with this practice. Therefore, it is relevant to analyze the continuum of the consolidation of metacognitive practices in students as they progress, from the start of the tutoring experience to the moment they become tutors themselves. Roughly speaking, this is the trajectory followed by the participants of a Network, which lasts some 18 months.

Progressing from tutee to tutor is expected to allow students to develop metacognitive and autonomous learning skills; in other words, they should "learn how to learn" while also improving their reading comprehension and problem-solving skills. Also, at a school level, this initiative should make it possible to observe what Mexican researchers have described and documented: a transformation of school culture, supported by the democratization of learning.

One of the hypotheses regarding the implementation of the project is that, when students become the protagonists of their learning process (which is expected to occur during the tutoring relationship), they will attain better learning outcomes and develop higher-level skills such as metacognition.

This article cannot detail the whole evaluative framework of the Educación 2020 Tutoring Networks; however it should be pointed out that the metacognitive evaluation instrument is part of the following battery of indicators and tools to be applied when implementing this innovative strategy in schools:

a. Analysis of the internal efficiency indicators of each school.
b. Conversation-based workshops for students, teachers, and families focused on expectations and beliefs linked to teaching/learning processes in schools.
c. A survey of teachers' and students' motivation, self-efficacy, and perceptions of teaching practices.
d. Evaluation of teaching practices and school leadership.
e. Evaluation of student learning.
f. Metacognitive evaluation.

The instruments used in the evaluation of students' learning and metacognitive skills are largely based on the bimodal matrix presented in a prior section (Anderson et al., 2001; in Arratia, Flotts, & Rodríguez, 2012), covering the four types of knowledge (conceptual, factual, procedural, and metacognitive) and the six cognitive processes (remembering, understanding, applying, analyzing, evaluating, and creating).

Description of the instrument. It is a written test that can be administered to students aged 11 - 16, although it can be used with younger children, older adolescents, and adults given that, rather than the nature and complexity of the tasks presented, researchers are interested in examining the metacognitive skills that respondents apply when solving them.

It comprises two sections. The first includes two logic exercises, one that requires students to arrange geometrical figures according to certain criteria and another in which they need to arrange elements following clues. Both exercises require: (a) understanding the task presented, (b) developing an answer through inference and the use of tentative paths or solutions, and (c) testing the answer by rereading the task. With respect to the conceptual framework presented, this first section evaluates cognitive and metacognitive skills.

The second section is a self-reflection worksheet for making the metacognitive process visible. Students are expected to express in writing the steps and strategies that they used to manage their knowledge and complete one of the tasks included in section 1. They are asked to choose one of the two tasks and answer a series of questions on it as a way of reflecting their awareness of the mental processes that they used to direct certain sets of knowledge and skills toward the execution of this task. This section refers to the declarative metacognitive knowledge that students employ when working on a task. This concept is explained in the conceptual framework section.

The Self-Reflection Worksheet comprises eight open-ended questions, two of which are divided into sub-items, which were designed and organized around three domains: planning, follow-up, and evaluation. These domains are aligned with the theoretical metacognitive construct defined previously. With respect to the specifications table of the evaluation instrument, they are weighted at 30, 30, and 40% respectively. The questions included in this Worksheet and their associated domains can be found in Table 1 (see Appendix 1):

- Planning:
    - (1) In your own words, explain what the task that you just completed is about.
    - (2) When you read the instructions, did you think the task would be easy or hard? Why?
- Follow-up:
    - (3) What knowledge of yours helped you solve the task?
    - (6) Write down what you thought and what you did while you solved the task.
- Evaluation:
    - (4) What was the hardest part of the task?
    - (5) What was the easiest part of the task?
    - (7) Write down what you thought and what you did after solving the task.
    - (8) How would you evaluate your performance in this task? Tick one of the options and support your choice.
        - ( ) Very good, because…
        - ( ) Good, because…
        - ( ) So-so, because…
        - ( ) Poor, because...

The evaluation instrument includes a rubric for reviewing, scoring, and delivering results based on the above questions. This rubric covers the eight questions included in the Self-Reflection Worksheet, plus two aspects to be examined: (a) Did the student solve the exercises correctly? (b) Did the student adjust his/her answers to one of the exercises completed? Item (a) is not scored; however, it is used to score question 8, which asks students to self-assess their performance while solving the task.

The rubric provides a choice of three performance levels: not yet achieved - achieved - remarkable. Per-question levels were adjusted after a set of piloting sessions that made it possible to better align theoretical descriptors with respondents' actual answers. The rubric was calibrated in order to optimize the scoring process, which consists in assigning a certain number of points to each answer according to the performance level reached. Based on the score assigned to each answer, a per-domain performance report is provided, associated with the definitions and descriptions of the levels included in the rubric.

Administration of the metacognitive evaluation instrument. Although the instrument described here can be used in a variety of classroom environments as long as the teacher is interested in working on his/her students' metacognition, Educación 2020 has designed it to be administered as part of the implementation of the Tutoring Networks program, working on the assumption that tutoring practice makes visible and trains processes linked to reflection about one's personal resources for solving a task and about the learning achieved while solving it.

Thus, competent learners use their knowledge to self-regulate their learning effectively (Mateos, 2001), which highlights the strategic value of knowing how to plan, monitor, and determine which techniques should be used, at which moments, and how, depending on the learning challenges encountered. Each activity implemented as part of the Tutoring Networks program is linked to key tasks that can be fulfilled by a competent learner who uses his/her metacognitive knowledge and self-regulates his/her learning. So, for instance, in a tutoring relationship, the tutee develops plans that are consistent with the [aims suggested by the] topic to be studied, is able to identify and change the strategy used, and can describe the process conducted to obtain a certain result or solution. With practice, individuals should be able to perform these activities at a higher level of proficiency.

From this perspective, metacognitive knowledge is also considered to be learnable and trainable; therefore, it can be evaluated at multiple moments. In essence, Tutoring Networks allows students to become gradually familiar with this practice. Therefore, it is relevant to analyze the continuum of the consolidation of metacognitive practices in students as they progress, from the start of the tutoring experience to the moment they become expert tutors. Roughly speaking, an expert tutor is a student who is able to engage autonomously and proactively with a topic, developing varied and contextualized repertoires and proposals aimed at familiarizing another learner with the topic.

The metacognitive evaluation instrument proposed within the context of the Tutoring Networks program is administered at an initial stage, when the school and the students who participate voluntarily join the training process and experience the Networks. This first evaluation is expected to provide a preliminary snapshot of students' metacognitive perceptions and skills. Later on, after having experienced the training process as tutors, the instrument should be administered again, in order to determine how much progress students have made. On this second occasion, the tasks to be completed will be different (first section of the instrument), but the questions included in the Self-Reflection Worksheet (second section) will remain unchanged in order to compare both evaluation instances.

Considering that the metacognitive evaluation proposed takes place in a school context, that Tutoring Networks involve the participation of both students and teachers, and that every evaluation is also a learning opportunity, multiple evaluation stages or phases can be determined:

1. Teachers learn about the evaluation instrument and its theoretical-methodological basis.
2. Teachers administer the metacognition instrument to students, following a previously discussed evaluation protocol.

3. Teachers receive a report of students' metacognitive level in several domains.
4. Students reflect on metacognition based on the results received.

In line with one of the core principles of the Tutoring Networks program, i.e. the importance of person-to-person relationships, the instrument should ideally be administered one-on-one. However, it can also be administered to groups if the necessary conditions are met. In both cases, certain suggestions are included in the administration protocol and discussed with the participating teachers beforehand: (a) Explain the context of the evaluation session, its aims, and its main characteristics, stressing that it is not graded. (b) The administration of the instrument can last no more than 50 minutes. (c) Students are allowed to use the material in any way they wish (highlighting, coloring, underlining, erasing) and use blank sheets if they need them. It is important to boost students' confidence in their ability to solve the tasks presented and think about their own learning. (d) Teachers cannot assist students when they are completing the Self-Reflection Worksheet, unless they have reading comprehension (decoding) problems. If this occurs, the incident must be recorded in the administration log. (e) At the end of the evaluation session, teachers must guide a conversation about the exercise completed and the answers reported in the Worksheet, in order to allow students to explain their responses in more detail. (f) At the end of the session, the teacher thanks students for participating and stores the administration records.

The instruments are processed in a spreadsheet that is used to assign a score to each answer. This yields a quantitative report of per-group results and a qualitative report of individual results, which specifies the level attained by the student in each domain. This information forms the basis of the analysis of the results obtained, a process conducted with teachers and students according to the phases of the instrument administration process.

Analysis of results. Before analyzing the results, it is necessary to: (a) know the underlying construct of the evaluation instrument and the focus of its dimensions; (b) consider that the format of the evaluation instrument has limitations, and (c) bear in mind that a one-school sample is not significant enough to draw general conclusions. Taking these factors into account, the results obtained provide information that invites learning communities to reflect on various aspects of learning:

- What metacognitive domains are the most and least developed in students?
- What is needed to plan actions within the context of the execution of a task?
- How to monitor or supervise the execution of a task?
- How to evaluate the results of a task?
- Is it easy to identify factors that boost and/or hinder the effectiveness of a strategy?
- What pedagogical actions are being conducted at an institutional or individual level to foster the development of these metacognitive processes?

Analyzing the results obtained through an instrument of this type invites teachers and students to acknowledge their role as learners, which should help them identify their learning difficulties and preferences. This is expected to make it possible to better adjust attainment expectations while also facilitating the development of learning challenges that are better suited to the characteristics of the people taking part in the process. Encouraging students to reflect on how they learn allows them to make decisions about their learning and improves the self-regulation of the cognitive processes involved. This requires identifying the aims of a given learning opportunity and establishing substantial connections with the new information or skills needed to tackle it. All these elements can contribute to the generation of more meaningful learning, which is necessary for progressing toward quality education.

Discussion and conclusions

An instrument for the evaluation of metacognition in the classroom can greatly aid the transformation of this learning space and the learning processes that emerge in it, while also facilitating the development of the skills needed in the 21st century. Although this is a socially well-established and accepted ideal, it is still necessary to produce theoretical and empirical research to support a type of training that will enable teachers to innovate in the classroom, in order to improve the performance of today's students and cater to their needs, characteristics, and interests. It has been determined that developing processes and instruments that allow students to reflect on their own learning is indeed helpful for both teachers and students.

In order to continue developing the metacognitive evaluation instrument presented here, the next step is to administer it to significant samples of students, which will make it possible to determine its validity and reliability. As of this writing, the instrument has been administered to about 190 people, which has made it possible to calibrate and adjust it; however, it is still necessary to conduct a validity and reliability analysis with a larger N. In this vein, future challenges include performing a content validity analysis with expert raters, an inter-rater validity analysis of the rubric, and a correlation analysis comparing the instrument's results with those of others, such as the self-report tool developed by Jaramillo and Osses (2012). In line with the conceptual framework presented, it would be interesting to conduct a correlation analysis with an observation method based on video recordings (Whitebread et al, 2009) or other audiovisual materials in order to analyze the discussion process at the end of the evaluation session.

The instrument proposed is not free from methodological limitations common to evaluations of this nature. Even though it is targeted at participants who, due to their stage of development (approximately 11 to 16 years of age), have enough verbal skills to report their thought processes, the instrument requires writing skills, which can pose difficulties for students, or at least some of them. Some studies have demonstrated the importance of implicit metacognitive processes, which do not necessarily translate into a written report (Reder & Schunn, 1996; Siegler, 1996, in Whitebread et al. 2009). The administration protocol provides guidelines to address these limitations, which can be better controlled in individual administration settings and not in a group session. The choice to develop a paper-based instrument with open-ended questions was influenced by feasibility and scope criteria: it is a low-cost, easy-to-administer instrument that can be used with large numbers of participants. Also, the rubric does not assign higher scores to higher-quality texts, focusing instead on the content of the idea presented. To some extent, this decision also offsets the limitations of writing as an evaluation device.

It should also be considered that one of the aims of the Tutoring Networks program is the development of verbal skills for expressing and becoming aware of one's own thought process. Therefore, one of the challenges for Educación 2020 is to research this aim and generate evidence related to it, for instance, through the inclusion of complementary methodologies such as video or audio observation and coding; however, analysis guidelines have yet to be developed, which is a task that must be completed in order to improve this evaluation tool.

Another challenge that the Tutoring Networks methodology entails and which goes beyond the instrument presented is the evaluation of the tutor as a mediator of metacognition *during* his/her tutoring work. The instrument presented can be used to collect information leading to hypotheses about the role of tutors and the strategies employed to generate or facilitate metacognition; yet, it is still necessary to directly evaluate the tutor-tutee relationship and the key moments or actions that trigger the development of skills.

Finally, future goals derived from this metacognitive evaluation instrument include drawing more precise conclusions about its results and systematizing the work and reflection activity conducted with the teachers and students who take part in the process. Both these elements are sources of valuable information for enriching the development of materials that, by strengthening certain processes, can improve students' connection with their own learning.

The original article was received on November 23rd, 2016
The article was accepted on April 10th, 2018

References

Agencia de la Calidad de la Educación (2014). Sobre Resolución de Problemas: Análisis de los Resultados de la Prueba PISA 2012. *Apuntes sobre la Calidad de la Educación, 15.* Recuperado de https://s3.amazonaws.com/archivos.agenciaeducacion.cl/documentos-web/Papers/2014_15_Sobre_resolucion_de_problemas_analisis_resultados_prueba_PISA_2012.pdf

Agencia de la Calidad de la Educación (2016). Resultados Educativos 2016. Recuperado de: http://archivos.agenciaeducacion.cl/ResultadosNacionales2016_.pdf

Klimenko, O. & Alvares, J. (2009). Aprender cómo aprendo: la enseñanza de estrategias metacognitivas. *Educación y Educadores, 12*(2), 11-28. Recuperado de http://www.redalyc.org/articulo.oa?id=83412219002

Kuhn, D. & Dean, D. (2004). A bridge between cognitive psychology and educational practice. *Theory into Practice, 43*(4), 268-273.

Lai, E. (2011). Metacognition: A Literature Review. Pearson Research Reports. Recuperado de http://images.pearsonassessments.com/images/tmrs/metacognition_literature_review_final.pdf

Ley General de Educación No.20.370 (2009). Biblioteca del Congreso Nacional de Chile / BCN. Santiago, Chile.

López, D. (2016). La relación tutora y la mejora del aprendizaje. *Revista Red: Cubo de Ensayo*. Recuperado de http://www.inee.edu.mx/images/stories/2016/revista_red/articulos/02Cubo.pdf

Mateos, M. (2001). *Metacognición y Educación*. Argentina: Aique Grupo Editorial.

Ministerio de Educación de Chile (Mineduc) (2015). Análisis de indicadores educativos de Chile y la OCDE en el contexto de la Reforma Educacional. *Serie Evidencias, 31.* Recuperado de http://centroestudios.mineduc.cl/tp_enlaces/portales/tp5996f8b7cm96/uploadImg/File/Evidencias/Evidencias%20final_noviembre_2015.pdf

Monereo, C. (2003). La evaluación del conocimiento estratégico a través de tareas auténticas. *Revista de Investigación Educacional Latinoamericana, 32*(1), 71-89. Recuperado de http://pensamientoeducativo.uc.cl/index.php/pel/article/view/246/516

OECD (2013). Innovative Learning Environments. *Educational Research and Innovation*. Francia, París: autor.

Osses, S. & Jaramillo, S. (2008). Metacognición: un camino para aprender a aprender. *Estudios Pedagógicos, 34*(1), 187-197. Recuperado de https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-07052008000100011

PREAL (2009). Tendencias en la evaluación del logro escolar: la experiencia de cinco países con alto rendimiento educativo. *Serie mejores prácticas, 11*(31). Recuperado de http://www.oei.es/historico/noticias/spip.php?article5386

Ravela, P. (2013). La evaluación de y en la educación. *Boletín del Instituto Nacional de Evaluación Educativa.* Recuperado de: https://www.ineed.edu.uy/images/pdf/Aristas-La-evaluacin-de-y-en-la-educacin-I.pdf

Red Latinoamericana de Organizaciones de la Sociedad Civil por la Educación [REDUCA] (2017). *Súper Profes y Súper Directivos. Estudio de percepción estudiantil sobre la caracterización de competencias docentes y directivas en América Latina y el Caribe y sus marcos regulatorios*. Recuperado de: http://www.reduca-al.net/files/observatorio/reportes/Super_docentes_y_directivos.pdf

Reder, L. & Schunn, C. (1996). Metacognition does not imply awareness; strategy choice is governed by implicit learning and memory. *Implicit memory and metacognition, 3,* 45-77. Recuperado de http://memory.psy.cmu.edu/publications/96_lmr_cds.pdf

Rincón-Gallardo, S. (2011). La transformación del núcleo pedagógico como fuente de motivación docente en escuelas públicas mexicanas. *Didac, 59,* 30-36. Recuperado de http://revistas.ibero.mx/didac/uploads/volumenes/1/pdf/59.pdf

Rincón-Gallardo, S. (2014). Innovación pedagógica a gran escala: ¿lujo o imperativo moral?. *Didac, 65*, 11-18. Recuperado de http://revistas.ibero.mx/didac/articulo_detalle.php?pageNum_paginas=1&totalRows_paginas=8&id_volumen=19&id_articulo=227&pagina=1

Ruiz, C. (2002). Mediación de estrategias metacognitivas en tareas divergentes y transferencia recíproca. *Revista Investigación y Postgrado*, *17*(2), 53-82. Recuperado de http://www.scielo.org.ve/scielo.php?script=sci_arttext&pid=S1316-00872002000200003&lng=es&tlng=es.

Santos, M. (1993). La evaluación: un proceso de diálogo, comprensión y mejora. *Revista Investigación en la Escuela, 20*, 23-38. Recuperado de http://investigacionenlaescuela.es/index.php/revista-investigacion-en-la-escuela/435-la_evaluacin_un_proceso_de_dilogo_comprensin_y_mejora

Schraw, G. & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460-475.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36, 111-139. http://dx.doi.org/10.1007/s11165-005-3917-8

Secretaría de Educación Pública (2012). *Redes de Tutoría Académica: Orientaciones para su Gestión en las regiones y escuelas.* México, DF: autor.

Torrano, F. & González, M. (2004). El aprendizaje autorregulado: presente y futuro de la investigación. *Revista Electrónica de Investigación Psicoeducativa, 2*(1), 1-34. Recuperado de http://www.webdocente.altascapacidades.es/Aprendizaje%20Autorregulado/Art_3_27.pdf

UNESCO-OREALC. (2017). Reporte: Educación y habilidades para el siglo XXI. *Reunión Regional de Ministros de Educación de América Latina y el Caribe, Buenos Aires, Argentina, 24 y 25 de enero 2017*. Santiago: OREALC/UNESCO.

UNESCO-OREALC. [Bases del Futuro Construyendo el Currículum] (2017, marzo 14). Consulta Jóvenes ¿Qué quieres aprender? [Archivo de video]. Recuperado de https://www.youtube.com/watch?v=f4DbKR5NxYk

UNICEF (2015). *El aprendizaje bajo la lupa: Nuevas perspectivas para América Latina y el Caribe*. Panamá: autor. Recuperado de https://www.unicef.org/lac/UNICEF_Aprendizaje_bajo_la_lupa_nov2015(1).pdf

Veenman, M., Van Hout-Wolters, B. & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition Learning,1*, 3–14. http://dx.doi.org/10.1007/s11409-006-6893-0

Villalón, M. (2002). El proceso metacognitivo de la información escrita: un desafío para la educación contemporánea. *Revista de Investigación Educacional Latinoamericana, 30*(1), 177-196. Recuperado de http://pensamientoeducativo.uc.cl/index.php/pel/article/view/218

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *The Journal Of Educational Research*, *84*(1), 30-43. http://dx.doi.org/10.1080/00220671.1990.10885988

Whitebread, D., Bingham, S., Grau, V., Pasternak, D. P., & Sangster, C. (2007). Development of metacognition and self-regulated learning in young children: Role of collaborative and peerassisted learning. *Journal of Cognitive Education and Psychology, 6*(3), 433-455.

Whitebread, D., Coltman, P., Pino Pasternak, D., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition Learning, 4,* 63–85. http://dx.doi.org/10.1007/s11409-008-9033-1

Zepeda, S. (2008). Relaciones entre evaluación de aprendizajes y práctica pedagógica: explorando la estrategia de acompañamiento pedagógico. *Revista de Investigación Educacional Latinoamericana, 43*(2), 243-259. Recuperado de http://pensamientoeducativo.uc.cl/index.php/pel/article/view/440

Zimmerman, B. & Schunk, D. (2011). *Handbook of Self-Regulation of Learning and Performance*. New York, EEUU: Taylor & Francis.

Appendix

Table 1
Metacognitive Evaluation Instrument: Self-Reflection Worksheet questions.

Question	Associated domain
1. In your own words, explain what the task that you just completed is about.	Planning
2. When you read the instructions, did you think the task would be easy or hard? Why?	Planning
3. What knowledge of yours helped you solve the task?	Follow-up
4. What was the hardest part of the task?	Evaluation
5. What was the easiest part of the task?	Evaluation
6. Write down what you thought and what you did while you solved the task.	Follow-up
7. Write down what you thought and what you did after solving the task.	Evaluation
8. How would you evaluate your performance in this task? Tick one of the options and support your choice. ( ) Very good, because… ( ) Good, because… ( ) So-so, because… ( ) Poor, because...	Evaluation