Computers in Human Behavior 80 (2018) 441–459

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Extending the nomological network of computational thinking with non-cognitive factors

Marcos Román-González [a,*], Juan-Carlos Pérez-González [a], Jesús Moreno-León [b,c], Gregorio Robles [c]

[a] Universidad Nacional de Educación a Distancia (UNED), Faculty of Education, C/ Juan del Rosal, nº 14, 28040, Madrid, Spain
[b] Programamos.es, Seville, Spain
[c] Universidad Rey Juan Carlos (URJC), ETSI Telecomunicación, Camino del Molino s/n, 28943, Fuenlabrada (Madrid), Spain

A R T I C L E I N F O

Article history:
Received 30 April 2017
Received in revised form
9 September 2017
Accepted 23 September 2017
Available online 25 September 2017

Keywords:
Computational thinking
Computational thinking test
Code literacy
Computer science education
Personality
Self-efficacy
Educational psychology
Primary education
Secondary education

A B S T R A C T

Computational thinking (CT) is being consolidated as a key set of problem-solving skills that must be developed by the students to excel in our software-driven society. However, in psychological terms, CT is still a poorly defined construct, given that its nomological network has not been established yet. In a previous paper, we started to address this issue studying the correlations between CT and some fundamental cognitive variables, such as primary mental abilities and problem-solving ability. The current work deepens in the same direction as it aims to extend the nomological network of CT with non-cognitive factors, through the study of the correlations between CT, self-efficacy and the several dimensions from the 'Big Five' model of human personality: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. To do so, the Computational Thinking Test (CTt) and some additional self-efficacy items are administered on a sample of 1251 Spanish students from 5th to 10th grade ($N = 1251$), and the Big Five Questionnaire-Children version (BFQ-C) is also taken by a subsample from the above ($n = 99$). Results show statistically significant correlations between CT and self-efficacy perception relative to CT performance ($r_s = 0.41$), in which gender differences in favor of males are found ($d = 0.42$). Moreover, results show statistically significant correlations between CT and: Openness to Experience ($r = 0.41$), Extraversion ($r = 0.30$), and Conscientiousness ($r = 0.27$). These findings are consistent with the existing literature except for the unexpected correlation between CT and the Extraversion factor of personality, which is consequently discussed in detail. Overall, our findings corroborate the existence of a non-cognitive side of CT that should be taken into account by educational policies and interventions aimed at fostering CT. As a final contribution, the extended nomological network of CT integrating cognitive and non-cognitive variables is depicted.

1. Introduction

Computer programs are increasingly mediating our existence. More and more, we are experimenting an algorithmic life (Sadin, 2015), as plenty of objects driven by software surround us and condition our interactions with reality (Manovich, 2013). Given this disruptive scenario, it is becoming indispensable to handle, in a broad sense, the language of computers to participate fully and effectively in the digital reality (Rushkoff, 2010).

Within this context, the term 'code-literacy' has been coined to define the set of actions of teaching and learning to read-write with computer programming languages (Prensky, 2008; Rushkoff, 2012). Thus, we consider that a person is code-literate when he/she is able to read and write in the language of computers, and to think computationally (Román-González, 2014). While code-literacy refers ultimately to an emerging read-write practice, computational thinking (CT) refers to the underlying problem-solving (only cognitive?) process that supports and allows it. In other words, computer programming is the fundamental way that enables CT to come alive (Lye & Koh, 2014), although CT can be projected on

* Corresponding author.
E-mail addresses: mroman@edu.uned.es (M. Román-González), jcperez@edu.uned.es (J.-C. Pérez-González), jesus.moreno@programamos.es (J. Moreno-León), grex@gsyc.urjc.es (G. Robles).

https://doi.org/10.1016/j.chb.2017.09.030
0747-5632/

different kinds of problems that may not involve directly programming tasks (Wing, 2008).

Therefore, CT is becoming considered all around the world as a key set of problem-skills that should be acquired by the emerging generations of students (Balanskat & Engelhardt, 2014; Bocconi et al., 2016; García-Peñalvo, 2016; García-Peñalvo, Reimann, Tuul, Rees, & Jormanainen, 2016). However, up to now there is no consensus about a formal definition of CT (Gouws, Bradshaw, & Wentworth, 2013; Kalelioglu, Gülbahar, & Kukul, 2016). We are also witnessing disagreements on how CT should be incorporated in educational curricula (Lye & Koh, 2014), and there is a lack of tools to measure and assess CT (Grover & Pea, 2013; Grover, Cooper, & Pea, 2014; Kalelioglu et al., 2016). Furthermore, in psychometric terms, CT is still a poorly defined construct, given that its nomological network has not been established yet. That is, the correlations between CT and other psychological variables, whether cognitive or non-cognitive, have not been completely reported by the scientific community.

In a previous paper we started to address this issue studying the correlations between CT and some fundamental cognitive variables (Román-González, Pérez-González, & Jiménez-Fernández, 2017). There, we reported statistically significant correlations at least moderately intense between CT and problem-solving ability ($r = 0.67$), reasoning ability ($r = 0.44$), and spatial ability ($r = 0.44$). Overall, the primary mental abilities could explain only 27% of the CT variance. This result led us to affirm that CT is, to some extent, an independent psychological construct, distinct from the traditional aptitudes; and to wonder about the unexplained remaining 73%. In summary, this first work empirically corroborated the conceptualization of CT as mainly a problem-solving ability, linked with *g* or fluid intelligence; a fact that had been theoretically stated by many other authors in earlier years (Brennan & Resnick, 2012; Lye & Koh, 2014; Wing, 2006, 2008). Moreover, these prior results are consistent with the framework just described by Kalelioglu et al. (2016), in which CT is defined as a complex and high-order thinking skill involved in problem-solving processes. Finally, they are also consistent with recent theoretical proposals (Ambrósio, Xavier, & Georges, 2014) linking CT to some components of the Cattel-Horn-Carroll (CHC) model of intelligence (McGrew, 2009; Schneider & McGrew, 2012), such as *fluid reasoning* (G_f), *visual processing* (G_v), and *short-term memory* (G_{sm}).

Nevertheless, to the best of our knowledge no empirical research has been conducted studying the correlations between CT and non-cognitive variables. In order to fill this gap, we have already presented a preliminary study (Román-González, Pérez-González, Moreno-León, & Robles, 2016), which will be significantly extended along the present paper. Thus, in this work we specifically attempt to answer the following research questions:

- *RQ_1: Does CT correlate with self-efficacy?*
- *RQ_2: Does CT correlate with personality?*
- *RQ_3: What are the personality profiles of top and low computational thinkers?*
- *RQ_4: How much variance of CT can be explained by personality?*

These questions are plausible since in the literature there is prior evidence of relationships between cognitive and non-cognitive variables, such as self-efficacy and personality, which are described in subsections 2.2 and 2.3. In addition, answering these research questions, even tentatively, is relevant as it may contribute to extend the nomological network of CT and consolidate its consistency as an emerging psychological construct. Finally, the answers may help to understand better which non-cognitive factors underlie CT, in order to promote and optimize its development in educational settings.

2. Background

2.1. Computational thinking

A decade ago, Jeannette Wing defined that CT "involves solving problems, designing systems, and understanding human behavior, by drawing on the concepts fundamental to computer science" (Wing, 2006, p. 33). Thus, we could state that CT's essence is thinking like a computer scientist when confronted with a problem. The phrase received some criticism and, even though new proposals emerged, the computer science education community has had difficulties in finding an agreement in defining the term (Grover & Pea, 2013; Kalelioglu et al., 2016). As a response, in 2011 Wing clarified that CT "is the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent" (Wing, 2011, on-line). One year later, this definition was simplified by Aho, who conceptualizes CT as the thought processes involved in formulating problems so "their solutions can be represented as computational steps and algorithms" (Aho, 2012, p. 832).

However, these generic definitions quoted above are not enough from a psychometric approach. It is necessary to count on operational definitions of CT to enable and guide its development, measurement and assessment as a solid psychological construct. In this direction, the Computer Science Teachers Association (CSTA) and the International Society for Technology in Education (ISTE) stated in 2011 an operational definition of CT that provides a framework and common vocabulary for Computer Science K-12 educators. This definition states that CT is a "problem-solving process that includes (but is not limited to) the following characteristics: formulating problems in a way that enables us to use a computer and other tools to help solve them; logically organizing and analyzing data; representing data through abstractions such as models and simulations; automating solutions through algorithmic thinking (a series of ordered steps); identifying, analyzing, and implementing possible solutions with the goal of achieving the most efficient and effective combination of steps and resources; generalizing and transferring this problem solving process to a wide variety of problems" (CSTA & ISTE, 2011).

Moreover, the aforementioned operational definition continues saying that "these (problem-solving) skills are supported and enhanced by a number of dispositions or attitudes that are essential dimensions of CT. These dispositions or attitudes include:

- Confidence in dealing with complexity.
- Persistence in working with difficult problems.
- Tolerance for ambiguity.
- The ability to deal with open-ended problems.
- The ability to communicate and work with others to achieve a common goal or solution" (CSTA & ISTE, 2011).

In other words, the existence of a non-cognitive side of CT is accepted in a broad sense, which we have already started to confirm empirically (Román-González et al., 2016). In this paper, we continue to investigate on this regard.

2.2. Self-efficacy

Self-efficacy is defined as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performance" (Bandura, 1986, p. 391). In other words, we can state that self-efficacy is an individual's judgement of his/her ability to perform a task within a specific domain.

Self-efficacy is crucial in learning given that a "competent functioning requires both skills and self-beliefs of efficacy to use them effectively" (Bandura, 1986, p. 391). In learning contexts, self-efficacy influences decisively the amount of effort expended, which type of coping and solving problems strategies are adopted, how much persistence is sustained in the face of failure, and finally the performance outcomes. Following the model of Bandura (1977, 1986), it is assumed that judgments and expectations of self-efficacy rely on four principal sources of information: the previous subject's performance attainments; vicarious experiences derived from the observation of the performance of others (especially if these 'others' are considered as peers); verbal persuasion and associated types of social influences; and physiological states from which people partly judge their capableness, strength, and vulnerability. This model has been validated in multiple domains, including the specific one of learning to program (e.g., Kukul, Gökçearslan, & Günbatar, 2017; Ramalingam, LaBelle, & Wiedenbeck, 2004).

The impact of self-efficacy on students' academic performance is also well documented in the literature. Overall, higher levels of self-efficacy increase the chances of academic success, since people with confidence in their abilities are more likely to cope with difficulties (Bandura, 2001). Thus, a very large meta-analysis on the psychological correlates of university students' academic performance showed a medium-sized correlation between 'academic self-efficacy' and 'grade point average' (GPA), and a large correlation between 'performance self-efficacy' and GPA (Richardson, Abraham, & Bond, 2012). Also with undergraduate students, the research of Komarraju and Nadler (2013) showed that self-efficacious students are able to achieve academically because they monitor and self-regulate their impulses and persist in the face of difficulties. Moreover, several studies conducted at university level proved that self-efficacy is a significant predictor not only of academic achievement but also of learner satisfaction (Joo, Lim, & Kim, 2013). Finally, similar results have also been found in investigations involving middle and high school students, which proved again that self-efficacy is a significant predictor of academic achievement (e.g., see Zuffianò et al., 2013).

With respect to gender differences in academic self-efficacy, the large meta-analysis of Huang (2013) identified an overall effect size of 0.08, with a small difference in favor of males. However, their further analysis demonstrated that content domain was a significant moderator in explaining the variation of the effect size. Thus, females showed higher language arts self-efficacy than males. Meanwhile, males displayed higher mathematics, computer, and social sciences self-efficacy than females. Another interesting result was that gender differences in academic self-efficacy also varied with age (e.g., for mathematics self-efficacy, the significant gender differences emerged in late adolescence).

Therefore, it should be noted that self-efficacy is very specific to a certain activity: a person may have high self-efficacy in one domain, such as drawing, and low self-efficacy in another, such as computer programming. Accordingly, some authors have investigated the specific effects of students' self-efficacy on learning to program (e.g., see Ramalingam et al., 2004, for novice programmers enrolled in a CS1 course, where a significant correlation $r = 0.23$ between self-efficacy and grade is found).

Furthermore, specific differences in terms of levels of self-efficacy between male and female university students regarding their computing ability have been widely investigated in the last decades. In general, most studies have found lower levels of confidence for females in spite of achieving similar levels of performance than their male counterparts. This pattern of results appear both in the context of undergraduate students of business administration (Busch, 1995), within students enrolled in STEM degrees (Askar & Davenport, 2009), and even for PhD candidates, as female doctoral students in Computer Science have less confidence than male students in that they can achieve their educational goals (Cohoon, 2007). These gender differences in computer self-efficacy seem to be the in the base of the current women underrepresentation in Computer Science field, along with other gender differences in stereotypes, interests, values, interpersonal orientation, and personality (Beyer, 2014).

Finally, when we focus on K-12 education, which is the specific scope of this paper, we find a recent wave of research related on how the new visual programming languages, such as Scratch or Blockly, are contributing to increase motivation and self-efficacy of primary and secondary students in programming tasks and subsequently facilitating their transition to professional textual programming languages (Armoni, Meerbaum-Salant, & Ben-Ari, 2015). Nevertheless, even within these new block-based 'learning to program' contexts, gender differences in self-efficacy are found again. This is the case in an intervention with 49 young Spanish students between 7 and 14 years old who participated in a coding workshop (Espino & González, 2016). After the training, boys showed higher levels of confidence in their ability to create their own informatics content. Similar results are found in a study with 340 fifth and sixth-grade students from 7 elementary schools in Greece, where "boys had more positive self-efficacy and value beliefs about computers compared to girls, and were more likely to engage in *hard-core* computer activities such as programming" (Vekiri & Chronaki, 2008, p. 1400). The Hour of Code performed a survey study aimed to assess attitudes towards and self-efficacy with computer science in over 8000 participants in this initiative (Code.org, 2017). Again, male students showed higher levels of self-efficacy both before and after the activity. In summary, although visual block-based languages are lowering the entry barriers to computer programming in early ages, engaging students and promoting an optimal self-efficacy in coding tasks, it seems that gender differences tend to appear alike.

Therefore, it is necessary to deepen on how these self-efficacy differences are generated from the K-12 level if we desire to build an equal computer science education. This paper tries to shed some light on it.

2.3. 'Big Five' model of personality

In the last 30 years, a huge body of research has been gathered supporting the validity of a five-factor structure to describe human personality (the so-called 'Big Five') (Barbaranelli, Caprara, Rabasca, & Pastorelli, 2003; Goldberg, 1990). The Big Five model, whose validity has also been demonstrated in early childhood (Abe, 2005), late childhood and adolescence (Mervielde, Buyst, & De Fruyt, 1995), states the following labels for these five dimensions:

- *Openness to Experience* (O), also called *Intellect*: it refers to self-reported intellect, especially in the school domain, and broadness or narrowness of cultural interests, self-reported fantasy/creativity, and interest in other people.
- *Conscientiousness* (C): it refers to aspects such as autonomy, dependability, orderliness, precision, persistence, and the fulfilling of commitments.
- *Extraversion* (E), also called *Energy*: it refers to aspects such as sociability, activity, enthusiasm, assertiveness, and self-confidence.
- *Agreeableness* (A): it refers to aspects such as concern and sensitivity towards others and their needs, tendency to cooperation.

- *Neuroticism* (N), also called *Emotional Instability*: it refers to feelings of anxiety, depression, discontent, irritability, and anger.

Examples of items aimed at assessing each of the five factors are reported in paragraph 3.2.3, in which the Big Five Questionnaire-Children version (BFQ-C) is described.

Given that CT is conceptualized as mainly a problem-solving ability, linked with *g* or fluid intelligence (Román-González, Pérez-González, et al., 2017), its relationships pattern with the Big Five is expected to be similar to that found for the Big Five and the two main indicators of intellectual ability, namely general intelligence and academic performance (AP). Thus, it is necessary to review these issues in the following lines.

On the whole, correlations between personality factors and cognitive abilities are typically low ($r \leq 0.40$) (Von Stumm, Chamorro-Premuzic, Quiroga, & Colom, 2009). Regarding the relationships between personality and intelligence, *Openness* is the personality factor most frequently associated with intelligence as well as with creativity (Chamorro-Premuzic & Furnham, 2014; Higgins, Peterson, Pihl, & Lee, 2007; Von Stumm & Ackerman, 2013). Thus, positive correlations between *Openness* and Intelligence Quotient (IQ) have been typically reported (e.g., see the study of John, Caspi, Robins, Moffitt, and Stouthamer-Loeber (1994) with adolescent boys). Conversely, *Conscientiousness* is associated with persistence, self-discipline and achievement striving but it is usually only weakly related to intelligence (Chamorro-Premuzic & Furnham, 2014). An interesting approach to the particular case for *Conscientiousness* comes from the Moutafi, Furnham, and Crump (2003) 'compensation hypothesis', according to which the (slight) negative relationship between *Conscientiousness* and intelligence may be a consequence that less intelligent people would become more conscientious as a result of attempting to compensate for their low intellectual ability.

Regarding the relationships between personality and AP, literature shows that the two main personality factors positively correlated with AP are *Openness* and *Conscientiousness*; and that can be stated alike for childhood (Barbaranelli et al., 2003; Mervielde et al., 1995; Poropat, 2014) and adulthood (Chamorro-Premuzic & Furnham, 2008). Furthermore, *Conscientiousness* seems to be the strongest predictor of AP, both in K-12 education (Abe, 2005) and university (Higgins et al., 2007). In this sense, the meta-analysis by Poropat (2009, 2014) reports that *Conscientiousness* has similar validity to traditional intelligence measures as predictor of AP. Less frequently, some studies can be found reporting positive correlations between AP and the *Extraversion* personality factor, especially in primary and secondary education (Mervielde et al., 1995; Zuffianò et al., 2013); and between AP and *Agreeableness*, in tertiary education (Poropat, 2009).

Moreover, *Extraversion* and *Neuroticism* are positively correlated with externalizing problematic behavior syndromes (hyperactivity, transgressive/disruptive conduct, inattentiveness, and aggression); and *Neuroticism* is also positively correlated with internalizing problematic behavior syndromes (depression, anxiety, somatic complaints, and obsessiveness) (Barbaranelli et al., 2003).

From another approach, if we intersect the dispositions or attitudes that underlie CT according to the aforementioned operational definition from CSTA & ISTE (2011) with the five personality dimensions stated from the Big Five model (Barbaranelli et al., 2003; Goldberg, 1990), the crosstab in Table 1 is obtained.

Consequently, according to the above-mentioned throughout this subsection, it is expected that CT will correlate positively with *Openness* (O); and, to a lesser extent, with *Conscientiousness* (C) and with *Extraversion* (E). In subsections 4.2, 4.3, and 4.4 an empirical study aimed at verifying these hypotheses is conducted.

3. Method

3.1. Participants

Our main sample is composed by $N = 1251$ Spanish students from 24 different schools, and enrolled from 5th to 10th grade. From the total sample, 825 (65.9%) students belong to public schools, and 426 (34.1%) belong to private schools. The students enrolled from 7th to 10th grade, which belong to the Spanish Secondary Education, are coming from the region of Valencia, where Computer Science is an elective subject along this educational stage. The students enrolled in 5th and 6th grade, which belong to the Spanish Primary Education, are coming from the regions of Madrid and Seville, where Computer Science contents can be optionally integrated from a cross-curricular approach along these ages. Table 2 shows the distribution of the subjects by gender, grade and age. The sampling procedure was not probabilistic and intentional, and it was already detailed in our previous paper (Román-González, Pérez-González, et al., 2017).

The Computational Thinking Test (CTt) and the additional self-efficacy items, which will be described in paragraphs 3.2.1 and 3.2.2, were both administered in the whole sample when none of the $N = 1251$ subjects had been exposed before to programming lessons in the school. Besides, the Big Five Questionnaire-Children version (BFQ-C), which will be described in paragraph 3.2.3, was also taken by a subsample from the above composed by $n = 99$ individuals, distributed along both genders and all the grades involved in our research (Table 3).

3.2. Instruments

3.2.1. Computational Thinking Test (CTt)

The Computational Thinking Test[1] (CTt) is a multiple-choice instrument composed by 28 items, which are administered on-line (via non-mobile or mobile electronic devices) in a maximum time of 45 min. It aims to measure the development level of CT in the subject, and it is based on the following operational definition of the construct assessed: "CT involves the ability to formulate and solve problems by relying on the fundamental concepts of computing, and using the inherent logic of programming languages: basic sequences, loops, iteration, conditionals, functions and variables" (Román-González, 2015, p. 2438).

The CTt is mainly designed and intended for Spanish students between 12 and 14 years old (7th and 8th grade), although it can also be used in lower grades (5th and 6th) and upper grades (9th and 10th). Each item of the CTt is presented either in a 'maze' or in a 'canvas' interface, and it is designed according to the following three principal dimensions (Román-González, 2015; Román-González, Pérez-González, et al., 2017):

- **Computational concept addressed**: each item addresses one or more of the following seven computational concepts, ordered in increasing difficulty: Basic directions and sequences; Loops–repeat times; Loops–repeat until; If–simple conditional; If/else–complex conditional; While conditional; Simple functions. These 'computational concepts' are progressively nested along the test, and are aligned with the CSTA Computer Science Standards for the 7th and 8th grade (Seehorn et al., 2011).
- **Style of answers:** in each item, responses are presented in any of these two styles: 'visual arrows' or 'visual blocks'.

[1] Sample copy available at: https://goo.gl/P7fkFw.

Table 1
Crosstab intersecting CT dispositions/attitudes with the Big Five model of personality.

CT dispositions or attitudes	Big Five model				
	O	C	E	A	N
Confidence in dealing with complexity	*	–	/	–	–
Persistence in working with difficult problems	–	*	–	–	–
Tolerance for ambiguity	*	/	–	–	–
The ability to deal with open ended problems	*	–	–	–	–
The ability to communicate and work with others to achieve a common goal	–	–	*	/	–

*: Yes; /: Partly; -: No.

- **Required task:** depending on which cognitive task is required for solving the item: 'sequencing' ≈ stating in an orderly manner a set of commands, 'completion' of an incomplete set of commands, or 'debugging' an incorrect set of commands.

The content validation of the CTt has been already reported (Román-González, 2015), as well as its criterion validity regarding other cognitive constructs such as primary mental abilities (verbal, spatial, reasoning, and numerical) and problem-solving ability (Román-González, Pérez-González, et al., 2017). The reliability of the CTt in terms of internal consistency has been also reported ($\alpha_{overall} \approx 0.80$; increasing as it does the grade, from $\alpha_{5th\&6thgrade} = 0.72$ to $\alpha_{9th\&10thgrade} = 0.82$).

In our previous paper (Román-González, Pérez-González, et al., 2017), it has been already demonstrated that the CTt score, which can range from 0 to 28 points, is quasi-normally distributed along the measured population (i.e., symmetrically distributed with skewness ≈ 0). The mean of the CTt score in the validation sample (Table 2) was placed very close to the middle point of the scale ($\bar{X} = 16.38$), and the CTt score showed proper variability ($SD = 4.82$) so that it was possible to construct sensitive and discriminant percentiles for the target population. Finally, the CTt demonstrated an appropriate degree of difficulty (medium) for the aforementioned population, with an increasing difficulty along its items, as recommended in the design of abilities' tests (e.g, Carpenter, Just, & Shell, 1990; Elithorn & Telford, 1969). Furthermore, it has been already reported that the performance in the CTt increases with the grade (with an effect size of the increment $d \approx 0.40$–0.60 between each couple of grades). This finding is consistent with our assumption that CT is mainly a problem-solving ability that should therefore be linked to the cognitive development and maturity of the subjects (Ackerman & Rolfhus, 1999; Mayer, Caruso, & Salovey, 1999). Finally, a statistically significant difference was found in the CTt score in favor of the male group ($t = 5.374$; $p < 0.01$), resulting an effect size measured through Cohen's d (Cohen, 1992) equal to $d = 0.31$, that can be considered as a low-moderate effect.

Examples of CTt items translated into English are shown in Figs. 1–3, with their specifications detailed below.

In summary, we can affirm that the CTt is a reliable and valid test for assessing CT in students from 10 to 16 years. However, some of its limitations have been already pointed out. For instance, in terms of the CT framework of Brennan and Resnick (2012), the CTt is heavily focused on 'computational concepts', only covers 'computational practices' partly, and ignores 'computational perspectives'. Moreover, the CTt only provides a *summative-aptitudinal* assessment, which should be complemented with other tools designed from a *skill-transfer* approach such as Bebras Tasks[2] (Dagiene & Futschek, 2008; Dagiene & Stupuriene, 2014), and from a *formative-iterative* assessment perspective such as Dr. Scratch[3] (Moreno-León & Robles, 2015; Moreno-León, Robles, & Román-González, 2015). The convergent validity between the CTt and the aforementioned complementary assessment tools has been already reported (Román-González, Moreno-León, & Robles, 2017).

Finally, the CTt seems to be moderately gender biased as its items have a large visuospatial load that could be favoring males, given that there are several meta-analysis in literature that demonstrate higher male spatial ability, especially in tasks that involve mentally rotation of figures (Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995).

3.2.2. Additional items of self-efficacy in the CTt

At the end of the CTt, two additional items of self-efficacy are included. Both are Likert-type items in where the subject reports his/her perception according an 11-point scale (*Awful=0; Excellent=10*). The first item aims to collect the subject's self-efficacy perception regarding his/her performance on the CTt (*"From 0 to 10, how do you think you did on the test?"*). The variable that stores the answers to this first item will be called *'CT-SE'* (*'Computational Thinking Self-Efficacy'*) along the next sections. The second item is focused on the subject's self-efficacy perception in terms of his/her general competence with computers (*"From 0 to 10, how do you think you get on with computers?"*), and its corresponding variable will be called *'ICT-SE'* (*'Information and*

Table 2
Distribution of the total sample ($N = 1251$) by gender, grade and age.

			Grade			Total
			5th & 6th	7th & 8th	9th & 10th	
			Age			
			10-12 y/o	12-14 y/o	14-16 y/o	
Gender	Boys	Count	78	450	202	730
		% of Total	6.2%	36.0%	16.1%	58.4%
	Girls	Count	98	285	138	521
		% of Total	7.8%	22.8%	11.0%	41.6%
Total		Count	176	735	340	1251
		% of Total	14.1%	58.8%	27.2%	100.0%

Table 3
Distribution of the subsample ($n = 99$) by gender, grade and age.

			Grade			Total
			5th & 6th	7th & 8th	9th & 10th	
			Age			
			10-12 y/o	12-14 y/o	14-16 y/o	
Gender	Boys	Count	22	15	12	49
		% of Total	22.2%	15.2%	12.1%	49.5%
	Girls	Count	24	13	13	50
		% of Total	24.2%	13.1%	13.1%	50.5%
Total		Count	46	28	25	99
		% of Total	46.5%	28.3%	25.3%	100.0%

[2] http://www.bebras.org/.
[3] http://www.drscratch.org/.

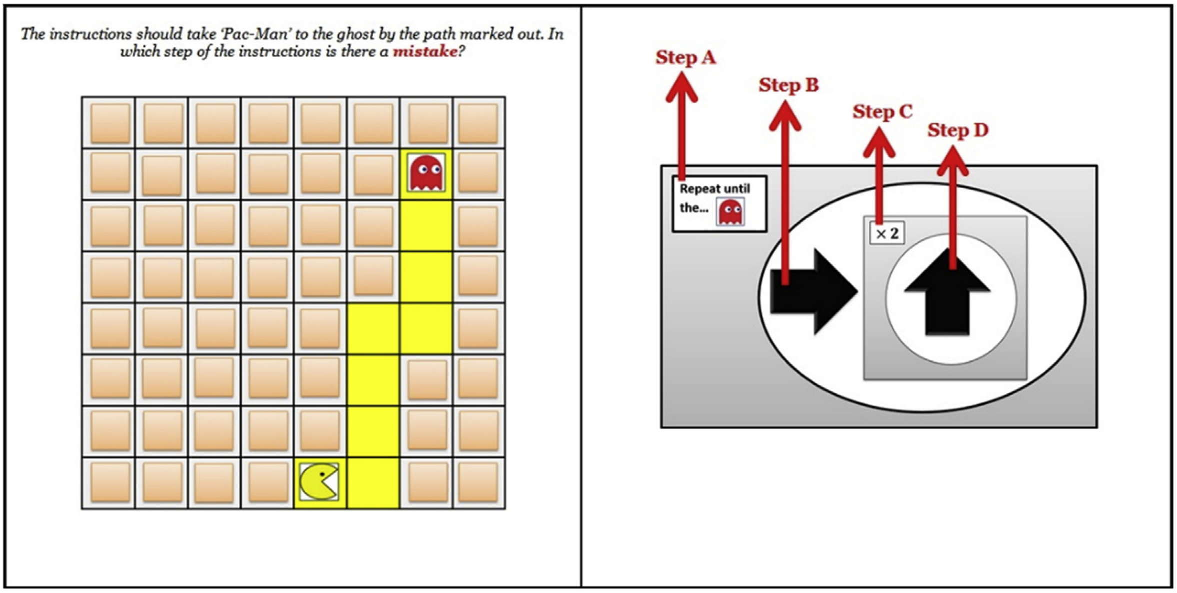

Fig. 1. CTt, item #11 ('maze'): loops 'repeat until + repeat times' (nested); 'visual arrows'; 'debugging'.

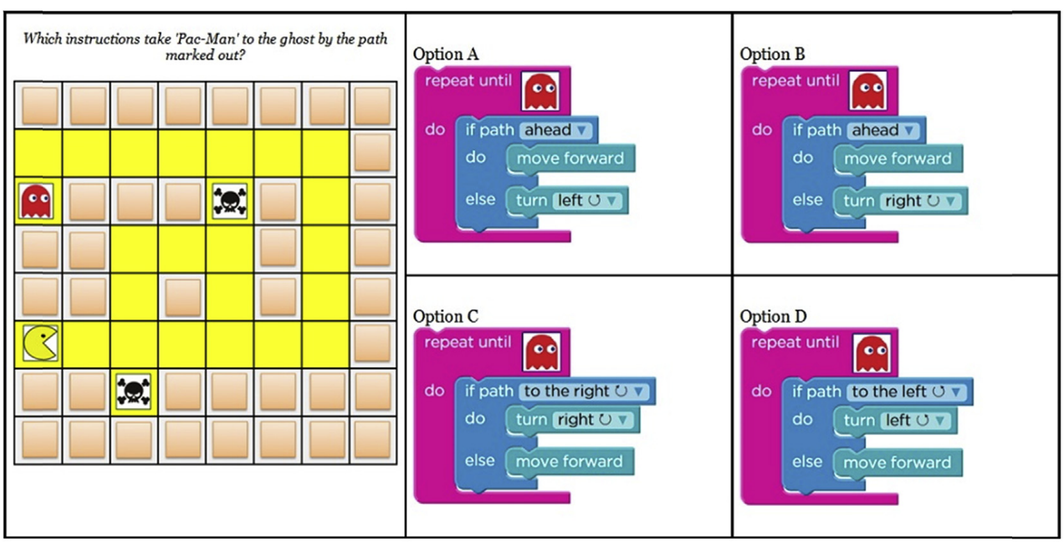

Fig. 2. CTt, item #18 ('maze'): loops 'repeat until' + if/else conditional (nested); 'visual blocks'; 'sequencing'.

Communication Technology Self-Efficacy').

As seen in subsection 2.2, given that self-efficacy is a psychological construct that is very specific to a certain domain, it is expected that the correlation between the CTt score and CT-SE will be stronger than the one between the CTt score and ICT-SE. In addition, higher levels of CT-SE and ICT-SE in the male group are also hypothesized.

3.2.3. Big Five Questionnaire-Children version (BFQ-C)

The Big Five Questionnaire-Children version (BFQ-C) (Barbaranelli et al., 2003) is an adaptation for child and adolescent population (8–15 years old) derived from the original 'adult' BFQ (Caprara, Barbaranelli, Borgogni, & Perugini, 1993), and it is aimed at assessing the personality of the subject within the Big Five model. The BFQ-C is a questionnaire without time limit and composed by 65 items; for each of them, the individual rates the occurrence of the behavior reported in the item using a 5-point Likert scale ranging from 1 (=*Almost never*) to 5 (=*Almost always*). Below, some item examples relative to each of the five personality factors are given:

- *Openness to Experience* (O): "I know many things"; "I have a great deal of fantasy"; "I easily learn what I study at school"; etc.
- *Conscientiousness* (C): "I work hard and with pleasure"; "I engage myself in the things I do"; "During class-time I am concentrated on the things I do"; etc.
- *Extraversion* (E): "I like to meet with other people"; "I like to compete with others"; "I like to move and to do a great deal of activity"; "I like to be with others"; etc.
- *Agreeableness* (A): "I share my things with other people"; "I behave correctly and honestly with others"; "I understand when others need my help"; "I like to give gifts"; etc.

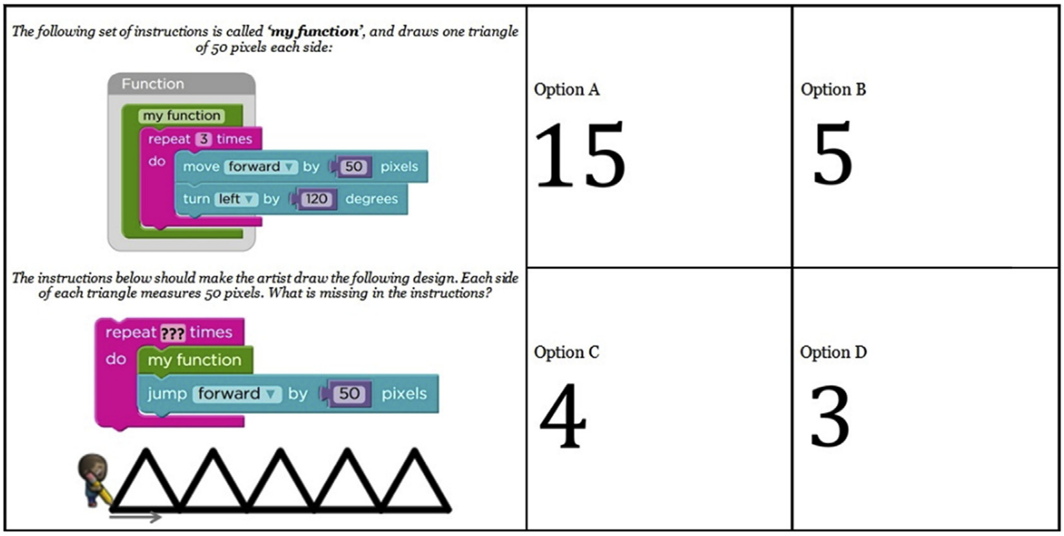

Fig. 3. CTt, item #26 ('canvas'): loops 'repeat times' + simple functions (nested); 'visual blocks'; 'completing'.

- *Neuroticism* (N): "I get nervous for silly things"; "I am in a bad mood"; "I argue with others with excitement"; "I easily get angry"; "I easily get offended"; etc.

In our research, the Spanish version (Barbaranelli et al., 2006) of the BFQ-C was administered as a 'self-report' form (students answer the items referring to themselves). The technical manual reports good reliability for all the factors ($\alpha > 0.80$), and statistically significant positive correlations between *Openness* (O) and Academic Achievement ($r = 0.51$), and between *Conscientiousness* (C) and Academic Achievement ($r = 0.13$) (Barbaranelli et al., 2006).

3.3. Procedure

Participants in our research attended the elective subject of Computer Science (in Secondary School), or received optional Computer Science contents (in Primary School), with a frequency of twice a week. Typically, the CTt (including the additional self-efficacy items) was administered during the first of the two weekly classes, and the BFQ-C during the second weekly class. None of the subjects had prior computer programming formal experience when the CTt was administered.

For the CTt collective administration, the Computer Science teacher followed the instructions that were sent by email the week before, containing the URL to access the on-line test. The student's direct answers to the CTt items were stored in the Google Drive database linked with the instrument, which was subsequently downloaded as an Excel *.xls* file.

For the collective administration of the BFQ-C, students were previously registered in the on-line platform from the publishing house[4] holder of this questionnaire's commercial rights. On the day of the questionnaire, subjects had to login into the platform and perform the questionnaire. Afterwards, from our administrator profile, we could download the subjects' results as an Excel *.xls* file.

Finally, all *.xls* files generated during data collection were exported to a single *.sav* file, which constitutes the data matrix under analysis with the SPSS software (version 24). Results shown below arise from this analysis.

[4] https://www.teaediciones.net/portal/e-teaediciones/default.aspx.

4. Results and discussion

4.1. Correlations between computational thinking and self-efficacy

Firstly, the descriptive statistics relative to the self-efficacy items are shown in Table 4, and the corresponding histograms are depicted in Fig. 4. As it can be seen, both variables range along the 11-points of the Likert scale and display a negative skewness, i.e., none of them follow a normal distribution ($Z_{\text{k-s (CT-SE)}} = 0.183$; $Z_{\text{k-s (ICT-SE)}} = 0.164$; $p < 0.001$). Additionally, the mean of ICT-SE is significantly higher than the CT-SE one (using the Wilcoxon signed-rank non-parametric test for large and paired samples, $Z_{\text{Wilcoxon}} = -6.76$; $p < 0.001$), but just with a small effect size ($d = 0.16$).

When the sample is split by grade, we find that in both variables the levels of self-efficacy decline as we move to upper grades (Table 5). These differences are statistically significant also in both items (using the Kruskal-Wallis non-parametric test for $K_{(=3)}$ large and independent samples, $\chi^2_{\text{(CT-SE, df=2)}} = 14.27$; $\chi^2_{\text{(ICT-SE, df=2)}} = 43.61$; $p < 0.001$). This finding could be somehow counter-intuitive, since the CTt score increases along the grades (Table 5). Nevertheless, there is a plausible explanation related to the highest accuracy and consistency of the answers coming from older students (Anastasi, 1968), which will be clearer when we study the correlations between CT and self-efficacy in the paragraphs below.

Table 6 shows the non-parametric correlations (Spearman's r) between the CTt score, CT-SE and ICT-SE for the whole sample. As it was expected, a positive and moderately intense correlation is found between the CTt score and the specific CT-SE variable ($r_s = 0.41$); and this value is notably higher than the one found between the CTt score and the general ICT-SE variable ($r_s = 0.16$). Furthermore, the correlation between both self-efficacy items is just moderate ($r_s = 0.51$), showing that CT-SE and ICT-SE are relatively independent (i.e., approximately only a quarter of the variance of the one is explained by the other). All correlations are statistically significant ($p < 0.01$).

When the sample is split by grade again, we find that the correlation values between the CTt score and the self-efficacy variables increase with the grade (Table 7), especially in the case of CT-SE (Fig. 5). In other words, the self-efficacy perceptions of the older students become more consistent with respect to their real performance on the test. As the older students become more aware

Table 4
Descriptive statistics relative to the self-efficacy items ($N = 1251$).

		"From 0 to 10, how do you think you did on the test?"	*"From 0 to 10, how do you think you get on with computers?"*
		Computational Thinking Self-Efficacy (CT-SE)	Information and Communication Technology Self-Efficacy (ICT-SE)
N	Valid	1237	1232
	Missing	14	19
Mean		7.07	7.38
Std. Error of Mean		.052	.056
Median		7.00	8.00
Mode		7	8
Std. Deviation		1.812	1.955
Variance		3.283	3.824
Skewness		−1.034	−1.091
Kurtosis		2.097	1.639
Range		10	10
Minimum		0	0
Maximum		10	10
Percentiles	10	5.00	5.00
	20	6.00	6.00
	25	6.00	6.00
	30	6.00	7.00
	40	7.00	7.00
	50	7.00	8.00
	60	8.00	8.00
	70	8.00	9.00
	75	8.00	9.00
	80	8.00	9.00
	90	9.00	10.00

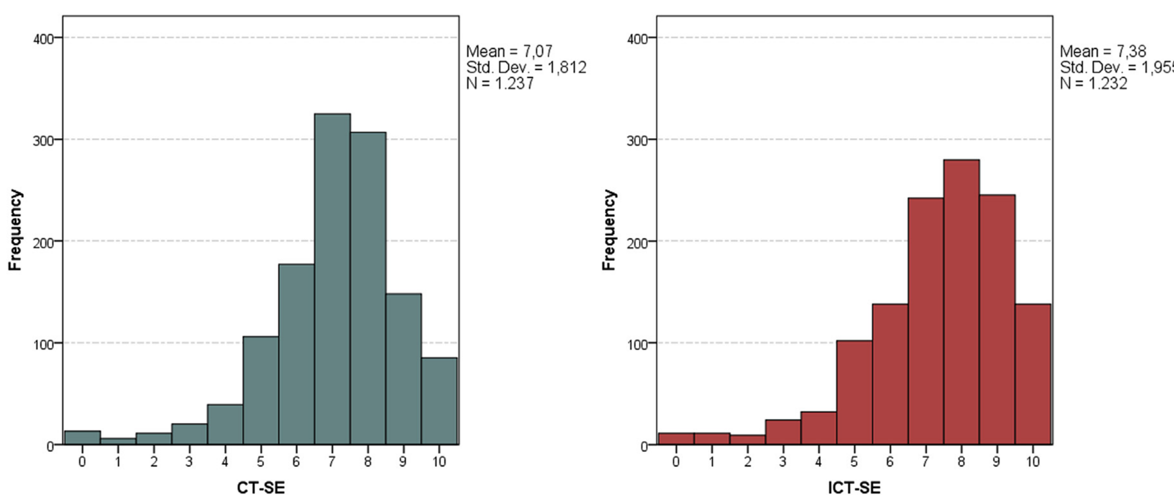

Fig. 4. Histograms relative to CT-SE (left) and ICT-SE (right).

about the accuracy of their answers in the CTt, it also occurs that the reliability of the test, measured as internal consistency (Cronbach's α), increases when we move to the upper grades (Table 7). All these related effects are depicted in Fig. 6.

Furthermore, when the whole sample is split by gender, statistical differences in favor of males are found both in CT-SE and in ICT-SE, and both using parametric and non-parametric tests (Table 8). As it can be seen, the effect size of the gender difference is greater in CT-SE ($d = 0.42$) than in ICT-SE ($d = 0.24$). In other words, it seems that the self-efficacy gender gap is bigger when it is referred specifically to computing tasks that when it does to

Table 5
Mean, median, and standard deviation of CT-SE, ICT-SE and CTt score, split by grades.

		Grade		
		5th & 6th (n = 176)	7th & 8th (n = 735)	9th & 10th (n = 340)
CT-SE	Mean	7.46	7.11	6.78
	Median	8.00	7.00	7.00
	Std. Deviation	1.589	1.646	2.182
ICT-SE	Mean	8.08	7.37	7.03
	Median	9.00	8.00	7.00
	Std. Deviation	1.855	1.867	2.097
CTt Score	Mean	13.76	16.24	18.05
	Median	14.00	16.00	18.00
	Std. Deviation	4.330	4.519	5.049

Table 6
Non-parametric correlations (Spearman's r) between CTt score, CT-SE and ICT-SE ($N = 1251$).

	CT-SE	ICT-SE
CTt score	.407**	.159**
CT-SE		.510**

**p-value < 0.01.

general digital competence. Moreover, the effect size of the gender difference in CT-SE is even bigger than the one found in the CTt score ($d = 0.31$), which was already reported in our previous paper (Román-González, Pérez-González, et al., 2017). That is, an extra subjective CT gender gap is added to the objective difference found in the performance on the CTt. This additional effect might inhibit the subsequent learning progress of the female group.

Finally, when we analyze the aforementioned gender differences along the grades, we find that the gender gap in the CTt score grows as we move to the upper grades, and that occurs in parallel with the sharp decline in self-efficacy that is shown by the oldest girls (Fig. 7). Given that our approach is correlational, we cannot state any causality of the CTt performance on the CT-SE, or vice versa. However, the joint covariance of both variables along grades and genders seems to be clear. Thus, one of the plausible ways of improving the CT of our students could be to foster their self-efficacy in specific computing tasks, and that might be critical with adolescent girls. We will return to this issue in section 5, relative to the implications and limitations of our research.

When it comes to the factors that may cause these differences based on gender, we have identified in the literature three aspects that could be present in our investigation. On the one hand, girls tend to prefer and yield higher in open-ended, storytelling activities instead of closed, guided ones, when learning to program; and that occurs specially from middle-school ages onwards (Howland & Good, 2015). These authors report significantly higher values in the computational complexity of scripts written by girls from 7th and 8th grade in comparison with their male peers within narrative tasks; this result is consistent with the (slight) female superiority in tasks involving verbal ability reported in the classical literature (Hyde & Linn, 1988). Also, in an investigation comparing girls' experiences learning to program using Storytelling Alice and a version of Alice without storytelling support (Generic Alice), participants who used Storytelling Alice showed more evidence of engagement with programming and expressed greater interest in future use of coding than participants who used Generic Alice (Kelleher, Pausch, & Kiesler, 2007).

On the other hand, it is common that male students consider time, or coding speed, as the main measure of ability in programming courses. In this sense, as proved in an investigation studying women in the first systems course in Stanford's CS, “perceptions of coding speed are biased by variations in accepted norms of self-report and self-display” (Irani, 2004). Thus, while men admit that they say ‘I'm done’ even when the task is not complete, female students tend to submit the exercises only when programs meet all requirements in the specifications. Consequently, women get frustrated with their coding speed, which has an impact on their level of self-efficacy.

Lastly, gender inequality and stereotypes in programming materials also have an impact on girls' self-efficacy perception. In an evaluation conducted at California State University, researchers compared the results obtained in a security class using two different materials: classic cryptography protocol characters (Bob, Alice, Eve…) and a gender equitable alternative using animal characters and the singular ‘they’ instead of the generic ‘he’ (Medel & Pournaghshband, 2017). Results indicate an improvement for female students' confidence in understanding the material, while no difference was detected for male students.

In consequence, the kind of programming items included in the CTt (visuospatial and closed-ended items), may have contributed to a lack of engagement of girls. A different set of questions, following an open-ended approach and including storytelling activities might show different results. In addition, the CTt is a maximum performance test (i.e., it must be completed in a limited time). This fact might encourage a biased coding speed competition with an impact on girls' self-efficacy perception. Finally, a different version of the CTt in which a special focus on the representation, imagery and language used in the questions was placed, might also increase girls' interest and confidence, which could therefore enhance their test score.

4.2. Correlations between computational thinking and personality

Table 9 shows the correlations between the CTt score and the five personality factors assessed through the BFQ-C. Given that both are interval measures and at least quasi-normally distributed, parametric correlations (Pearson's r) are calculated.

The results show that the CTt has a positive statistically significant correlation ($p < 0.01$) with three of the five personality factors assessed through BFQ-C: moderately intense with the *Openness* (O) factor, and slightly intense with the *Extraversion* (E) and *Conscientiousness* (C) factors. There is no statistically significant correlation between CTt and the *Agreeableness* (A) and *Neuroticism* (N) factors. Overall, these results are partially consistent with the literature regarding the links between cognitive and personality variables, as it will be discussed in the next paragraphs. Furthermore, our results fit notably with our expectations after intersecting the dispositions or attitudes that underlie CT (CSTA & ISTE, 2011) with the five personality dimensions stated from the Big Five model (Barbaranelli et al., 2003; Goldberg, 1990) (Table 1).

On the one hand, the positive and moderately intense correlation found between CT and the *Openness* (O) factor is consistent with the prior research, which states that the ‘O’ factor is the one that has the strongest relation with cognitive variables such as intelligence (Chamorro-Premuzic & Furnham, 2014; Higgins et al., 2007; John et al., 1994; Von Stumm & Ackerman, 2013) or academic performance (Barbaranelli et al., 2003; Mervielde et al., 1995; Poropat, 2014). Finding this correlation value was also expected as there are several attitudes underlying CT (CSTA & ISTE, 2011) linked with *Openness* (e.g., “The ability to deal with open ended problems”). Moreover, the positive but in a lesser extent slight correlation between CT and the *Conscientiousness* (C) factor is consistent with the body of research, which links lightly the ‘C’ factor with the aforementioned cognitive variables (Abe, 2005; Chamorro-Premuzic & Furnham, 2014). This was expected too, as there are also some dispositions supporting CT related with *Conscientiousness* (e.g., “Persistence in working with difficult problems”).

On the other hand, and following an analogous discussion thread, the absence of correlation between CT and the

Table 7
Non-parametric correlations (Spearman's r) CTt score * SE, and reliability of the CTt, split by grade.

Grade	n		Spearman's r		Reliability of the CTt (Cronbach's α)
			CT-SE	ICT-SE	
5th & 6th	176	CTt Score	.387**	.187*	0.721
7th & 8th	735		.418**	.186**	0.762
9th & 10th	340		.557**	.297**	0.824

**p-value < 0.01; *p-value < 0.05.

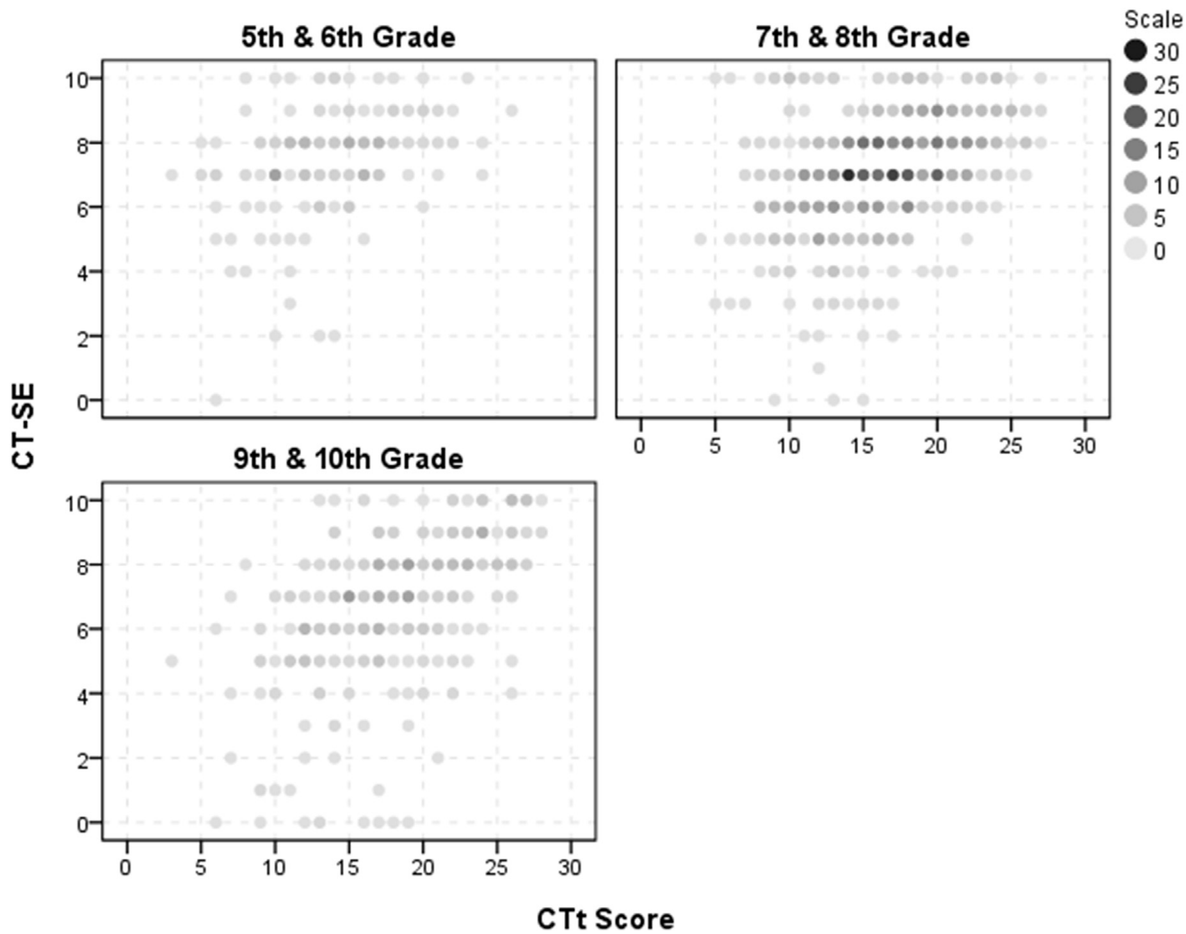

Fig. 5. Scatterplots CTt score * CT-SE, by grade.

Agreeableness (A) and *Neuroticism* (N) factors is also consistent with literature visited in subsection 2.3. This result was also expected according to Table 1, as there are no attitudes enhancing CT related to these factors (except "The ability to communicate and work with others to achieve a common goal or solution", which can be just partly linked with the 'A' factor).

Finally, and this is the point which requires deeper discussion, the positive and slight (but considerable $r = 0.30$) correlation between CT and the *Extraversion* (E) factor seems surprising. Although the 'E' factor intersects with some CT attitudes or dispositions from the CSTA & ISTE operational definition (Table 1), the traditional body of research does not show any positive correlation between the 'E' and cognitive variables (as CT is mainly supposed to be). Actually, some evidence of slight negative correlation exists in the classic literature between the 'E' factor and intelligence (Austin et al., 2002) or academic performance ($r = -0.13$) (Barbaranelli et al., 2003). Just recently, coinciding with the appearance of the social learning environments, some evidence of positive correlations specifically in middle school students between 'E' and AP can be found (Poropat, 2011; Zuffianò et al., 2013).

Following this discussion argument, we also wonder if the *Extraversion***Intelligence* correlation might be development-dependent, since it has been observed that relationship between personality and intelligence change from younger to older adulthood (Baker & Bichsel, 2006). The conclusions of a meta-analysis carried out by Ackerman and Heggestad (1997) state that although *Extraversion* and intelligence are, positively and significantly, related in a weak way ($r = 0.08$), this correlation may be larger in samples with younger subjects, reaching $r = 0.20$. Furthermore, the two factors of the Big Five showing marked changes as a result of intervention are *Neuroticism* and *Extraversion*, according to the meta-analysis by Roberts et al. (2017). This finding highlights the malleability of traits such as *Extraversion*, which probably does not remain unchanged across major developmental stages and across different learning environments, being more salient in childhood than in adulthood, and more prominent within social learning contexts than within isolated ones. All of the above might produce the aforementioned inconsistencies concerning its relationship with cognitive variables.

The previous discussion line leads us to speculate that *Extraversion* (also called *Energy*) might be a specific personality trait of the present top computational thinkers in middle school, which are being educated in the age of social platforms and media. Or maybe the opposite, *Introversion* might be characteristic of actual low computational thinkers along this range of ages. Subsequently, the personality profiles of top and low computational thinkers are studied in the next subsection.

4.3. *Personality profiles of top and low computational thinkers*

In order to build the personality profiles of top and low computational thinkers, we standardize the CTt score and the BFQ-C scores of each individual with respect to his/her reference group by grade. Then, each subject is categorized in one of the three following classes: 'Low CT Thinkers' (CTt score < $-1SD$), 'Medium CT Thinkers' ($-1SD$ < CTt score < $+1SD$), or 'Top CT Thinkers' (CTt

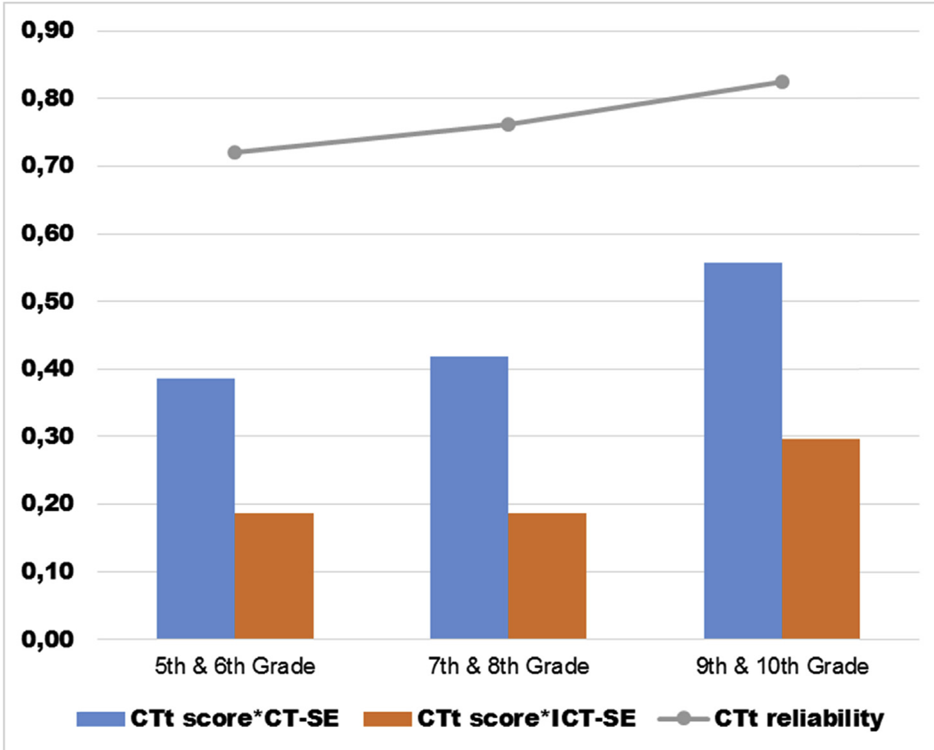

Fig. 6. Related effects between self-efficacy and reliability along the grades.

score > +1*SD*). The distribution of the subsample ($n = 99$) along these three CT categories and by gender is shown in Table 10. A statistical significant association between gender and the CT categories is found ($C = 0.286$; $\chi^2_{(df=2)} = 8.788$; $p < 0.05$).

Fig. 8 shows the personality profiles of low, medium and top CT thinkers. As it can be seen, top CT thinkers display high standardized means ($z > 0.50$) in *Openness* and *Conscientiousness*, while low CT thinkers show a low standardized mean ($z < -0.50$) in *Extraversion*. In other words, the specific relation between CT and the 'E' factor seems to be not specially based on the extraversion of the top CT thinkers, but in the introversion of the low ones. The ANOVA (Table 11) shows statistically significant differences among the CT categories only in *Openness* ($F_{(2, 96)} = 6.429$; $p < 0.01$) and *Extraversion* ($F_{(2, 96)} = 3.442$; $p < 0.05$). Finally, using the Bonferroni's post-hoc test, statistically differences (at least $p < 0.05$) are found between the means of the couple 'top-medium' in *Openness*, and between the couple 'top-low' in *Openness* and in *Extraversion*.

These findings contradict in some way the traditional body of research supporting introversion as a characteristic personality trait of the top computer programmers. Thus, introversion of software developers has been studied since the 1980s. Most of the research articles before the year 2000 state that programmers tend to introversion, as excellent programmers do not like social interaction (Schott & Selwyn, 2000). In this regard, some studies provide scientific evidence that shows programmers to be more introverted than the general population (Ketler & Smith, 1992), that extroverts have lower grades in computing exams (Kagan & Esquerra, 1984), or that most programmers are identified as introverted, serious, quiet and logical (Bush & Schkade, 1985). Other studies from those years have however not found a relation between introversion and computer programming skills (Kagan & Douthat, 1985; Newsted, 1975). Summarizing this first wave of research, Gnambs (2015) performed a meta-analysis, where he studied 19 samples from the last 40 years. He reported that "the most important personality predictor [of programmers] was introversion. Introverts are reserved individuals with low levels of sociability; they tend to focus on their inner self instead of their social surrounding" (Gnambs, 2015, p. 34).

However, in the last years we have witnessed many studies in the opposite direction, i.e., showing the superiority of extroverts in programming tasks. The superiority of extrovert students while e-learning Java (Law, Lee, & Yu, 2010), the increasing number of extrovert individuals in industrial software teams (Yilmaz & O'Connor, 2012), the better performance of those developers who

Table 8
Differences by gender in CT-SE, ICT-SE and CTt score.

	Gender	N	Mean	Std. Deviation	Student's *t*	$Z_{\text{Mann-Whitney } U}$	Effect size Cohen's *d*
CT-SE	Boys	723	7.38	1.787	7.261**	−8.095**	0.42
	Girls	514	6.64	1.759			
ICT-SE	Boys	723	7.57	1.968	4.136**	−4.971**	0.24
	Girls	509	7.10	1.906			
CTt Score	Boys	730	16.99	4.802	5.374**	−5.090**	0.31
	Girls	521	15.52	4.727			

**p-value < 0.01.

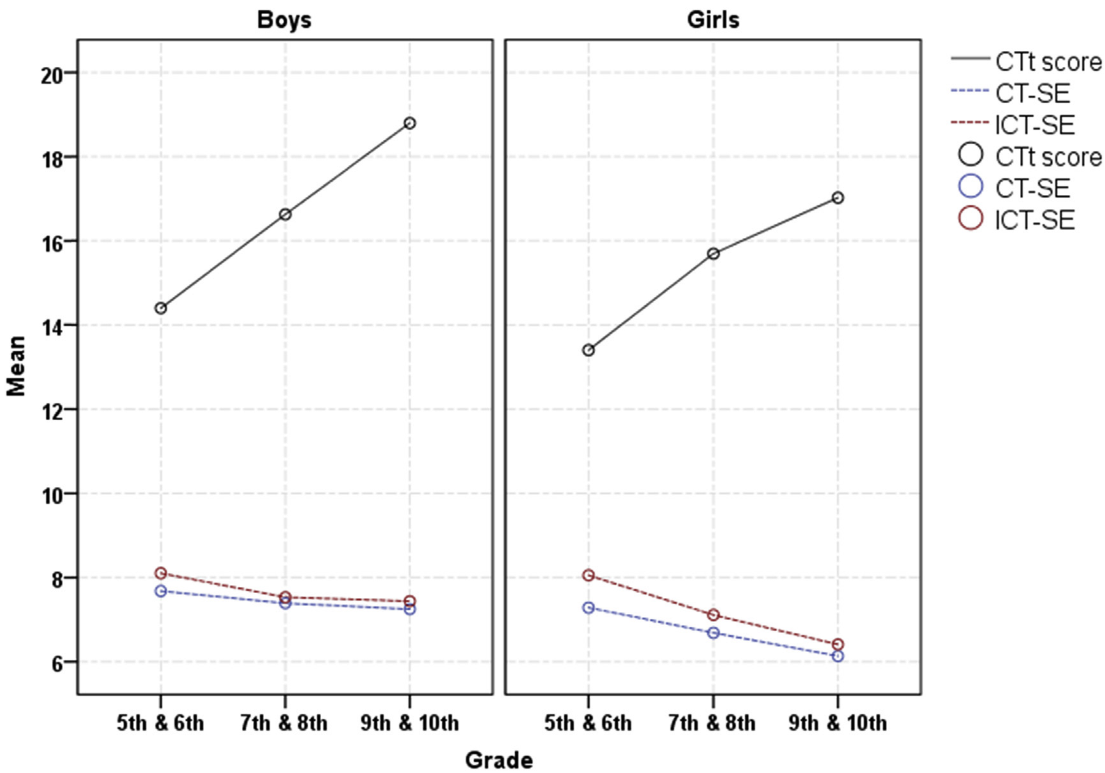

Fig. 7. Evolution of CTt score, CT-SE and ICT-SE by gender and grade.

demonstrated more openness in global software development environments (Licorish & MacDonell, 2014), or the identification of highly extroverted developers in usability-related tasks (Licorish & MacDonell, 2014) have been reported. In addition, a study of the popular StackOverflow Q&A website showed that more prominent participants (i.e., those with a higher "karma" or reputation) are more extroverted than those with medium or lower reputation (Bazelli, Hindle, & Stroulia, 2013). Hence, the extroversion trait is becoming much more predominant in software developers than previously suggested in the literature (Yilmaz, O'Connor, Colomo-Palacios, & Clarke, 2017).

The change in personality traits has also to be situated in the moving technological context. In this sense, although open source software has always put special emphasis on community and sharing (Lakhani & Von Hippel, 2003), the appearance of so-called social software development platforms, such as GitHub or GitLab, has fostered this new trend (Dabbish, Stuart, Tsay, & Herbsleb, 2012). Thus, in some way, developing software has become an inherently social task, which seems to favor extroverts.

In a similar vein, but now descending from adult software development to young coders, traditionally programming has been learned in programming languages and environments that better fit introverted learners, with difficult syntaxes and isolated development contexts for children and teenagers (e.g., Logo or BASIC). In contrast, the new environments, such as the Scratch platform, allow learners to easily take projects from others ('remix' in the Scratch jargon), and adapt and enhance them (Resnick et al., 2009). Evidence from the Scratch platform exist that learners who perform social actions are able to create more complex projects than those who do not (Moreno-León, Robles, & Román-González, 2016), being this an indicator of the higher CT of the most social 'scratchers'. We might be witnessing a transformation of the young and adult programmer from a 'logical-formal', 'inside-oriented' to an 'expressive-communicative', 'outside-oriented' subject.

A similar academic discussion on the extroversion of young programmers can be found in the research of gamers. Until a decade ago, the involvement in gaming activities was thought as a compensation of inefficiencies in real life, for instance, loneliness, introversion, poor social skills and lack of relationship abilities (Caplan, 2005; Charlton & Danforth, 2007). However, a more recent study of gamers in Massively Multiplayer Online Role-playing games (MMORPGs), such as the popular World of Warcraft, has started to question the myth of the socio-emotionally dissatisfied gamer, revealing, "that gamers are avidly social individuals"

Table 9
Correlations (Pearson's r) between the CTt and BFQ-C ($n = 99$).

	BFQ-C				
	Openness (O)	*Conscientiousness* (C)	*Extraversion* (E)	*Agreeableness* (A)	*Neuroticism* (N)
CTt	.407**	.267**	.304**	.133	.092

**p-value < 0.01.

Table 10
Distribution of the subsample ($n = 99$) in low, medium and top CT thinkers.

		Gender		Total
		Boys	Girls	
Low CT Thinkers	Count	4	11	15
	% of Total	4.0%	11.1%	15.2%
Medium CT Thinkers	Count	33	36	69
	% of Total	33.3%	36.4%	69.7%
Top CT Thinkers	Count	12	3	15
	% of Total	12.1%	3.0%	15.2%
Total	Count	49	50	99
	% of Total	49.5%	50.5%	100.0%

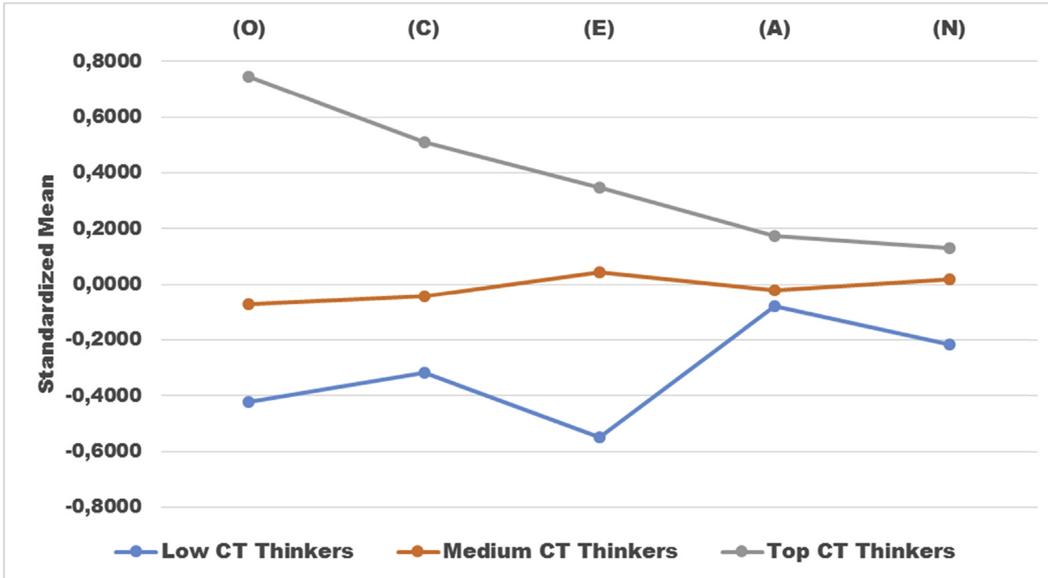

Fig. 8. Personality profiles of low, medium and top CT thinkers.

Table 11
ANOVA of the personality factors along the CT categories (‘low’, ‘medium’ and ‘top’).

		N	Standardized Mean	Std. Deviation	95% Confidence Interval for Mean		*F*
					Lower Bound	Upper Bound	
(O)	Low CT	15	−0.421	1.086	−1.023	0.180	6.429**
	Medium CT	69	−0.070	0.921	−0.292	0.151	
	Top CT	15	0.745	0.865	0.266	1.224	
(C)	Low CT	15	−0.317	1.050	−0.899	0.264	2.924
	Medium CT	69	−0.042	0.988	−0.279	0.196	
	Top CT	15	0.509	0.785	0.074	0.944	
(E)	Low CT	15	−0.548	1.134	−1.176	0.081	3.442*
	Medium CT	69	0.044	0.970	−0.189	0.277	
	Top CT	15	0.346	0.738	−0.063	0.754	
(A)	Low CT	15	−0.080	1.373	−0.840	0.680	0.295
	Medium CT	69	−0.021	0.888	−0.234	0.193	
	Top CT	15	0.175	1.050	−0.406	0.757	
(N)	Low CT	15	−0.215	1.240	−0.902	0.471	0.494
	Medium CT	69	0.018	0.954	−0.211	0.247	
	Top CT	15	0.132	0.908	−0.371	0.635	

**p-value < 0.01; *p-value < 0.05.

(Herodotou, Kambouri, & Winters, 2014, p. 23). In other words, as the coding and gaming environments are becoming on-line and collaborative, an emerging personality trait of extraversion is emerging both in coders and in gamers.

In summary, when we try to explain the unexpected correlation between CT and *Extraversion* in our middle school subjects, two types of plausible explanations are found. First, from a development approach, middle schoolers are more likely to be extrovert than their analogous adult peers. Second, from an environmental approach, the new collaborative e-learning tools are favoring the ways of (computational) thinking of extroverts. Thus, updating the terms of Vygotsky (1978), the external digital-world and the internal psychological-world are influencing each other, and conforming this new human-machine hybrid reality.

Finally, we find two more types of additional evidence that support that *Extraversion* might be a specific personality trait of top computational thinkers:

- There are some emerging and comprehensive CT assessment frameworks that take into account leadership and collaboration skills (Snow, 2014), or assertiveness and effective communication skills (Grover, 2015) (which are at the core of ‘E’ factor), as important ingredients of CT.
- We have some recent qualitative studies (Román-González, 2016), in which teachers report the unexpected brilliant performance and behavior of students usually disruptive and inattentive, when faced with computer programming experiences (e.g., Code.org courses[5]). In other words, students with some externalizing problems (hyperactivity, disruptive conduct, inattentiveness), linked with high values in the ‘E’ factor (Barbaranelli et al., 2003), seem to respond especially well to programming tasks.

[5] https://studio.code.org/s/20-hour.

4.4. Regression model of computational thinking based on personality

In order to answer our last research question (i.e., how much variance of CT can be explained by personality?), we perform a multiple linear regression over the CTt score (considered as the dependent variable) based on the BFQ-C scores (considered as predictors). Table 12 summarizes the regression model, which is calculated through the 'enter' method. This regression model, based on the BFQ-C, correlates $r = 0.529$ with the CTt, which means an adjusted $R^2 = 0.24$. That is, 24.0% of the CTt scores' variance is explained from a linear combination of the Big Five personality factors measured through the BFQ-C. Normality of the regression model residuals was verified.

The regression model is able to explain, statistically significant, the differences in the CTt scores, as $F_{(5,\ 93)} = 7.216$ ($p < 0.01$). However, as shown in Table 13, which contains the coefficients of the regression model, only *Openness* and *Extraversion* are capable, specifically and statistically significant ($p < 0.01$), to explain differences in the dependent variable (CTt). Fig. 9 shows the scatterplot between the observed CTt scores and the corresponding predicted values according to the regression model.

Table 12
Summary of the regression model of the CTt onto the BFQ-C factors.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.529[a]	.280	.241	3.663

[a] Predictors: (Constant), Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism.

4.5. Extending the nomological network of computational thinking with non-cognitive factors

In our previous paper (Román-González, Pérez-González, et al., 2017), we found that 27% of the CT variance was explained by the primary mental abilities. Our new findings, from a personality perspective, show that the Big Five factors do so with 24%. When we consider both studies together, strong evidence about the importance of taking into account both cognitive and non-cognitive factors in the explanation and enhancement of the CT of our students appears.

Merging the results of our previous paper (Román-González, Pérez-González, et al., 2017) with the present work, the extended nomological network of CT can be depicted (Fig. 10).

5. Implications and limitations

Given that skills and self-beliefs are so intertwined, one way of improving the performance of a subject is to improve his/her self-efficacy. This statement implies that one of the plausible ways of improving the CT skills of our students could be to foster their self-efficacy in specific computing tasks, and that might be critical for girls in their early adolescence.

Then, the most effective strategy to enhance the self-efficacy perception of an individual is to provide him/her with positive and personal learning experiences in the corresponding domain (Bandura, 1986). Subsequently, if we want to guarantee an equitable computer science education, it is essential to offer students who start learning to program a wide range of coding and CT development environments, not limited and biased ones. Thus, we may maximize the chance that every student, both male and female, could find a significant and self-reinforcing computing experience.

Since the appearance of the new generation of visual programming languages, some of the 'learning to code' environments have been skewed to 'close-ended' and visuospatial problems (e.g., Code.org). We consider that after this rapid inrush, the time has come to diversify the environments in which students learn to program, and to enlarge the types of problems on which they project their CT. In this sense, there is already evidence about the use of CT for solving problems with different features in middle and high school, such as: modeling scientific simulations (Weintrop et al., 2016); algorithmic composition of computational music (Aaron & Blackwell, 2013) and choreographies (Daily, Leonard, Jörg, Babu, & Gundersen, 2014); or digital interactive storytelling (Burke, 2012; Howland & Good, 2015). The latter seems to be specially engaging for teenage girls, given their verbal load, as well as recent experiences where adolescent females showed deeper learning of computational concepts through the programming of conversational 'chatbots' (Benotti, Martinez, & Schapachnik, 2017; Benotti, Martínez, & Schapachnik, 2014). In addition, it is worth noting that, although early adolescence seems to be the critical moment for the gender gap in CT self-efficacy and performance, recent studies show that providing positive programming experiences promotes higher technology interest and self-efficacy even in 1st grade girls (6 years-old) (Master, Cheryan, Moscatelli, & Meltzoff, 2017).

Another way to enhance the self-efficacy, and subsequently the performance, of our students may be the use of behavior modeling and verbal persuasion (Bandura, 1986). Accordingly, gender stereotypes about computer science and programming must be tackled to avoid their detrimental effect on the CT development of females. As noted in Cheryan, Plaut, Handron, and Hudson (2013, p. 58), stereotypes of software developers have an impact in the inclusion of women; in an experiment it was found that those women "who read that computer scientists no longer fit the stereotypes expressed more interest in computer science than those who read that computer scientists fit the stereotypes". Hence, more women could be interested in computer science if programmers are depicted more accurately in the media and in the schools.

Finally, there is an extra strategy stated by Bandura (1986) to enhance self-efficacy, which consists in implementing cooperative learning environments. Thus, it is more likely to reach higher levels of CT self-efficacy when the subject learns in a social and collaborative way. This fact has been recently pointed out by Brady et al. (2017), who argue that computer science is becoming an increasingly diverse domain, and that this growing diversity should be addressed through 'participatory an social computing'. Furthermore, this implication links with the malleability of the extraversion personality factor, as developing CT in collaborative and social contexts might push introverted female 'low CT thinkers' to higher levels of extraversion and, subsequently, to perform better in CT tasks.

Overall, our results give empirical support to the statement of Kafai and Burke (2013). These relevant authors argue that "recent developments in K-12 programming education are suggestive of what can be called a 'social turn' (…) learning to code has shifted from being a predominantly individualistic and tool-oriented approach to now one that is decidedly sociologically and culturally grounded in the creation and sharing of digital media" (Kafai & Burke, 2013, p. 603).

The main limitations of this work are related with the instruments used for the assessments. Thus, the CTt has itself some limitations that have been already described in paragraph 3.2.1. Moreover, the self-efficacy items were designed *ad-hoc* for our research. In consequence, different results might have been found if we had used a different CT assessment tool or an already validated measure for CT self-efficacy, such as the one by Kukul et al. (2017).

Table 13
Standardized coefficients[a] of the regression model of the CTt onto the BFQ-C personality factors.

Model		Unstandardized Coefficients		Standardized Coefficients	Student's *t*
		β	Std. Error	β	
1	(Constant)	−1.472	3.785		−.389
	Openness	.510	.136	.618	3.745**
	Conscientiousness	−.077	.072	−.220	−1.066
	Extraversion	.294	.096	.365	3.050**
	Agreeableness	−.180	.097	−.270	−1.861
	Neuroticism	.071	.057	.122	1.243

**p-value < 0.01.
[a] Dependent Variable: CTt score.

Finally, it must be noted the small size of the subsample of the CTt * BFQ-C analysis, being formed by just 99 students.

6. Conclusions and further research

In this paper, we have studied the correlations between CT and several non-cognitive variables, such as self-efficacy and personality. We have found a moderate positive correlation between CT and the specific self-efficacy perception of the students relative to their performance in CT tasks, in which a medium size difference in favor of males is found. We have also provided empirical evidence of the correlations between CT and the five factors of personality from the 'Big Five' model. We have found expected positive correlations with *Openness* and *Conscientiousness*, and unexpected positive correlation with *Extraversion*. Both results have led us to affirm that it is essential to assure a diversity of computing contexts so that every student, especially adolescent girls, can enjoy meaningful and self-reinforcing experiences. Moreover, our findings have led us to argue that extraversion might be an emerging and specific personality trait of present top computational thinkers in middle and high school, coinciding with the popularization of the social and collaborative coding platforms.

In addition, we have provided a regression model of CT build onto the personality factors of the Big Five model. We have found that 24% of CT can be explained through personality factors. This result complements the 27% of CT that was explained through the primary mental abilities in our previous work (Román-González, Pérez-González, et al., 2017). Overall, our findings corroborate the idea that, although CT is mainly a cognitive psychological construct close to problem-solving ability, there is also a complementary non-cognitive side of CT. Subsequently, both sides should be taken into account by educational policies and interventions aimed at fostering CT. As a final contribution, we depict the extended nomological network of CT integrating cognitive and non-cognitive

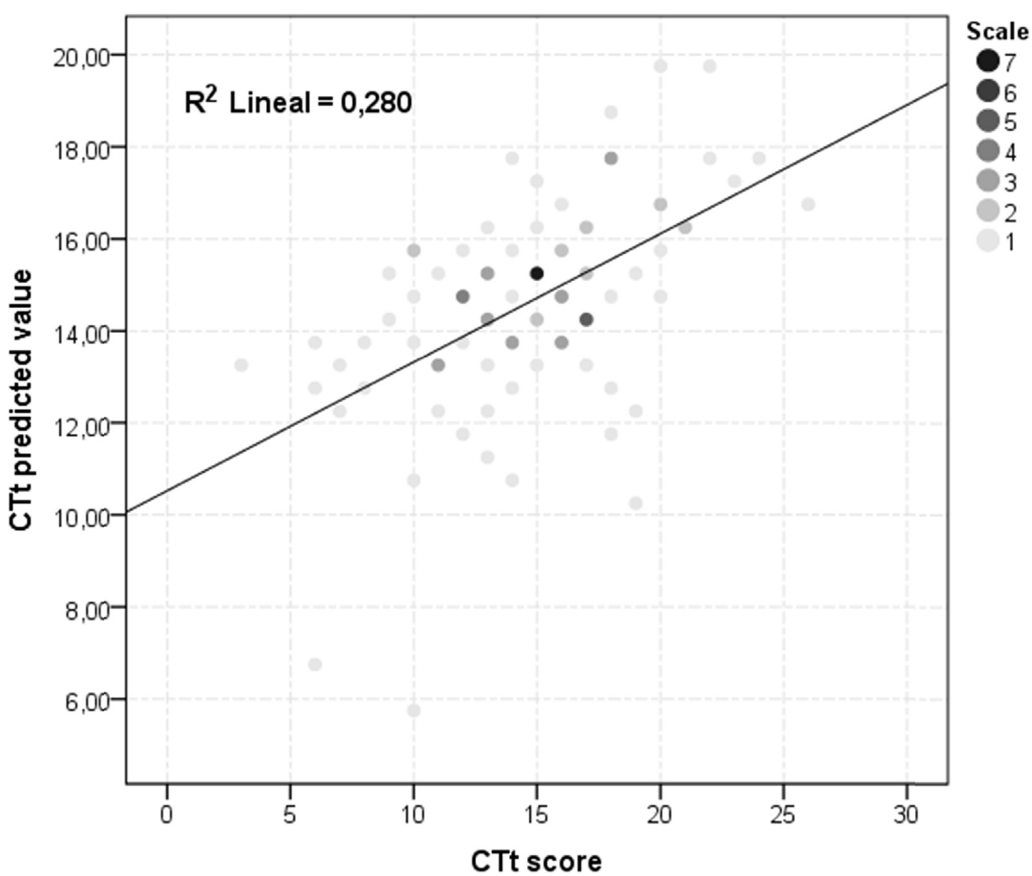

Fig. 9. Scatterplot between the observed CTt scores and the predicted values by the regression model.

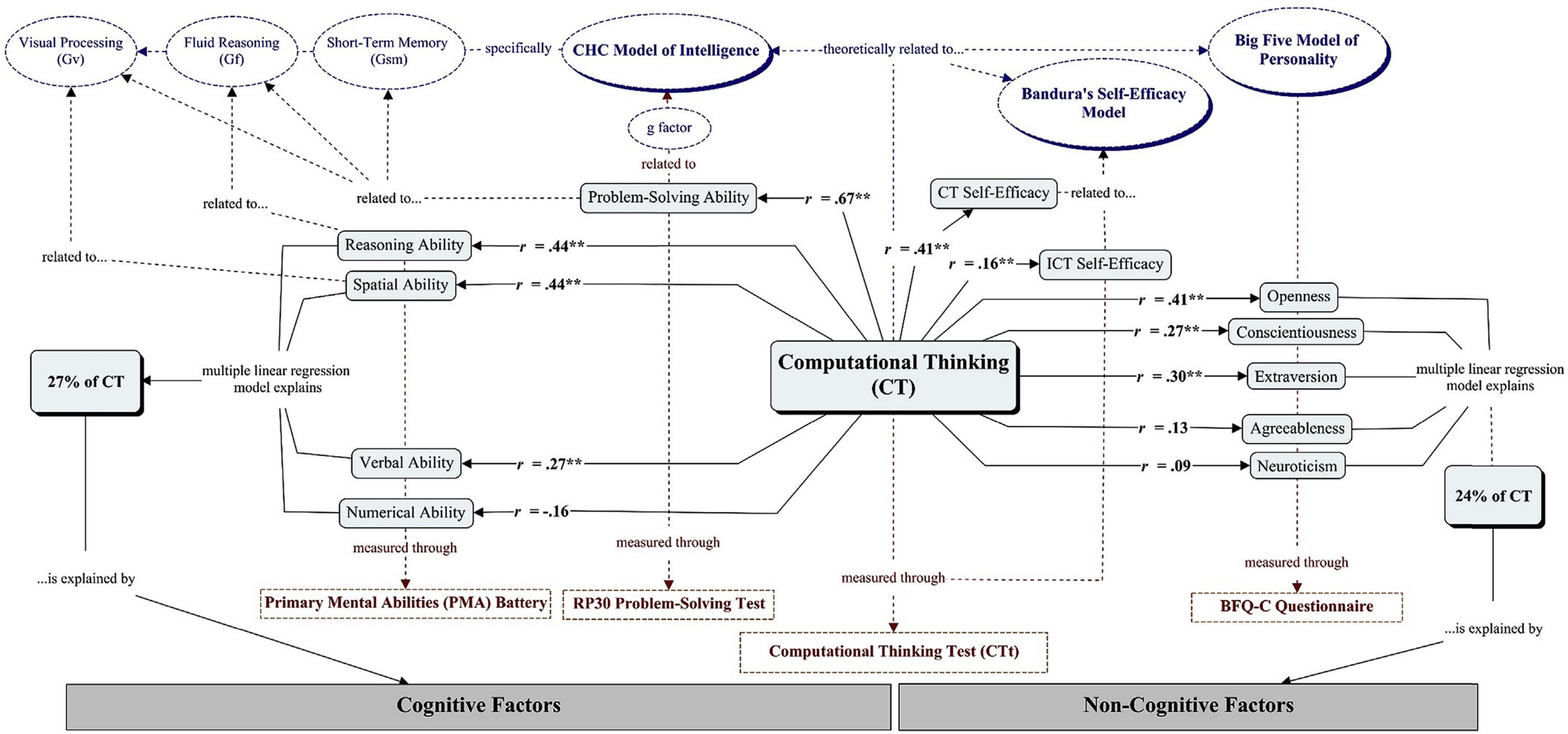

Fig. 10. Nomological network of CT including cognitive and non-cognitive factors.

variables (Fig. 10). We hope that this extension of the nomological network of CT will contribute to consolidate it as a solid psychological construct.

From this point, further research can be conducted in the following directions:

- Longitudinal studies aimed at studying if reinforcement of self-efficacy can actually improve CT performance.
- Analogously, longitudinal studies aimed at studying if reinforcement of behaviors related to *Openness*, *Conscientiousness*, and *Extraversion*, can actually foster CT learning and development in educational settings.
- Moreover, it is necessary to extend the repertoire of CT assessment tools, so the growing diversity of the field is reflected in the corresponding measurement instruments. Thus, we intend to enlarge the CTt to a 'multiple CT battery', in which CT will be assessed in different modalities (not only visuospatial, but also verbal, musical, kinesthetic, or ethical).
- Finally, case studies on the effect of computer programming tasks over students with externalizing problems (e.g., hyperactivity, disruptive conduct, inattentiveness) may be conducted.

Acknowledgements

The work of Jesús Moreno-León and Gregorio Robles has been funded in part by the Region of Madrid under project "eMadrid -Investigación y Desarrollo de tecnologías para el e-learning en la Comunidad de Madrid" (S2013/ICE-2715) and in part by the Spanish Government under project SobreVision (TIN2014-59400-R).

References

Aaron, S., & Blackwell, A. F. (2013). From sonic Pi to overtone: Creative musical experiences with domain-specific and functional languages. In *Proceedings of the first ACM SIGPLAN workshop on Functional art, music, modeling & design* (pp. 35–46). https://doi.org/10.1145/2505341.2505346.

Abe, J. A. A. (2005). The predictive validity of the Five-Factor Model of personality with preschool age children: A nine year follow-up study. *Journal of Research in Personality, 39*(4), 423–442. https://doi.org/10.1016/j.jrp.2004.05.002.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*(2), 219–245. https://doi.org/10.1037/0033-2909.121.2.219.

Ackerman, P. L., & Rolfhus, E. L. (1999). The locus of adult intelligence: Knowledge, abilities, and nonability traits. *Psychology and Aging, 14*(2), 314–330. https://doi.org/10.1037/0882-7974.14.2.314.

Aho, A. V. (2012). Computation and computational thinking. *The Computer Journal, 55*(7), 832–835. https://doi.org/10.1093/comjnl/bxs074.

Ambrósio, A. P., Xavier, C., & Georges, F. (2014). Digital ink for cognitive assessment of computational thinking. In *Frontiers in education conference (FIE), 2014 IEEE* (pp. 1–7). https://doi.org/10.1109/FIE.2014.7044237.

Anastasi, A. (1968). *Psychological testing* (3rd ed.). Oxford, England: Macmillan.

Armoni, M., Meerbaum-Salant, O., & Ben-Ari, M. (2015). From Scratch to "real" programming. *ACM Transactions on Computing Education (TOCE), 14*(4), 25: 1–25:15. https://doi.org/10.1145/2677087.

Askar, P., & Davenport, D. (2009). An investigation of factors related to self-efficacy for Java programming among engineering students. *The Turkish Online Journal of Educational Technology, 8*(1), 26–32. Retrieved from http://files.eric.ed.gov/fulltext/ED503900.pdf.

Austin, E. J., Deary, I. J., Whiteman, M. C., Fowkes, F. G. R., Pedersen, N. L., Rabbitt, P., … McInnes, L. (2002). Relationships between ability and personality: Does intelligence contribute positively to personal and social adjustment? *Personality and Individual Differences, 32*(8), 1391–1411. https://doi.org/10.1016/S0191-8869(01)00129-5.

Baker, T. J., & Bichsel, J. (2006). Personality predictors of intelligence: Differences between young and cognitively healthy older adults. *Personality and Individual Differences, 41*(5), 861–871. https://doi.org/10.1016/j.paid.2006.02.017.

Balanskat, A., & Engelhardt, K. (2014). *Computing our future: Computer programming and coding-Priorities, school curricula and initiatives across Europe*. Brussels: European Schoolnet. Retrieved from http://www.eun.org/c/document_library/get_file?uuid=521cb928-6ec4-4a86-b522-9d8fd5cf60ce&groupId=43887.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive perspective*. Englewood Cliffs, NJ: Princeton-Hall.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1.

Barbaranelli, C., Caprara, G. V., Rabasca, A., del Barrio, M. V., Carrasco Ortiz, M.Á., & Holgado Tello, F. P. (2006). *BFQ-NA: Cuestionario "Big Five" de personalidad para niños y adolescentes (Manual Técnico)* [*BFQ-NA: Big Five Personality Questionnaire for Children and Adolescents (Technical Manual)*]. Madrid, Spain: TEA Ediciones.

Barbaranelli, C., Caprara, G. V., Rabasca, A., & Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences, 34*(4), 645–664. https://doi.org/10.1016/S0191-8869(02)00051-X.

Bazelli, B., Hindle, A., & Stroulia, E. (2013). On the personality traits of stackoverflow users. In *Software maintenance (ICSM), 2013 29th IEEE international conference on* (pp. 460–463). https://doi.org/10.1109/ICSM.2013.72.

Benotti, L., Martínez, M. C., & Schapachnik, F. (2014). Engaging high school students using chatbots. In *Proceedings of the 2014 conference on Innovation & technology in computer science education* (pp. 63–68). https://doi.org/10.1145/2591708.2591728.

Benotti, L., Martinez, M. C., & Schapachnik, F. (2017). A tool for introducing computer science with automatic formative assessment. *IEEE Transactions on Learning Technologies*. https://doi.org/10.1109/TLT.2017.2682084 (in press).

Beyer, S. (2014). Why are women underrepresented in Computer Science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. *Computer Science Education, 24*(2–3), 153–192. https://doi.org/10.1080/08993408.2014.963363.

Bocconi, S., Chioccariello, A., Dettori, G., Ferrari, A., Engelhardt, K., & others. (2016). *Developing computational thinking in compulsory education-implications for policy and practice*. Seville: Join Research Center (European Commission). Retrieved from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC104188/jrc104188_computhinkreport.pdf.

Brady, C., Orton, K., Weintrop, D., Anton, G., Rodriguez, S., & Wilensky, U. (2017). All roads lead to computing: Making, participatory simulations, and social computing as pathways to computer science. *IEEE Transactions on Education, 60*(1), 59–66. https://doi.org/10.1109/TE.2016.2622680.

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American educational research association, Vancouver, Canada* (pp. 1–25). Retrieved from http://scratched.gse.harvard.edu/ct/files/AERA2012.pdf.

Burke, Q. (2012). The markings of a new pencil: Introducing programming-as-writing in the middle school classroom. *Journal of Media Literacy Education, 4*(2), 121–135. Retrieved from http://files.eric.ed.gov/fulltext/EJ985683.pdf.

Busch, T. (1995). Gender differences in self-efficacy and attitudes toward computers. *Journal of Educational Computing Research, 12*(2), 147–158. https://doi.org/10.2190/%2FH7E1-XMM7-GU9B-3HWR.

Bush, C. M., & Schkade, L. L. (1985). In search of the perfect programmer. *Datamation, 31*(6), 128–132. Retrieved from http://dl.acm.org/citation.cfm?id=15198.

Caplan, S. E. (2005). A social skill account of problematic Internet use. *Journal of Communication, 55*(4), 721–736. https://doi.org/10.1111/j.1460-2466.2005.tb03019.x.

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The "Big Five Questionnaire": A new questionnaire to assess the five factor model. *Personality and Individual Differences, 15*(3), 281–288. https://doi.org/10.1016/0191-8869(93)90218-R.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*(3), 404–431. https://doi.org/10.1037/0033-295X.97.3.404.

Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences, 44*(7), 1596–1603. https://doi.org/10.1016/j.paid.2008.01.003.

Chamorro-Premuzic, T., & Furnham, A. (2014). *Personality and intellectual competence*. New York: Psychology Press.

Charlton, J. P., & Danforth, I. D. W. (2007). Distinguishing addiction and high engagement in the context of online game playing. *Computers in Human Behavior, 23*(3), 1531–1548. https://doi.org/10.1016/j.chb.2005.07.002.

Cheryan, S., Plaut, V. C., Handron, C., & Hudson, L. (2013). The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women. *Sex Roles, 69*(1), 58–71. https://doi.org/10.1007/s11199-013-0296-x.

Code.org. (2017). *The hour of code: Impact on attitudes towards and self-efficacy with computer science*. Retrieved from https://code.org/files/HourOfCodeImpactStudy_Jan2017.pdf.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155.

Cohoon, J. M. (2007). Gendered experiences of computing graduate programs. In *Proceedings of the 38th SIGCSE technical symposium on Computer science education* (pp. 546–550). New York: ACM. https://doi.org/10.1145/1227504.1227494.

CSTA, & ISTE. (2011). *Operational definition of computational thinking for K–12 education*. Retrieved from http://www.iste.org/docs/ct-documents/computational-thinking-operational-definition-flyer.pdf.

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social coding in GitHub: Transparency and collaboration in an open software repository. In *Proceedings of the ACM 2012 conference on computer supported cooperative work* (pp. 1277–1286). New York: ACM. https://doi.org/10.1145/2145204.2145396.

Dagiene, V., & Futschek, G. (2008). Bebras international contest on informatics and computer literacy: Criteria for good tasks. In *International conference on informatics in secondary schools-evolution and perspectives* (pp. 19–30). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-540-69924-8_2.

Dagiene, V., & Stupuriene, G. (2014). Informatics education based on solving attractive tasks through a contest. In *KEYCIT 2014: Key competencies in informatics and ICT* (pp. 51–62). Postdam, Germany: Universitätsverlag Potsdam. Retrieved from http://www.bebras.org/sites/default/files/documents/publications/Dagiene.2014.pdf.

Daily, S. B., Leonard, A. E., Jörg, S., Babu, S., & Gundersen, K. (2014). Dancing alice: Exploring embodied pedagogical strategies for learning computational thinking. In *Proceedings of the 45th ACM technical symposium on Computer science education* (pp. 91–96). https://doi.org/10.1145/2538862.2538917.

Elithorn, A., & Telford, A. (1969). Computer analysis of intellectual skills. *International Journal of Man-Machine Studies, 1*(2), 189–209. https://doi.org/10.1016/S0020-7373(69)80021-0.

Espino, E. E., & González, C. S. (2016). Estudio sobre Pensamiento Computacional y Género [A study on computational thinking and gender]. *Versión Abierta Español Portugués – Revista Iberoamericana de Tecnologías Del Aprendizaje, VAEP-RITA, 4*(3), 119–128. Retrieved from http://rita.det.uvigo.es/VAEPRITA/201609/uploads/VAEP-RITA.2016.V4.N3.pdf#page=21.

García-Peñalvo, F. J. (2016). A brief introduction to TACCLE 3—Coding European project. In F. J. García-Peñalvo, & A. J. Mendes (Eds.), *2016 international symposium on computers in education (SIIE 16)* (pp. 1–4). USA: IEEE. https://doi.org/10.1109/SIIE.2016.7751876.

García-Peñalvo, F. J., Reimann, D., Tuul, M., Rees, A., & Jormanainen, I. (2016). *An overview of the most relevant literature on coding and computational thinking with emphasis on the relevant issues for teachers*. Belgium: TACCLE 3 Consortium. https://doi.org/10.5281/zenodo.165123.

Gnambs, T. (2015). What makes a computer wiz? Linking personality traits and programming aptitude. *Journal of Research in Personality, 58*, 31–34. https://doi.org/10.1016/j.jrp.2015.07.004.

Goldberg, L. R. (1990). An alternative" description of personality": The big-five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216.

Gouws, L. A., Bradshaw, K., & Wentworth, P. (2013). Computational thinking in educational activities: An evaluation of the educational game light-bot. In *Proceedings of the 18th ACM conference on Innovation and technology in computer science education* (pp. 10–15). https://doi.org/10.1145/2462476.2466518.

Grover, S. (2015). "Systems of Assessments" for deeper learning of computational thinking in K-12. In *Proceedings of the 2015 annual meeting of the American educational research association* (pp. 15–20). Retrieved from https://www.sri.com/sites/default/files/publications/aera2015-_systems_of_assessments_for_deeper_learning_of_computational_thinking_in_k-12.pdf.

Grover, S., Cooper, S., & Pea, R. (2014). Assessing computational learning in K-12. In *Proceedings of the 2014 conference on Innovation & technology in computer science education* (pp. 57–62). https://doi.org/10.1145/2591708.2591713.

Grover, S., & Pea, R. (2013). Computational thinking in K–12 a review of the state of the field. *Educational Researcher, 42*(1), 38–43. https://doi.org/10.3102/0013189X12463051.

Herodotou, C., Kambouri, M., & Winters, N. (2014). Dispelling the myth of the socio-emotionally dissatisfied gamer. *Computers in Human Behavior, 32*, 23–31. https://doi.org/10.1016/j.chb.2013.10.054.

Higgins, D. M., Peterson, J. B., Pihl, R. O., & Lee, A. G. M. (2007). Prefrontal cognitive ability, intelligence, Big Five personality, and the prediction of advanced academic and workplace performance. *Journal of Personality and Social Psychology, 93*(2), 298–319. https://doi.org/10.1037/0022-3514.93.2.298.

Howland, K., & Good, J. (2015). Learning to communicate computationally with Flip: A bi-modal programming language for game creation. *Computers & Education, 80*, 224–240. https://doi.org/10.1016/j.compedu.2014.08.014.

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education, 28*(1), 1–35. https://doi.org/10.1007/s10212-011-0097-y.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104*(1), 53–69. https://doi.org/10.1037/0033-2909.104.1.53.

Irani, L. (2004). Understanding gender and confidence in CS course culture. In *Proceedings of the 35th SIGCSE technical symposium on Computer science education* (pp. 195–199). https://doi.org/10.1145/1028174.971371.

John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The" little five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65*(1), 160–178. https://doi.org/10.2307/1131373.

Joo, Y. J., Lim, K. Y., & Kim, J. (2013). Locus of control, self-efficacy, and task value as predictors of learning outcome in an online university context. *Computers & Education, 62*, 149–158. https://doi.org/10.1016/j.compedu.2012.10.027.

Kafai, Y. B., & Burke, Q. (2013). The social turn in K-12 programming: Moving from computational thinking to computational participation. In *Proceedings of the 44th ACM technical symposium on computer science education - SIGCSE '13* (pp. 603–608). New York: ACM. https://doi.org/10.1145/2445196.2445373.

Kagan, D. M., & Douthat, J. M. (1985). Personality and learning FORTRAN. *International Journal of Man-Machine Studies, 22*(4), 395–402. https://doi.org/10.1016/S0020-7373(85)80046-8.

Kagan, D. M., & Esquerra, R. L. (1984). Personality and learning BASIC. *Journal of Instructional Psychology, 11*(1), 10–16. Retrieved from http://psycnet.apa.org/psycinfo/1985-05281-001.

Kalelioglu, F., Gülbahar, Y., & Kukul, V. (2016). A framework for computational thinking based on a systematic research review. *Baltic Journal of Modern Computing, 4*(3), 583–596. Retrieved from http://www.bjmc.lu.lv/fileadmin/user_upload/lu_portal/projekti/bjmc/Contents/4_3_15_Kalelioglu.pdf.

Kelleher, C., Pausch, R., & Kiesler, S. (2007). Storytelling alice motivates middle school girls to learn computer programming. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 1455–1464). https://doi.org/10.1145/1240624.1240844.

Ketler, K., & Smith, R. D. (1992). Differences between third and fourth generation programmers: A human factor analysis. *Information Resources Management Journal (IRMJ), 5*(2), 25–35. https://doi.org/10.4018/irmj.1992040103.

Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences, 25*, 67–72. https://doi.org/10.1016/j.lindif.2013.01.005.

Kukul, V., Gökçearslan, S., & Günbatar, M. S. (2017). Computer programming self-efficacy scale (CPSES) for secondary school students: Development, validation and reliability. *Eğitim Teknolojisi Kuram ve Uygulama, 7*(1). Retrieved from http://dergipark.ulakbim.gov.tr/etku/article/view/5000195912.

Lakhani, K. R., & Von Hippel, E. (2003). How open source software works:"free" user-to-user assistance. *Research Policy, 32*(6), 923–943. https://doi.org/10.1016/S0048-7333(02)00095-1.

Law, K. M. Y., Lee, V. C. S., & Yu, Y.-T. (2010). Learning motivation in e-learning facilitated computer programming courses. *Computers & Education, 55*(1), 218–228. https://doi.org/10.1016/j.compedu.2010.01.007.

Licorish, S. A., & MacDonell, S. G. (2014). Personality profiles of global software developers. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering* (pp. 45:1–45:10). New York: ACM. https://doi.org/10.1145/2601248.2601265.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*(6), 1479–1498. https://doi.org/10.2307/1130467.

Lye, S. Y., & Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Computers in Human Behavior, 41*, 51–61. https://doi.org/10.1016/j.chb.2014.09.012.

Manovich, L. (2013). *Software takes command*. UK: Bloomsbury Academic.

Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology, 160*, 92–106. https://doi.org/10.1016/j.jecp.2017.03.013.

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27*(4), 267–298. https://doi.org/10.1016/S0160-2896(99)00016-1.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004.

Medel, P., & Pournaghshband, V. (2017). Eliminating gender bias in computer science education materials. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education* (pp. 411–416). https://doi.org/10.1145/3017680.3017794.

Mervielde, I., Buyst, V., & De Fruyt, F. (1995). The validity of the Big-Five as a model for teachers' ratings of individual differences among children aged 4–12 years. *Personality and Individual Differences, 18*(4), 525–534. https://doi.org/10.1016/0191-8869(94)00175-R.

Moreno-León, J., & Robles, G. (2015). Dr. Scratch: A web tool to automatically evaluate scratch projects. In *Proceedings of the workshop in primary and secondary computing education* (pp. 132–133). New York: ACM. https://doi.org/10.1145/2818314.2818338.

Moreno-León, J., Robles, G., & Román-González, M. (2015). Dr. Scratch: Automatic analysis of scratch projects to assess and foster computational thinking. *RED. Revista de Educación a Distancia, 15*(46), 1–23. Retrieved from http://www.um.es/ead/red/46/moreno_robles.pdf.

Moreno-León, J., Robles, G., & Román-González, M. (2016). Examining the relationship between socialization and improved software development skills in the scratch code learning environment. *Journal of Universal Computer Science, 22*(12), 1533–1557. https://doi.org/10.3217/jucs-022-12-1533.

Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: A study using the NEO personality inventory and the Myers–Briggs type indicator. *European Journal of Personality, 17*(1), 79–94. https://doi.org/10.1002/per.471.

Newsted, P. R. (1975). Grade and ability predictions in an introductory programming course. *ACM SIGCSE Bulletin, 7*(2), 87–91. https://doi.org/10.1145/382205.382897.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338. https://doi.org/10.1037/a0014996.

Poropat, A. E. (2011). The Eysenckian personality factors and their correlations with academic performance. *British Journal of Educational Psychology, 81*(1), 41–58. https://doi.org/10.1348/000709910X497671.

Poropat, A. E. (2014). A meta-analysis of adult-rated child personality and academic performance in primary education. *British Journal of Educational Psychology, 84*(2), 239–252. https://doi.org/10.1111/bjep.12019.

Prensky, M. (2008). *Programming is the new literacy*. Edutopia. Retrieved from https://www.edutopia.org/literacy-computer-programming.

Ramalingam, V., LaBelle, D., & Wiedenbeck, S. (2004). Self-efficacy and mental models in learning to program. In *Proceedings of the 9th annual SIGCSE conference on Innovation and technology in computer science education* (pp. 171–175). New York: ACM. https://doi.org/10.1145/1007996.1008042.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., et al. (2009). Scratch: Programming for all. *Communications of the ACM, 52*(11), 60–67. https://doi.org/10.1145/1592761.1592779.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387. https://doi.org/10.1037/a0026838.

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin, 143*(2), 117–141. https://doi.org/10.1037/bul0000088.

Román-González, M. (2014). Aprender a programar "apps" como enriquecimiento curricular en alumnado de alta capacidad [To learn programming "apps" as curriculum enrichment on gifted students]. *Bordón. Revista de Pedagogía, 66*(4), 135–155. https://doi.org/10.13042/Bordon.2014.66401.

Román-González, M. (2015). Computational thinking test: Design guidelines and content validation. In *Proceedings of the 7th annual international conference on education and new learning technologies (EDULEARN 2015)* (pp. 2436–2444). https://doi.org/10.13140/RG.2.1.4203.4329.

Román-González, M. (2016). *Codigoalfabetización y pensamiento computacional en educación primaria y secundaria: validación de un instrumento y evaluación de programas [Code-literacy and computational thinking in primary and secondary education:...].* Spain: Universidad Nacional de Educación a Distancia. Retrieved from http://e-spacio.uned.es/fez/view/tesisuned:Educacion-Mroman.

Román-González, M., Moreno-León, J., & Robles, G. (2017). Complementary tools for computational thinking assessment. In S. C. Kong, J. Sheldon, & K. Y. Li (Eds.), *Proceedings of international conference on computational thinking education (CTE 2017)* (pp. 154–159). The Education University of Hong Kong. Retrieved from http://www.eduhk.hk/cte2017/doc/CTE2017.Proceedings.pdf.

Román-González, M., Pérez-González, J.-C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the computational thinking test. *Computers in Human Behavior, 72*, 678–691. https://doi.org/10.1016/j.chb.2016.08.047.

Román-González, M., Pérez-González, J.-C., Moreno-León, J., & Robles, G. (2016). Does computational thinking correlate with personality?: The non-cognitive side of computational thinking. In *Proceedings of the fourth international conference on technological ecosystems for enhancing multiculturality* (pp. 51–58). New York: ACM. https://doi.org/10.1145/3012430.3012496.

Rushkoff, D. (2010). *Program or be programmed: Ten commands for a digital age*. New York: OR Books.

Rushkoff, D. (2012). *Code literacy: A 21st-century requirement*. Edutopia. Retrieved from https://www.edutopia.org/blog/code-literacy-21st-century-requirement-douglas-rushkoff.

Sadin, É. (2015). *La vie algorithmique: critique de la raison numérique* [*The Algorithmic Life: A Critique of Digital Rationality*]. Paris: L'Echappée.

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. Flanagan, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99–144). New York: Guilford.

Schott, G., & Selwyn, N. (2000). Examining the "male, antisocial" stereotype of high computer users. *Journal of Educational Computing Research, 23*(3), 291–303. https://doi.org/10.2190/V98R-5ETX-W9LY-WD3J.

Seehorn, D., Carey, S., Fuschetto, B., Lee, I., Moix, D., O'Grady-Cunniff, D., et al. (2011). *CSTA K–12 computer science standards: Revised 2011*. Retrieved from http://c.ymcdn.com/sites/www.csteachers.org/resource/resmgr/Docs/Standards/CSTA_K-12_CSS.pdf.

Snow, E. (2014). *Assessing computational thinking*. SRI International. Retrieved from https://pdfs.semanticscholar.org/8c9b/7d3f0f5ab5ce5e1eb2a77613772a1561fe59.pdf.

Vekiri, I., & Chronaki, A. (2008). Gender issues in technology use: Perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. *Computers & Education, 51*(3), 1392–1404. https://doi.org/10.1016/j.compedu.2008.01.003.

Von Stumm, S., & Ackerman, P. L. (2013). Investment and intellect: A review and meta-analysis. *Psychological Bulletin, 139*(4), 841–869. https://doi.org/10.1037/a0030746.

Von Stumm, S., Chamorro-Premuzic, T., Quiroga, M. A., & Colom, R. (2009). Separating narrow and general variances in intelligence-personality associations. *Personality and Individual Differences, 47*(4), 336–341. https://doi.org/10.1016/j.paid.2009.03.024.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250–270. https://doi.org/10.1037/0033-2909.117.2.250.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., et al. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25*(1). https://doi.org/10.1007/s10956-015-9581-5.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35. https://doi.org/10.1145/1118178.1118215.

Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 366*(1881), 3717–3725. https://doi.org/10.1098/rsta.2008.0118.

Wing, J. M. (2011). *Research notebook: Computational thinking—What and why?*. The Magazine of the Carnegie Mellon University School of Computer Science. Retrieved from http://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why.

Yilmaz, M., & O'Connor, R. V. (2012). Towards the understanding and classification of the personality traits of software development practitioners: Situational context cards approach. In *Software engineering and advanced applications (SEAA), 2012 38th EUROMICRO conference on* (pp. 400–405). https://doi.org/10.1109/SEAA.2012.62.

Yilmaz, M., O'Connor, R. V., Colomo-Palacios, R., & Clarke, P. (2017). An examination of personality traits and how they impact on software development teams. *Information and Software Technology, 86*, 101–122. https://doi.org/10.1016/j.infsof.2017.01.005.

Zuffianò, A., Alessandri, G., Gerbino, M., Kanacri, B. P. L., Di Giunta, L., Milioni, M., et al. (2013). Academic achievement: The unique contribution of self-efficacy beliefs in self-regulated learning beyond intelligence, personality traits, and self-esteem. *Learning and Individual Differences, 23*, 158–162. https://doi.org/10.1016/j.lindif.2012.07.010.