Research Article doi: 10.12973/ijem.5.3.421

International Journal of Educational Methodology
Volume 5, Issue 3, 421 - 432.
ISSN: 2469-9632
http://www.ijem.com/

Development of Computational Thinking Scale: Validity and Reliability Study

Buket Ertugrul-Akyol*
Erciyes University, TURKEY

Received: June 23, 2018 ▪ Revised: July 28, 2018 ▪ Accepted: July 30, 2019

Abstract: Computational thinking is a way of thinking that covers 21st century skills and includes new generation concepts such as robotics, coding, informatics and information construction. Computational thinking has reached an important point especially in the field of science in line with the rapid developments in technology. Robotics applications, software-based activities, STEM (Science, Technology, Engineering, Math) education and problem-based studies are some of the areas where this thinking is used. In this study, which is based on this point, it is aimed to develop a scale for computational thinking. Exploratory sequential design, one of the mixed research methods, was used in the study. First of all, a detailed literature review was conducted and needs analysis was carried out. This study consists of two stages. In the first stage, exploratory factor analysis was performed and analyzed with SPSS 23 program. In the second stage, confirmatory factor analysis was performed and analyzed with LISREL 9.2 program. As a result of the study, the goodness of fit indexes of the scale was found. According to this; X^2/df value 1.81; NNFI value 0.97; NFI value 0.93; CFI value 0.98; RMR value 0.05; SRMR value 0.04; AGFI value 0.91 and GFI value was found to be 0.93. When the reliability values of the study were examined, Cronbach's Alpha value was found to be 0.86. As a result of the research, a computational thinking scale consisting of 3 factors and 30 items was developed. This scale was developed for prospective teachers and can be used at all levels of prospective teachers.

Keywords: *Computational thinking, scale development, 21st century skills, science education.*

To cite this article: Ertugrul-Akyol, B. (2019). Development of Computational Thinking Scale: Validity and reliability study. *International Journal of Educational Methodology*, *5*(3), 421-432. https://doi.org/10.12973/ijem.5.3.421

Introduction

Changes in science and technology have also affected individuals' thinking and behavioral patterns (Dalrymple, 2011). Critical thinking, analytical thinking and problem-solving became particularly important in the 21st century (Yilmaz, Gulgun, Cetinkaya & Doganay, 2018). Today, in addition to these developments, another area of thinking called computational thinking has emerged. According to Wing (2006, p.33), computational thinking can be defined as "To solve problems by using the basic concepts of computer science, to design systems and to think like a computer scientist". The concept of computational thinking has a structure based on the idea of "calculation". When the development process from the past to the present is examined, the machines that make the first calculation are actually people (Light, 1999). In the 1900s, especially during the war periods, there were only officials working for calculations. In addition, ENIAC, the first programmable computer, was developed and introduced to humanity in 1946 (Schneider & Gersting, 2016). The calculation here is not only to achieve a certain result by performing four basic operations. Computing in computer science means; constructing algorithms, making logical inferences and making a choice as a result of conditional propositions (Denning, 2016). Today, many of the interfaces used in social media accounts work with this logic and proceed with a certain algorithmic workflow (Cinar & Tuzun, 2017; Feurzeig & Papert, 2011). Based on these statements, computational thinking, "It can be expressed as a process of creating new information or decisions that make sense through a certain algorithmic process by making calculations and inferences" (Cetin & Toluk Ucar, 2017; Kalelioglu, Gulbahar & Kukul, 2016). There are some behavioral patterns that students, individuals and people of all ages are expected to gain with computational thinking skills (Ozden, 2015; Wing, 2006). These are;

* **Correspondence:**
Buket Ertugrul Akyol, Erciyes University, Institute of Educational Sciences, Department of Science Education, Kayseri- Turkey.
✉ bukertakyol@gmail.com

1. Re-formulating relatively large and difficult problems in a simpler and easier manner,
2. To gain the ability of recursive thinking (thinking over and over again, thinking continuously by renewing),
3. Gaining the ability of abstraction and analysis,
4. Separating the focal points of the study and focusing on the parts, not the whole,
5. Determining the behavior of the system examined by determining the variables,
6. Use of heuristic reasoning,
7. Working as a computer scientist and gaining the ability to make multi-level abstraction constitute these patterns.

The concept of computational thinking has a new name in recent years with the increase in computer technologies and artificial intelligence applications (Selby & Woollard, 2013). "Thinking Like a Computer Scientist" is concept that actually summarizes the whole process because a computer scientist does not approach events from an ordinary perspective. As a matter of fact, he has to think like a computer and has to act according to the working principles of a computer. When this statement is examined carefully, the manner in which a computer scientist approaches the events can be defined as follows (Burton, 2010; Cetin & Toluk Ucar, 2017; Kramer, 2007);

1. To formulate and disassemble problems and problem situations using existing and alternative tools,
2. Analyze and organize the available data using a logical process,
3. To be able to create fast and automatic solutions by using algorithmic thinking patterns,
4. Calculate and analyze possible solutions and put them into practice,
5. There are approaches to adapting and transferring to a number of possible problems that they may face in the future by structuring and storing a problem-solving process they face.

Innovations in technology now affect individuals' habits and learning activities. The most concrete examples of this are the increase in robotic applications as power passes and the introduction of software-based applications into all areas of our lives (Yilmaz & Ertugrul Akyol, 2017). Education programs in our country are constantly updated and efforts are made to keep up with the era we live in. In this context; textbooks are renewed, alternative measurement and evaluation systems are used, technology-based applications are adapted to courses and course environments. The important point here is how individuals will adapt to these innovations. Existing learning systems and ways of thinking are no longer as effective (Barr, 2014). As a natural consequence, new and effective ways of thinking are preferred. Computational thinking is a frequently used form of thinking (Denning, 2014). The following suggestions were made about what the components of computational thinking skills are (Aho, 2012; Cetin & Toluk Ucar, 2017);

1. Having problem solving skills,
2. Recognizing and distinguishing the types of problems,
3. Subdividing and analyzing problems,
4. Abstraction and metacognitive thinking,
5. Ability to think algorithmically,
6. Preparing and evaluating algorithms,
7. Pattern identification and generalization are examined under seven sub-headings.

Considering the use of computational thinking, it is clear that this way of thinking has an indispensable importance in the 21st century (Brennan & Resnick, 2012). For this purpose, in order to contribute to education scientists and related field, it was decided to conduct a scale development study to measure the computational thinking tendencies of individuals. When the related literature is examined for the computational thinking approach, a scale was developed by Whetton and Cameron (2002). The name of this scale is "How Creative Are You?". In this scale, it is aimed to measure the creative thinking skills of students and their ability to process information. Korkmaz, Cakir and Ozden (2017) developed the "Computational Thinking Scale". This scale has sub-factors such as creativity, algorithmic thinking, critical thinking, cooperativity and problem solving. Gulbahar, Bahadir Kert and Kalelioglu (2019) developed "The Self-Efficacy Perception Scale for Computational Thinking Skill". Algorithm design, data processing competence, basic programming competence, self-confidence competence and problem solving competence are the sub-factors of this study. As can be seen, these studies are not directly related to the subject of the researcher in robotics, coding, software, STEM education and professional career. Therefore, a new scale has been developed to serve the purpose.

Methodology

In this study, a mixed research method was used in which qualitative and quantitative data collection tools were used together. In the qualitative phase of the study, document analysis and thematic content analysis were performed, and a detailed needs analysis was performed. In the quantitative stage of the study, computational thinking scale was developed by using the survey method. The research method, which consists of a combination of these two different processes, is an exploratory sequential pattern (Tabachnick & Fidel, 2007). The exploratory sequential pattern is the studies in which the research problem was first started with a qualitative process and then continued with a quantitative process, and as a result, a measurement tool was produced (Acar, 2017). Since this study is mainly a scale development study, qualitative parts are used only in needs analysis and information is given about the analyses in the discussion section. Therefore, mainly quantitative processes were explained in this study.

Participant Characteristics and Sampling

The research process consists of two stages. The first is exploratory factor analysis, and the second is confirmatory factor analysis. Therefore, the study has two different groups of participants. The computational thinking scale was administered to prospective teachers (since the study process was carried out at university level, the scale was conducted on prospective teachers) studying in science teaching at universities. In this context, universities in the realization of exploratory factor analysis 1, 2, 3 and 4th in Turkey and has provided a total of 426 prospective teachers studying in various universities of participation. Confirmatory factor analysis is also provided for the realization of the previous total of 342 participants who read a variety of university teachers from different groups in Turkey as a candidate for accession.

Appropriate sampling method and snowball sampling method were used together in the determination of prospective teachers. The purpose of using these sampling methods is to provide the researcher with ease in terms of time, labor and cost and to reach a wide range of research. First of all, prospective teachers the university where the application was made were reached. Later, with the help of colleagues working in this field, the scales were applied in many universities in the country. The aim is to reach as many people as possible.

Data Collection Tools

Within the scope of the study, “Computational Thinking Scale” was developed by the researcher as a data collection tool. This scale consisted of 3 factors and 30 items. Scale factors were robotic coding and software, computational thinking, professional development and career planning. A number of procedures were applied during the preparation phase of the scale developed by the researcher.

First, exploratory factor analysis was performed. At this stage, the pool of items, expert opinion, content and appearance validity, pilot implementation and data collection, data analysis (SPSS 23) were obtained. In the second confirmatory factor analysis, construct validity, convergent and divergent validity values were calculated. Finally, the expert opinion was re-applied, and the scale was finalized.

Results

In the scope of the study, the first draft items were presented to expert opinion. Table 1 presents the results of the expert opinion (Content Validity Ratio – CVR / Content Validity Indexes - CVI) using the Lawshe (1975) technique.

Table 1. Expert opinion results

Item	CVR	Item	CVR	Item	CVR	Item	CVR	Item	CVR
1	0.87	12	0.93	23	0.87	34	0.93	45	0.87
2	0.93	13	0.93	24	0.87	35	0.87	46	0.93
3	0.87	14	0.87	25	0.93	36	0.40	47	0.87
4	0.93	15	0.93	26	0.87	37	0.93	48	0.93
5	0.93	16	0.87	27	0.93	38	0.87	49	0.93
6	0.40	17	0.93	28	0.93	39	0.93	50	0.40
7	0.40	18	0.93	29	0.40	40	0.87	51	0.40
8	0.93	19	0.40	30	0.40	41	0.93	52	0.87
9	0.87	20	0.40	31	0.87	42	0.93	53	0.93
10	0.40	21	0.93	32	0.87	43	0.40	54	0.93
11	0.40	22	0.93	33	0.93	44	0.40	55	0.93
Overall CVI = 0.87									

Table 2 shows the number of expert opinions and acceptable content validity values used in Lawshe technique.

Table 2. Lawshe (1975) technique experts and acceptable value ranges

Number of Experts	Minimum CVR Value	Number of Experts	Minimum CVR Value
5	0.99	13	0.54
6	0.99	14	0.51
7	0.99	15	0.49
8	0.78	20	0.42
9	0.75	25	0.37
10	0.62	30	0.33
11	0.59	35	0.31
12	0.56	40+	0.29

In this study, 55 items were determined by experts and doctoral theses in the field of scale development, 1 Professor, 4 Associate Professors, 5 Doctors and 5 research assistants who are similar in line with their opinions, do not fit the scale structure, do not enter the subject area and thought to serve the purpose the number of items was reduced to 42 by re-examining the related literature. When the CVR value of the scale items was examined, the lowest item was 0.87, and the highest item was 0.93. In addition, the CVI value of the overall scale was found to be 0.87. This shows that the scale items explain the structure by 87%. In the light of the expert opinions of the prepared draft scale items, pilot applications were made first, and the results obtained were examined. In this context, there are prerequisites to be performed and some procedures to be performed for factor analysis (Cokluk, Sekercioglu & Buyukozturk, 2014). These operations were data set and determination of lost data, control of the assumption of normality, determination of extreme values and examination of the multi-connection problem. First of all, the data set was examined, missing data were determined, normality assumption was checked and extreme values (eight extreme values) were determined. Then, multiple connection cases were examined (Tolerance Value = 0.79; 0.85; 0.91 / Variance Inflation Factor Value = 1.12; 1.87; 1.96) and exploratory factor analysis phase was started after the related arrangements were made.

Table 3. Kaiser-Meyer-Olkin (KMO) and Bartlett sphericity test results

KMO Coefficient		0.91
Bartlett Sphericity Test	Chi-square value	7013.07
	df	435
	p (p<0.05)	0.00

When Table 3 was examined, it was found that Kaiser-Meyer-Olkin (KMO) value was 0.91 and Bartlett Sphericity value was found to be significant at $p < 0.05$ level. The fact that KMO value is greater than 0.50, and Bartlett Sphericity test result is significant ($p < 0.05$) indicates that there is a sufficient sample size for factor analysis and that there is a desired relationship between scale items (Tabachnick & Fidel, 2007).

Table 4. Eigenvalue and variance ratio results

Number of factors	Eigenvalue	Variance (%)	Cumulative (%)
1	8.26	25.05	25.05
2	4.89	17.82	42.88
3	2.83	10.44	53.33
4	0.88		
5	0.75		

When Table 4 is examined, by approximately 5 factors whose eigenvalues were greater than 1 and 1. In this context, the explained variance ratio of factor 1 was 25.05%, the explained variance ratio of the second factor was 17.82%, and the explained variance ratio of the third factor was 10.44%. The explained variance ratio of the whole scale was determined as 53.33%. It is stated that this value should be at least 30% in single-factor studies and should be at least 40% and above in multi-factor studies (Simsek, 2007). It can be said that the total variance of the scale developed in this context is sufficient (Tavsancil, 2006). Figure 1 shows the scree plot graph of the scale.

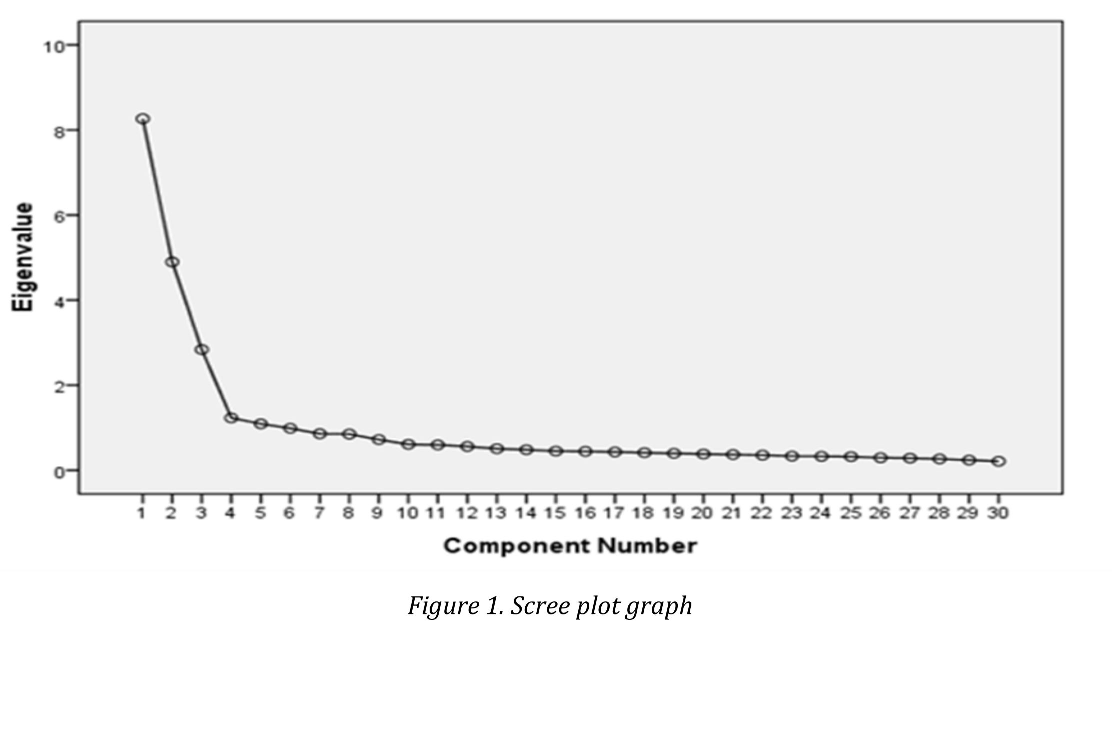

Figure 1. Scree plot graph

Figure 1 shows that the eigenvalue for the scale can be composed of 4 factors at the breaking point where there are many values greater than 1 and 1. However, the scree plot graph should be evaluated and made significant with the eigenvalue ratios and explained variance ratio in the factor determination process (Cokluk et al., 2014). For this reason, it can be said that it is appropriate to use a 3-factor structure because the factor groups determined by the researcher (determined as 3 factors) are sufficient, and the results obtained are within the desired value ranges (McMillan & Schumacher, 2006). Another point that should be checked in exploratory factor analysis is item factor loads (Fraenkel & Wallen, 2003). In many studies, it is stated that this value is accepted as 0.30 or above (Buyukozturk, 2010). Selecting this value at higher rates will cause the research to have better quality scale items. In this case, it will make your work qualified. Because of this functionality, this value was determined as 0.40 or above. In Table 5, item factor loads and Cronbach's Alpha values of the factors were presented.

Table 5. Item factor loads and Cronbach's Alpha values

	Factor Loads				
Item No	**1**	**2**	**3**	**Rotated Loads**	**Cronbach's Alpha**
A1	0.67			0.79	
A2	0.78			0.78	
A3	0.30			0.73	
A4	0.11			0.72	
A5	0.79			0.72	
A6	0.30			0.70	
A7	0.73			0.69	
A8	0.65			0.69	**0.92**
A9	0.70			0.68	
A10	0.72			0.67	
A11	0.22			0.66	
A12	0.66			0.65	
A13	0.59			0.64	
A14	0.69			0.61	
A15	0.64			0.59	
A16	0.68				
A17	0.61				
A18	0.69				
A19	0.22				
B1		0.68		0.74	
B2		0.73		0.74	
B3		0.68		0.73	
B4		0.67		0.72	
B5		0.69		0.72	
B6		0.74		0.70	**0.84**
B7		0.70		0.69	
B8		0.37		0.68	
B9		0.74		0.68	
B10		0.15		0.67	
B11		0.72			
B12		0.29			
B13		0.72			
B14		0.22			
C1		0.33		0.75	
C2		0.70		0.74	
C3		0.72		0.72	**0.88**
C4		0.67		0.70	
C5		0.74		0.67	
C6		0.36			
C7		0.75			
C8		0.11			
C9		0.04			
			Overall Cronbach's Alpha (42 items)		**0.86**

Table 5 shows that the value range of item factor loads varies between 0.59 and 0.79. Item factor loads are expected to be higher than 0.32 when the related literature is examined. However, this criterion was determined as 0.40 in our study. In this context, a total of 12 items were removed and 30 items remained. When the scale is examined, it is seen that item factor loads are within the acceptable value range. In Table 6, item-scale correlations and t-test results between groups were presented.

Table 6. Item-scale correlations and t-test results between group means

Item no	Item-total Correlations	The t-value of the difference between the Sub/upper group means	Item no	Item-total Correlations	The t-value of the difference between the Sub/upper group means
1	0.62**	8.88*	**22**	0.55**	7.45*
2	0.76**	10.15*	**23**	0.57**	6.87*
3	0.11	-	**24**	0.60**	8.25*
4	0.19	-	**25**	0.77**	8.74*
5	0.84**	11.24*	**26**	0.71**	8.42*
6	0.21	-	**27**	0.10	-
7	0.77**	10.02*	**28**	0.72**	10.75*
8	0.64**	9.05*	**29**	0.21	-
9	0.79**	11.74*	**30**	0.86**	11.97*
10	0.69**	9.94*	**31**	0.22	-
11	0.20	-	**32**	0.88**	12.04*
12	0.70**	9.21*	**33**	0.15	-
13	0.82**	8.96*	**34**	0.14	-
14	0.63**	8.54*	**35**	0.75**	11.74*
15	0.57**	7.46*	**36**	0.74**	11.05*
16	0.73**	10.25*	**37**	0.73**	11.32*
17	0.63**	9.45*	**38**	0.69**	10.41*
18	0.55**	7.14*	**39**	0.22	-
19	0.11	-	**40**	0.65**	9.75*
20	0.67**	9.45*	**41**	0.24	-
21	0.66**	9.85*	**42**	0.13	-

* $p < 0.01$ ** Correlation is significant at 0.01 level.

Table 6 shows the item total correlation values and the 27% subgroup upper group averages. In this context, the questions numbered 3, 4, 6, 11, 19, 27, 29, 31, 33, 34, 39, 41 and 42 were obtained because the correlation values and t-test results were considered, and the results are similar to Table 5. Table 7 shows the correlation results of scale factors.

Table 7. Correlation results for factors

Factors	Factor 1	Factor 2	Factor 3
Factor 1	1.00	0.89	0.91
Factor 2	0.89	1,00	0.92
Factor 3	0.91	0.92	1,00

When Table 7 is examined, it is seen that there were three different factors that make up the scale. The correlation values of these factors were quite high. This shows that the construct validity and appearance validity of the scale are provided and the scale items are meaningful as a whole. In the literature review, it is seen that scale development studies have a two-stage structure (George & Mallery, 2010; Ozturk, 2010). First, exploratory factor analysis was performed. The resulting scale structure was then subjected to confirmatory factor analysis. The main purpose of confirmatory factor analysis was to check the accuracy of the structure performed by exploratory factor analysis and to ensure the construct validity (convergent and divergent validity) of the scale (Cokluk et al., 2014). Figure 2 shows the structural equation model which is the result of confirmatory factor analysis. Table 8 shows the goodness of fit index values obtained by the CFA analysis.

Table 8. Goodness of fit indexes values

Fit Indexes	Values	Comment
X^2	1414.14	Appropriate value range
Sd	780	Appropriate value range
X^2/df	1.81	Perfect fit
p	0.57	0,05 pointless
RMSEA	0.03	Perfect fit
NFI	0.93	Perfect fit

Table 8. Continued

Fit Indexes	Values	Comment
NNFI	0.97	Perfect fit
CFI	0.98	Perfect fit
RMR	0.05	Perfect fit
SRMR	0.04	Perfect fit
AGFI	0.91	Perfect fit
GFI	0.93	Perfect fit
CN	304.68	Perfect fit

When the general evaluation of the structure revealed as a result of EFA analysis is made as a result of CFA and SEM analysis, it was seen that the construct validity of the scale developed was provided and found values were within the target value ranges.

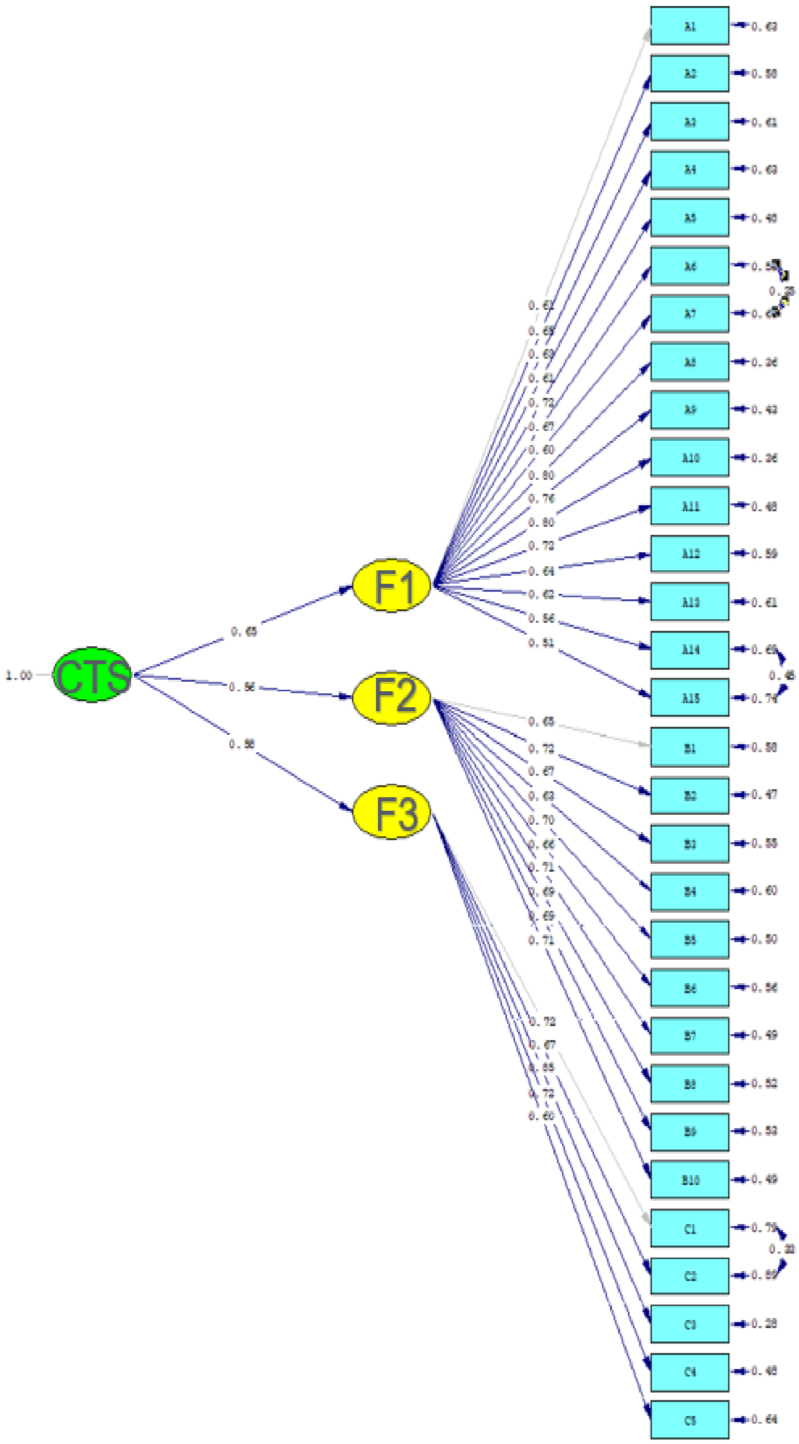

Figure 2. Confirmatory factor analysis model

Discussion and Conclusion

When the research findings related to the validity analyzes of the scale development studies conducted in the field of educational sciences are examined (Gul & Sozbilir, 2015; Kucuk, Yilmaz, Baydas & Goktas, 2014), it was seen that the scope and appearance validity is one of the most preferred validity types due to the nature of the scale development studies. While many studies (Dalgic, 2008; Kurnaz & Yigit, 2010; Yilmaz & Aydin, 2017) have included findings and information regarding the content validity of the literature, it has been seen that this information is often explained in a short way and no clear information can be found. Another issue that attracted the attention of the researcher in the literature review was content validity and appearance validity were often thought to serve the same purpose, and in most of the studies, the first stage of scale development was done. While the content validity helps to evaluate the whole structure as a whole, the appearance validity helps the researchers in terms of the fact that the completed structure is measuring the structure that it wants to measure and it seems to serve the purpose (Gul & Sozbilir, 2015). In this context, it is possible to use the content validity from the first stage to the last stage of the study, while the appearance validity should be used as the type of validity that should be made after the scale structure is completed. In this respect, it is considered appropriate to provide a detailed information about the content and appearance validity of the study by the researcher.

When the scale development studies were performed on the content validity are examined, there were two types of investigations as stated by Erkus (2012). The first of these is the logical way (non-statistical), and the second is the statistical way. Logical examinations are often studies in which a general assessment is made by interview or written and oral notification (Yurdagul & Bayrak, 2012). Statistical studies are the studies that use statistical procedures such as content validity ratio and content validity index to understand the developed scale items, the appropriateness of the collected data to the targeted sampling and so on. In order for the expert opinions obtained from the preliminary studies to be valid and in harmony, content validity ratio and content validity index values developed by Lawshe (1975) and updated by Wilson, Pan and Donald (2012) should be examined. Lawshe technique requires at least 5 and at most 40 experts (Yurdagul & Bayrak, 2012). Each item that is thought to be included in the scale is rated in the form of expert opinions. In this study, 15 experts were consulted, and a very large number of experts were reached. In the scope of the study, approximately 55 items were examined by experts, 1 professor, 4 associate professors, 5 assistant professors and 5 research assistants. The number of items was reduced to 42 by re-examining the related literature.

Within the scope of the research, first of all, the data whose content and appearance validity was completed, and which was made suitable for pilot application as a result of expert opinions and as a result of this, pilot data were analyzed. The analysis phase of the data consists of 2 parts. In the first part, exploratory factor analysis was performed with the help of SPSS 23 package program, and then the scale structures determined were subjected to confirmatory factor analysis by means of LISREL 9.2 package program, and necessary analyses were performed to verify their structures. There were some prerequisites and some procedures to be performed in order to perform exploratory factor analysis (Cokluk et al., 2014). Data set was examined, lost data were determined, normality hypothesis was checked, extreme values were determined, and multiple connection problem was examined. After necessary corrections were made, exploratory factor analysis was started (Tabachnick & Fidell, 2007). When the results of factor analysis were analyzed, the first data were found to be suitable by using Kaiser-Meyer-Olkin (KMO) coefficient and Bartlett sphericity test (Cokluk et al., 2014). The fact that KMO value was greater than 0.50 and Bartlett sphericity test result was significant ($p < 0.05$) indicated that there was a sufficient sample size for factor analysis, and that there was a desired level of relationship between scale items (Tabachnick & Fidel, 2007). In addition, a significant Bartlett sphericity test indicated that the data fulfills the characteristics of linearity and multiple homogeneity. As a natural consequence of this situation, it can be said that the data comes from a normal distribution (Tabachnick & Fidell, 2007; Yilmaz & Aydin, 2017). Correlation matrix and anti-image correlation matrices of the scale were also examined, and it can be said that the relations of the scale items with each other were at the desired level, and there was not a very high relationship between the scale items (this is a desired case). As a result, there was no multiple connection problem (Simsek, 2007). If there was a relationship between 0,90 and above, the scale items should be combined because both scale items serve almost the same purpose. For this purpose, when the first subscale was examined, it was seen that all scale items are in the desired value ranges.

After examining the suitability of the data for factor analysis, it was decided to determine how many factors the scale would consist of. The most commonly used method for determining factors is the eigenvalue statistics and scree plot results. Determining the number of factors in studies conducted to develop the scale is often among the departments where the researchers spend the most time (Cokluk et al., 2014). The point to be considered here is to pay attention to the selection of factors with eigenvalue results greater than 1 and 1 (Buyukozturk, 2010). When the eigenvalue table of our scale was examined, by approximately 4 factors with eigenvalues greater than 1 and 1. However, when the field literature was examined, it was stated that it was appropriate to terminate the factor groups when an eigenvalue ratio starts to appear less than the multiples of 2 and 2 (this ratio was seen as 3 or 4 times in some sources) (Tabachnick & Fidell, 2007). When the eigenvalue and variance ratios table continues to be examined, it was seen that the total variance rate explained is 53.33%.

It was stated that this value should be at least 30% in single-factor studies and should be at least 40% and above in multi-factor studies (Simsek, 2007). It can be said that the total variance of the scale developed in this context was sufficient (Tavsancil, 2006). When the studies carried out by developing scales, factor groups were often determined beforehand, and whether the results obtained were in compliance with these determined factors, and often the items of the scale were not subject to examination at the factor level. It is also observed. When the Figure 1 is examined, it can be seen that the eigenvalue for the scale can be composed of 4 factors at the breaking point where there are many values greater than 1 and 1. However, the scree plot graph should be evaluated and made significant with the eigenvalue ratios and explained variance ratio in the factor determination process (Cokluk et al., 2014). For this reason, it can be said that it is appropriate to use a 3-factor structure because the factor groups determined by the researcher (determined as 3 factors) are sufficient and the results obtained are within the desired value ranges (McMillan & Schumacher, 2006).

When item factor loads and common variance values were examined, it was stated that this value should be 0.32 or above (Buyukozturk, 2010). Selecting this value at higher rates will provide the research to have higher quality scale items, and in this case, it will make your work qualified. Because of this functionality, this value was determined as 0.40 or above. After determining the factors in the study, it was decided to have a three-factor structure within the framework of the related field literature. The factors created were named as computational thinking, robotic coding and software, professional development and career planning. Considerations for selecting and eliminating items for the scale developed by the researcher are detailed below;

1. First of all, the item pool was created by supporting the literature and presented to the expert opinion. Lawshe (1975) technique was used for content and appearance validity.
2. After the expert's opinion, the pool of items was applied to a group of 426 participants as a pilot application and the data obtained were examined. At this stage, substances which did not show normal distribution were determined, but not in the first stage. After determining these items, correlation matrices and anti-image matrices, lower and upper group 27% item total correlations, common variances, item factor loads and whether or not having overlapping values were examined. After all these evaluation stages, the items that were decided to be removed were evaluated in several different aspects and subjected to removal procedures and finally, the expert opinion was re-applied (Buyukozturk, 2010).
3. As a result of the examination of item 3 factor loads, the lower limit of 0.32, which has a general validity, was taken into consideration and this value was determined as 0.40 by the researchers in order to have higher quality of the study.
4. Finally, 27% of the subgroup and the upper group were examined and item-total correlations were excluded from the scale structure of $p < 0.05$. As a result, none of the items extracted in the scale structures were directly screened, and their specific processes were evaluated together and subjected to a detailed screening process.

After exploratory factor analysis, confirmatory factor analysis was performed. The results of the confirmatory factor analysis should be compared with the goodness of fit index values. The first process to be performed as a result of CFA will be to examine the X^2/df ratio. It was stated that this value is acceptable in some sources as 5 or less, that it is at a good level of harmony if it is 3 or less, and that it indicates perfect fit when it is 2 or less (Calvini, Fini & Ranieri, 2008). This value indicates that the sample size was sufficient and that the items of the scale could come together under certain groups in a meaningful way. It corresponds to some kind of KMO and Bartlett test. The second step in the CFA analysis is to check *p* significance level and RMSEA index value. *P* significance value is expected to be $p > 0.05$ here (Kline, 1994). This is a problem arising from the assumption that the scale items prepared in likert type are assumed to be continuous data (Cokluk et al., 2014). Therefore, many goodness of fit indices, especially RMSEA, should be examined respectively. First of all, SEM model and path diagram are obtained in CFA analysis. Here, standard values and t-values should be examined respectively. Item factor loads are reached with standard values and error rate is determined for each variable. The error rate was expected to be 0.90 and above. When t-values are considered, all of these values were expected to be higher than the limit value of 1.96 (Yilmaz, 2018). When the fit indexes obtained from the CFA analysis of the scale are examined, it was seen that the X^2/df value had a perfect fit. This shows that the sample size was sufficient to test the construct validity of the road analysis generated by SEM and that the scale items could be collected under certain groups. When the relevant literature on the sample size was examined, several different views emerge (Ardies, Maeyer & Gijbels, 2013; Gul & Sozbilir, 2015). These thoughts that prefer a sample size of at least 300 or more in the Likert type scale applications of the sample size, and that the sample should be used between at least 5 and 10 times the number of items in the scale. In this respect, our sample size ($n = 342$) complies with both views, but supports both X^2/df and CN = 304.68. When other fit indices obtained from CFA analysis were examined, RMSEA, NFI, NNFI, CFI, SRMR, RMR, CN, AGFI and GFI values were found to have excellent compatibility. When the general evaluation of the structure revealed as a result of EFA analysis was made as a result of CFA and SEM analysis, it was seen that the construct validity of the scale developed was provided, and the values found were within the target value ranges. Within the framework of the results of the study, the following recommendations can be made to the researchers;

1. Computational thinking has a very broad framework. Therefore, limit your topics when creating an item pool.

2. Good statistical knowledge is required in the scale development process. It is recommended that you first receive a sufficient level of statistical training before you begin.

References

Acar, I. H. (2017). Karma yontem arastirmalarina giris [Introduction to mixed method research]. In M. Sozbilir (Eds.). *Temel ve gelismis karma yontem desenleri* [Basic and advanced mixed method designs] (pp. 35-51). Ankara, Turkey: Pegem Academy.

Aho, A. V. (2012). Computation and computational thinking. *The Computer Journal, 55*(7), 832-835.

Ardies, J., Maeyer, S.D., & Gijbels, D. (2013). Reconstructing the pupils attitude towards technology survey. Design and Technology Education: *An International Journal, 18*(1), 8-19.

Barr, V. (2014). Computational thinking. In T. Gonzalez & A. Tucker (Eds.), *Computation handbook* (3rd ed.). Abingdon, UK: Chapman & Hall/CRC Press.

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.

Burton, B. A. (2010). Encouraging algorithmic thinking without a computer. *Olympiads in Informatics*, *4,* 3-14.

Buyukozturk, S. (2010). *Sosyal bilimler icin veri analizi el kitabi: Istatistik, arastirma deseni SPSS uygulamalari ve yorum* [Data analysis manual for social sciences: Statistics, research design SPSS applications and interpretation] (11th ed.). Ankara, Turkey: Pegem Academy.

Calvini, A., Fini, A., & Ranieri M. (2008). Models and instruments for assessing digital competence at school. *Journal of E-Learning and Knowledge Society, 3*(4), 183-193.

Cetin, I., & Toluk Ucar, Z. (2017). Bilgi islemsel dusunmeden programlaya [From computational thinking to programming]. In Y. Gulbahar (Eds.), *Bilgi islemsel dusunme tanimi ve kapsami* [Definition and content of computational thinking] (pp.42-78). Ankara, Turkey: Pegem Academy.

Cinar, M., & Tuzun, H. (2017). *Bilgisayimsal dusunme surecinin dogasina iliskin nitel bir analiz* [A qualitative analysis of the nature of computational thinking]. Paper presented at the 19th Academic Informatics Conference. Aksaray University, Aksaray, Turkey.

Cokluk, O., Sekercioglu, G., & Buyukozturk, S. (2014). *Sosyal bilimler icin cok degiskenli istatistik: SPSS ve LISREL uygulamalari* [Variable statistics for social sciences: SPSS and LISREL applications]. Ankara, Turkey: Pegem Academy.

Dalgic, G. Y. (2008). *Turk yuksekogretiminde ogretim elemanlarinin Bologna sureci kapsamindaki uygulamalara iliskin gorusleri* [The views of the academic staff in Turkish higher education about the applications within the Bologna process] (Unpublished doctoral dissertation). Gazi University, Ankara, Turkey.

Dalrymple, P. W. (2011). Data, information, knowledge: The emerging field of health informatics. *Bulletin of the American Society for Information Science and Technology, 20*(2), 307-315.

Denning, P. J. (2014). Structure and organization of computing. In T. Gonzales, J. Diaz-Herrera & A.Tucker (Eds), *Computing handbook*: *Computer Science and Software Engineering* (3rd ed.) (pp. 1-14). Boca Raton, FL: Chapman & Hall/CRC.

Denning, P. J. (2016). Remaining trouble spots with computational thinking. *Communications of the ACM, 60*(6), 33-39. https://doi.org/ 10.1145/2998438

Erkus, A. (2012). *Psikolojide olcek ve olcek gelistirme-I: Temel kavramlar ve islemler* [Scale and scale development in psychology-I: Basic concepts and operations] (1st ed.). Ankara: Pegem Academy.

Feurzeig, W., & Papert, S. A. (2011). Programming-languages as a conceptual framework for teaching mathematics. *Interactive Learning Environments, 19*(5), 487-501. https://doi.org/10.1080/10494820903520040

Fraenkel, J. R., & Wallen, N. E. (2003). *How to design and evaluate research in education*, (5th ed.). New York, NY: McGraw-Hill.

George, D., & Mallery, M. (2010). *SPSS for windows step by step: A simple guide and reference.* Boston, MA: Pearson.

Gulbahar, Y., Kert, S. B., & Kalelioglu F. (2019). The self-efficacy perception scale for computational thinking skill: Validity and reliability study. *Turkish Journal of Computer and Mathematics Education, 10*(1), 1-29.

Gul, S., & Sozbilir, M. (2015). Thematic content analysis of scale development studies published in the field of science and mathematics education. *Education and Science, 40*(178), 85-102. http://dx.doi.org/10.15390/EB.2015.4070

Kalelioglu, F., Gulbahar, Y., & Kukul, V. (2016). A framework for computational thinking based on a systematic research review. *Baltic Journal of Modern Computing, 4*(3), 583-596.

Kert, S. B. (2017). Bilgi islemsel dusunmeden programlaya [From computational thinking to programming]. In Y. Gulbahar (Ed.), *Bilgisayar bilimi egitimine giris* [Introduction to computer science education] (pp.1-20). Ankara, Turkey: Pegem Academy.

Kline, P. (1994). *An easy guide to factor analysis*. Abingdon, UK: Routledge.

Korkmaz, O., Cakir, R., & Ozden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). *Computers in Human Behavior, 72*, 558-569. https://doi.org/10.1016/j.chb.2017.01.005

Kramer, J. (2007). Is abstraction the key to computing? *Communications of the ACM, 50*(4), 36-42.

Kucuk, S., Yilmaz, R. M., Baydas, O., & Goktas, Y. (2014). Okullarda artirilmis gerceklik uygulamalari tutum olcegi: Gecerlik ve guvenirlik calismasi [Augmented reality applications attitude scale in schools: Validity and reliability study]. *Education and Science, 39*(174), 383-392. http://dx.doi.org/10.15390/EB.2014.3590

Kurnaz, M. A., & Yigit, N. (2010). Fizik tutum olcegi: Gelistirilmesi, gecerliligi ve guvenilirligi [Physical attitude scale: Development, validity and reliability]. *Necatibey Faculty of Education, Electronic Journal of Science and Mathematics Education, 4*(1), 29-49.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563-575. http://dx.doi.org/10.1111/j.1744-6570.1975.tb01393.x

Light, J. S. (1999). When computers were women. *Technology and Culture, 40*(3), 455-483.

McMillan, J. H., & Schumacher, S. (2006). *Research in education: Evidence-based inquiry* (6th ed.). Boston, MA: Pearson.

Ozden, M. Y. (2015). *Bilgisayarca dusunme* [Computer Thinking]. Retrieved on June 2, 2019 from http://myozden.blogspot.com.tr

Ozturk, M. A. (2010). An exploratory study on measuring educators' attitudes toward educational research. *Educational Research and Reviews, 5*(12), 758-769.

Schneider, G. M., & Gersting, J. (2016). *Invitation to computer science.* Toronto, Canada: Nelson Education.

Selby, C., & Woollard, J. (2013). Computational thinking: The developing definition. In J. Carter, I. Utting, & A. Clear (Eds.), *Proceedings of 18th Annual Conference on Innovation and Technology in Computer Science education.* Canterbury: University of Southampton.

Simsek, O. F. (2007). *Yapisal esitlik modellemesine giris: Temel ilkeler ve LISREL uygulamalari* [Introduction to structural equation modeling: Basic principles and LISREL applications] Ankara, Turkey: Ekinoks.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn ve Bacon.

Tavsancil, E. (2006). Tutumlarin olculmesi ve SPSS ile veri analizi [Measurement of attitudes and data analysis with SPSS] (3rd ed.) Ankara, Turkey: Nobel.

Whetton, D. A., & Cameron, K. S. (2002). *Answers to exercises taken from developing management skills* (3rd ed.). Evanston, IL: Northwestern University.

Wilson, F. R., Pan. W., & Donald, A. S. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development, 45*(3), 197-210. https://doi.org/10.1177/0748175612440286

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33-35.

Yilmaz, A., Gulgun, C., Cetinkaya, M., & Doganay, K. (2018). Initiatives and new trends towards STEM education in Turkey. *Journal of Education and Training Studies, 6*(11), 1-10.

Yilmaz, A., & Ertugrul Akyol, B. (2017). Required quality standards for augmented reality applications. *International Journal on Lifelong Education and Leadership, 3*(2), 13-21.

Yilmaz, A., & Aydin, S. (2017). *Quality standards for the content of the program and admission to the students in science education teacher training programs*. Paper presented at the International Teacher Education Conference (ITEC), Harvard University in Cambridge, MA, USA.

Yilmaz, A. (2018). *The determination of the quality standards of teacher training programs related to teachers of science education: The study of scale development and application* (Unpublished doctoral dissertation). Kastamonu University, Kastamonu, Turkey.

Yurdagul, H., & Bayrak, F. (2012). Olcek gelistirme calismalarinda kapsam gecerlik olcutleri: Kapsam gecerlik indeksi ve Kappa istatistiginin karsilastirilmasi [Content validity criteria in scale development studies: Comparison of content validity index and Kappa statistics]. *Hacettepe University Journal of Education, 2*(Special Issue), 264-271.

Appendix-1

Factors		COMPUTATIONAL THINKING SCALE	Absolutely agree	I agree	Undecided	Do not agree	Strongly disagree
Computational Thinking	1	I can solve the problems I face with computational thinking skills in a more systematic way.					
	2	I can distinguish between the concept of computer science, construction and informatics.					
	3	I can show computational thinking and theoretical and applied behaviors together.					
	4	Computational thinking makes it easy for me to understand the concepts of data, information, information and technology.					
	5	I think my computational thinking and problem solving skills increase.					
	6	I think my computational skills, such as classification, classification and grouping, have improved through computational thinking.					
	7	I can say that my computational thinking skills and individual research independence have improved.					
	8	With computational thinking, I can focus more comfortably on the process.					
	9	When I practice with computational thinking in science class, I concentrate more easily.					
	10	With computational thinking, I can handle many of the problems in my life in a more logical way.					
	11	It offers a system of work focused on computational thinking process and product.					
	12	Computational thinking allows me to approach individual and group work in a more moderate way.					
	13	Computational thinking allows me to follow today's technology more closely.					
	14	I gain the behavior of systematically approaching problems through computational thinking.					
	15	Computational thinking is a long-term process that requires intensive attention and gives the ability to work disciplined for a long time.					
Robotic Coding and Software	16	I can assimilate information-oriented application processes (robotics, coding) more easily.					
	17	I can adapt more quickly to software-based content development.					
	18	I think my scientific process skills have improved with robotic coding, software skills, and computational thinking activities.					
	19	I'm not afraid of the complexity of software systems.					
	20	Computational thinking increases my interest and curiosity in software and robotics.					
	21	I believe that software education will be the education system of the future.					
	22	I'm more interested in coding and robotics every day.					
	23	I would like to develop my own software language if I have the opportunity.					
	24	I think it would be appropriate to include software courses at undergraduate level.					
	25	I think software training should start at a very early age.					
Professional Development and Career Planning	26	I can use computer science, software technology, hardware technology and internet technology in a multidisciplinary way with computational thinking.					
	27	I'm thinking of working for a big software company in the future.					
	28	Computational thinking provides students with the requirements of the digital age at undergraduate level.					
	29	I think to use the technologies I developed for the welfare of society.					
	30	I think that there should be mass awareness and state support on robotics, coding, software, information and information processing thinking.					