

Learning Difficulties in Computing Courses: Cognitive Processes Assessment Methods Research and Application

Vinicius Vieira Pessoni¹, Fernando Marques Federson², Auri Marcelo Rizzo Vincenzi³

Instituto de Informática
Universidade Federal de Goiás - UFG
Goiânia - Goiás, Brazil

¹viniciuspessoni@ufg.br, ²fernandofederson@inf.ufg.br, ³auri@inf.ufg.br

ABSTRACT

Learning difficulties in computing courses is a situation perceived in diverse universities from different countries, cultures and backgrounds. These difficulties directly affect achievement rates and increase course evasion. We believe in the existence of a foundation of cognitive processes, that without it, even the most motivated student would have trouble to transform the received information into knowledge. This work has focused mainly on the research of candidate methods for cognitive processes assessment with a strong background theory. With this kind of information would be possible to devise cognitive interventions, in order to evolve students cognitive level, and consequently, raise their success rates. A systematic review was conducted and among the many researched methods we selected Lawson Classroom Test of Scientific Reasoning – LCTSR. Authorized by its author, we conducted the first translation of LCTSR to Brazilian Portuguese and administered to students of three undergraduate computing courses: Information Systems, Computer Science and Software Engineering. We also present results of its administration that we consider important to reinforce the above suggested strategy.

Categories and Subject Descriptors

K.3.2 [Computers and Education]: Computer and Information Science Education – Information systems education

General Terms

Teaching, Learning, Cognition, Measurement, Human Factors.

Keywords

Computing Learning Difficulties, Cognitive Development, Formal Operational Stage, Scientific Reasoning.

1. INTRODUCTION

Student's learning difficulties in computing courses is not a recent, isolated and local preoccupation. Instead, it is a challenge perceived across the years in diverse universities from different countries, cultures and backgrounds [15, 28, 45, 52, 54]. These difficulties can be noted on student's low grades, achievement rates and course retention. They also influence negatively on course's engagement [6].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SBSI 2015, May 26–29, 2015, Goiânia, Goiás, Brazil.
Copyright SBC 2015.

Despite it is relatively new, the Computer Education Research (CER) is going beyond to only propose tools to support learning computing and is gaining attention and importance as a wide research field [46, 51]. We agree with Malmi et al. [45] that the Computer Science, where theories are few and the constructive design and formal approaches dominate, has not favored the extensive use of existing theory as a normal part of research, like in the natural and human sciences. Therefore, we share the vision that a more holistic understanding of CER, beyond our basic approach, is necessary not only to consolidate its identity and independence as a research area, but also to achieve long term results.

There are many available tools to improve teaching and learning experiences. Through the years, the authors have been studying and applying many of them like: Bloom's taxonomy [31], as a classification of the different learning objectives but also as a guide to students about the cognitive process of learning in association with Kolb's learning cycle [30]; Myers-Briggs Type Indicator (MBTI), to measure psychological preferences in how people perceive the world and make decisions [9]; Kolb's Learning Style [27] and Index of Learning Styles (ILS) [17, 18], to measure the way individuals process information or prefer to learn. Some methodologies have also been used, like: Problem-based Learning (PBL) [3, 21] and Peer Instruction [14, 48], among others. There have also been some collaborative efforts from other groups in order to mitigate these issues, for instance [15, 47].

All these tools and methodologies are known by our scientific community. However, the deep understanding of their background theories has been of extremely importance for the authors to move their researches toward the comprehension of how knowledge construction is established on human beings, especially on those who left infancy phase. This way, even though the knowledge construction partially depends on factors like how the information is transmitted, student motivation, personality and interest, we believe that there is an inner foundation of cognitive processes, that without it, even the most motivated student would have difficulties to transform information into knowledge.

The analysis of anatomical and physical foundations of learning and memory is one of the greatest achievements of modern neuroscience. Thirty years ago little was known about "what are the different types of memory?", "where in the brain is memory located?" or "how memory is retrieved to answer a question?". Surely nowadays, knowledge is advancing and new findings are published every day [11, 49].

Although the researchers did not present a single or uniform theoretical viewpoint, it seems that the current research on learning and memory could be brought under a unifying

theoretical framework. This framework can be used to build and test metacognition models and strategies [66]. In this scenario, there is less room for intuitive models. Simons and Chabris [61], for example, shown that substantial numbers of their survey respondents agreed with propositions that conflicts with expert consensus about memory. They also emphasized that incorrect beliefs about the memory properties have broad implications, including the students misunderstanding of the memory role in learning processes.

Despite our approach is not precisely new, we believe it is relevant, being necessary for the community to revisit the previous existent models supported by neuroscience refreshing point of view. Especially in our country, Brazil, where teaching and learning indicators are not satisfactory, this kind of research would be a valuable contribution in order to develop strategies to improve teaching and learning processes.

The first step in this direction would be the identification of some kind of cognitive processes assessment method with a strong background theory. Then, with this kind of instrument, its model and some experience, it would possible to devise and apply approaches to verify the possibility of acquisition and development of cognitive processes on young and adults (metacognition).

The present work has focused mainly on the research of candidates methods to support this mission. We expect to be possible not only identify, characterize and measure cognitive processes (levels), but also to offer approaches to acquire and develop them. Therefore, reducing students gap among their cognitive level and the one expected by the courses, increasing the students success on these courses.

Among many researched methods, we choose and translated Lawson Classroom Test of Scientific Reasoning [34]. We also show results of its administration considered important to reinforce the above suggested strategy.

2. METHODOLOGY

The research was conducted through the following phases: 2.1. Cognitive Processes Assessment Methods Review; 2.2. Translation and Pilot Testing of LCTSR; 2.3. Application of the Translated Method. These phases are detailed in the next subsections. Results and data analysis are detailed on section 3.

2.1 Cognitive Processes Assessment

Methods Review

A systematic review [29] was conducted in order to identify and select appropriate methods to assess University alumni cognitive processes. We researched relevant works in thirteen well settled computing and interdisciplinary digital libraries: ACM Digital library; IEEEExplore; ISI Web of Knowledge; ScienceDirect; Scopus; APA Psychnet; EBSCO (Academic Search Premier; CINAHL; SocINDEX; Library, Information Science & Technology Abstracts; Information Science & Technology Abstracts (ISTA); Academic Search Elite; Computers & Applied Sciences Complete; MEDLINE Complete).

In accordance with the objectives, we aimed to find instruments with a solid theoretical foundation, that would allow not only to identify, characterize and measure cognitive processes but would give us a more accurate information in order to develop approaches to generate permanent improvements on them. Among the instrument's background theories, Piaget's cognitive development [24] is one of the most solid and popular, being inspiration for many assessment methods.

Since Piaget developed his stage theory in the 50's [24, 53], its influences in teaching and learning science has pervaded numerous scientific productions, in multiple countries and in proportions that cannot even be measured. Therefore, today, many decades later, his insights are still very relevant in many ways [32].

Piaget [24], describes the cognitive development assessment method called "the clinical method" that consists of interviewing a person while presenting her a set of experiments with specific apparatus. Based on the person's answers it should be possible to classify her on the following cognitive development stages: sensorimotor, preoperational, concrete operations and formal operations [26]. Formal operations would be the stage in which the person would be more capable to deal with complex modes of reasoning, or in other words, abstraction.

Despite the researchers' preference for this method, the difficulties on applying it in large groups are listed as motivation for existence of almost all the posterior works that intends to perform Piaget's developmental level identification e.g [1, 34]. Some of recurrent critics on the clinical interview are the necessary expertise of the interviewer, the excessive space and time consumed, and the variation on influence of human factors in results, or lack of standardization.

The need of instruments that would be standardized, reliable and relatively simple to be administered to a large number of participants culminated in diverse evolutions and adaptations of the clinic interview, each one with its particularities. These adapted instruments differ in many aspects like: place of origin, presentation method, application field and distinction among cognitive stages.

The researches of methods are not restricted to one culture, country or continent, and each of them has its own values and contributions. Among these contributions, some notorious examples across the world are Shayer [60] in United Kingdom, Longeot [41] in France, Lawson [34] in United States of America, Bond [7] in Australia and Roadrangka [56] in Thailand.

The variations in presentation method occur in a way that some of them involve specific apparatus and experiments manipulation, and some do not. The experiments may be executed either by students themselves on their own kits or presented by the test administrator in front of class [4, 20, 44, 58, 60]. The apparatus and experiments required are basically the ones used by Piaget in the clinical method or some variations that are claimed to measure the same principles.

Still about the presentation variations, methods that do not use special apparatus or experiments and are fully described in paper, are generically known as pure or strict paper and pencils methods. These instruments are presented on questionnaire format, where the subject must read some breath paragraphs describing the question and sometimes analyze some graphical representation like pictures or diagrams. The answer sometimes open ended, other times objective, must be developed by himself, without test administrator intervention in the discovering process [1, 7, 10, 13, 41, 55-57, 59, 63-65].

In application field, the assessment methods may vary according to the age range of individuals. Some tests are adapted and validated for children [55], in others, adolescents age range are included [13, 19, 34, 63]. In a third variation are those that may be suitable from children to adults [4, 44, 56, 57, 64].

If the method is able to identify in which developmental stage individuals belong, it may be also a factor of distinction among tests. On the one hand, there are tests intended only to assess the presence of formal thinking, identifying some reasoning

patterns of this stage [4, 55, 64]. On the other hand there are tests that make a better distinction between formal operational and concrete, or previous stages, categorizing the subject in one of the four developmental stages [34, 56, 63].

Despite the previous variations, some others are described in [4]. For the present research, we preferred tests that could be administered to adolescents and adults, once its application is expected to be for University students, minimal age of 15 years old. Despite Shayer's [60] method have a great influence on scientific productions, it is preferable tests that do not involve specific apparatus manipulation and that could be objective scored. It is also desirable to distinct between formal operations and concrete or previous stages, to obtain some more complete information.

Thus, summarizing the criteria used for evaluating the assessment methods we expect that: 1. The method has a solid background theory; 2. The method does not involve any special apparatus manipulation; 3. Turns possible to clearly distinct the formal operational stage from previous stages; 4. Is suitable for adolescents and adults, people older than 15 years old; 5. Is objective scored, with no open ended questions.

According to these five criteria, after the systematic review, three methods of cognitive development assessment were selected as strong candidates for administration on our University: Arlin Test of Formal Reasoning – ATFR [1], Bond's Logical Operations Test – BLOT [7] and Lawson Classroom Test of Scientific Reasoning – LCTSR [34, 35]. Despite ATFR and BLOT demonstrated to be strong candidates, LCTSR has many advantages when compared to them. Some of these advantages are the massive cross-cultural previous applications and mainly the inclusion of the evolutions of cognitive theory through the years. This way, this exam was selected to be the first translated and used to assess alumni cognitive development on our University. The LCTSR is commented in the following section.

2.1.1 Lawson Classroom Test of Scientific Reasoning – LCTSR

The LCTSR is a paper and pencil instrument released in 1978. On its original version was comprised of 15 questions and some kind of demonstrations. In this format, the test administrator performs a demonstration in front of the class and poses a question. Students note their answers in test booklets, without sharing answers.

The instrument went through successive revisions and improvements across the years, with its most recent version released August 2000 [35]. In its current version, the exam is a pure paper and pencil instrument, with 24 two-tier items. Each of the test items require for the subject to select, from multiple alternatives and multiple justifications, in other words, the correct answer and justification for it. Scoring may be single or two tier type. This version has been designed to assess five reasoning modes of formal reasoning: controlling variables, proportional reasoning, combinatorial reasoning, probabilistic reasoning, and correlational reasoning. In addition, it has been extended to measure the hypothetical-deductive reasoning pattern, identifying the post-formal stage, of neo-piagetian theories [36].

Based on the student score on LCTSR, he is classified in one of the following cognitive levels: concrete (0-8); transitional (9-14); formal operational (15-20); post-formal (21-24).

Although some critics exist [8, 20, 42], LCTSR is wide validated and very popular among Science, Technology, Engineering, and Mathematics (STEM) educators and researchers not only in USA [16, 23]. It has been massively administered in several universities and countries, to students from different

graduation courses e.g. [2, 25, 43] including computing majors [50]. Due its strong validity and consistency, it has even become inspiration for many other tests [12, 33, 56, 64]. Its use is related not only to the intention to predict academic success, which is been concluded to be a valid predictor, but to develop materials and classes for cognitive interventions, helping students to improve their cognitive reasoning skills. We are particularly interested in this kind of use.

2.2 Translation and Pilot Testing of LCTSR

Since none of the researched methods were available in Brazilian Portuguese, translation and adjustments were necessary. After selecting LCTSR and verifying that no Portuguese version of it was available, the present authors communicated electronically with its author, Professor Lawson, requiring for permissions to translate and apply his test in our University. Once obtained the permissions, best practices on translation and methods for conducting it were researched.

To conduct a translation, especially of some assessment method, is necessary to follow a protocol to ensure that the translated method will be equivalent and assesses the same characteristics that the original does. Simple one direction translation, made by a single translator, is not recommended because the risks of translator bias and other personal inconsistencies that can be inserted during translating process [62]. This way, translation and cross-cultural adaptation must be carried out in a systematic process, following well defined phases.

In order to keep the translated instrument as equivalent as possible to the original, in the present research, the translation was conducted according to international well settled translation guidelines [5, 62] comprising the five following steps:

1) Independent Forwarding Translation: the English source test (ST) was translated by two bilingual independent translators (T1) and (T2). T1 was an informed translator, which means that he was familiar with the instrument background theories and objectives. T2 was an uninformed translator, meaning that he was not familiar with the method background theories neither its objectives. The results of this step were two translated distinct versions in Brazilian Portuguese (V1) and (V2).

2) Third Person Review: versions V1 and V2 were revised by a third bilingual person (T3), graduated in American Literature. V1 and V2 were compared with each other, and with the ST. The inconsistencies were noted in each translation, for further discussion in the next step. The results of this step were two revised and annotated distinct versions of the questionnaire (RV1) and (RV2).

3) Committee Review and Merge: the revised versions RV1 and RV2 were revised in a committee approaching, comprising four persons: the first independent translators T1 and T2, the T3 specialist reviewer, plus one of the present authors. Translation's divergences were discussed and resolved in group consensus, generating the merged candidate version (C).

4) Pilot Experimentation: the candidate version C, was then tested in a small group of voluntaries to verify translation and application inconsistencies. This group was very heterogeneous, with 25 subjects, with age range from 22 to 57, undergraduate, graduate and non undergraduate. Although there was no time limitation to accomplish the test, it was asked to voluntaries to annotate their start and end time, in order to measure how long they would take to answer it. The information of how much time they spent was used to calibrate the amount of time given to subjects on subsequent applications. The test was anonymous.

They were asked about the terminology used on the test and, if they had any difficulties on understanding about what was requested in each question. The participants' doubts on terminology and in any other aspect of the test were noted for further revision.

5) Final Revision and Version: the pilot group feedback was revised and some few mistakes on C were resolved. After including these adjustments, the final version (F) was generated. This final version was the one administered to computing courses alumni.

2.3 Subjects and LCTSR Application

After concluding the translation process, we administered the final version of the questionnaire to 210 subjects from seven different groups. 187 of these subjects were computing courses' students and 23 were from a special group. The computing courses' students comprise the following six groups:

- a) Information Systems First Semester (ISF);
- b) Information Systems Last Semester (ISL)
- c) Computer Science First Semester (CSF);
- d) Computer Science Last Semester (CSL);
- e) Software Engineering First Semester (SEF);
- f) Software Engineering Last Semester (SEL);

The seventh group was named "Outstanding Group" (OG). It is a very heterogeneous group formed by non-undergraduate, undergraduate and graduate people. The OG participants are different from the computing courses participants. Included in this group are people that popularly have been excelled, in some way, when compared to their colleagues. For example, they have the highest grades among their classmates, hold a high disputed position job or are known by its colleagues as the "smartest", just to name a few, from others perceptions of how they stands out.

The present authors administered the tests during regular classes. The test administrator presented himself in front of class and rapidly explained the objectives of our research. After that, he distributed the questionnaires to students who had fifty minutes to answer it. The fifty minutes' time was the calibrate amount of time concluded in the end of the pilot experimentation.

During application, if any student had doubts about what was being asked in a question, the test administrator answered him in front of class, to all the classmates to hear, according to the test recommendations. The students responded the test without discussion or sharing answers with classmates. All students finished to answer before the deadline.

Simple scoring was used, following Professor Lawson suggestions. To each correct answer, one point was assigned, ranging from 0 to a maximum of 24 points. According to this scoring system, the higher the score, the higher would be the students' abilities on scientific reasoning (formal reasoning).

3. Results and Data Analysis

The administration of a questionnaire like LCTSR to seven different groups allows the use of diverse analytical techniques. For example, the results can be obtained by analyzing all answers of some group (as a set), or only observing the differences between groups or even considering individual answers.

We highlight in the following subsections a few results we consider important for the objectives of this work. We also emphasize the authors' preference for nonparametric statistical methods because of, among other reasons, their reliance on fewer assumptions, what make them more robust and relatively insensitive to outlying observations [22].

3.1 The Outstanding Group

The Outstanding Group is known by its colleagues as the "smartest", as mentioned before. The need of such group in our work had two main objectives. The first objective was to verify if for the members of this group would be easier to achieve the highest score on the LCTSR. If it had happened, despite the questionnaire has its validity already verified by Lawson and other works, we would have to discard its use.

The results show that although 47.8% of this group are identified as post-formal reasoners and 52.2% as formal operational, only 2 of 23 participants has achieved a perfect score of 24 points. In addition, all participants reported that despite the questions were clearly stated and easy to understand, they were not easy to be solved. Therefore, we consider the instrument and its translation valid, in particular, with respect to the ability to distribute groups of our population in function of their correct answers.

The second objective of the OG was to be a reference to identify if the test is able to differentiate groups of students. In other words, this group is used to verify if the test is sensible enough to distinguish groups regarding their cognitive levels. The OG scores provide an upper limit or a challenging goal to be achieved by cognitive interventions.

3.2 First Semester versus Last Semester

The comparison between groups of first and last semester students from an undergraduate course gives us the possibility to verify whether, on average (as a group), it is possible to note a change in students cognitive ability during the course. The basic hypothesis is that, even as a group, students "should" develop (acquire, improve) their cognitive skills throughout the course.

Comparing all the answers of the groups SEF and SEL using the Wilcoxon-Mann-Whitney Test (WMW) [22], we obtained a p-value = 0.5748 ($W=679$), far above the traditional value of p-value < 0.05, so we can consider that the two groups belong to the same population. This way, when comparing these two groups, we cannot prove a change of cognitive abilities throughout the Software Engineering course. The percentage of individuals in each cognitive level proposed by Lawson, in the respective order (SEF, SEL) was: (22%, 20%) post-formal, (56%, 57%) formal operational and (22%, 23%) transitional.

The same conclusions can be drawn from the Computer Science course results. Comparing all the answers of the groups CSF and CSL, we obtained a p-value = 0.6442 ($W = 679$) and a distribution in the following order (CSF, CSL): (20%, 15%) post-formal, (48%, 55%) formal operational and (33%, 30%) transitional.

For the Information Systems course, the ISF group showed a lower percentage of individuals in the post-formal level (4%), but still we could achieve a p-value higher than 0.05, p-value = 0.0856 ($W = 91$) when compared to the ISL results. These results may be partially explained by the timing that the test was administered to the ISF group, very after another teaching activity of about one hour long. Differently, the test administration to the other groups were performed as the first academic day activity. Therefore it is important to be concerned with the way the test is administered (voluntary, in a relaxed and comfortable environment), the timing (preferably as the first academic activity) and duration (limited, but sufficient to avoid stress). Currently we have scheduled the retest of groups, spending more attention to these observations. The students' distribution following the order (ISF, ISL) was: (4%, 18%) post-formal, (62%, 64%) formal operational and (35%, 18%) transitional.

Despite the common sense, the results of LCTSR application do not permit us to affirm that the first semester and last semester groups of any courses show significant differences in their cognitive abilities.

3.3 Computing Courses

Here, we are interested in conduct comparisons among the three courses: Information Systems, Computer Science and Software Engineering. When comparing all the student's groups of the first semester, we tried to understand whether students who chose different courses would have distinct cognitive abilities. When comparing the student's groups of the last semester, we tried to identify in which cognitive level, on average, each course delivered to society their undergraduate students.

We used the Kruskal-Wallis test (KW) [22]. Thus we can decide whether the population distributions are identical without assuming them to follow the standard distribution. If p-value is less than 0.05, we can affirm that our collection of data samples are independent, so they come from unrelated populations.

Comparing the results of the questionnaire among the first semester students (SEF, CSF and ISF), we obtained a p-value = 0.0957 (KW chi-squared = 4.6937, df = 2). Figure 1 shows the box-plot for the three groups. Clearly the ISF group is responsible for a lower p-value. But still, we cannot claim to be found significant differences among populations regarding their cognitive abilities.

For the last semester students of the three courses (SEL, CSL and ISL) we obtained a p-value = 0.5360 (KW chi-squared = 1.2473, df = 2). Therefore, we cannot affirm that our last semester students, as a group, are concluding their graduation with differences in cognitive abilities in function of their courses. Figure 2 shows the box-plot for the three groups.

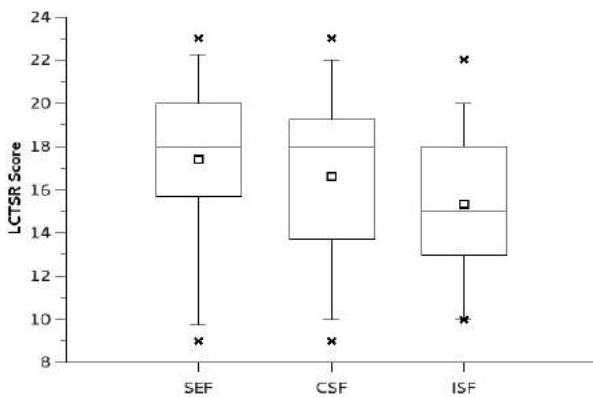


Fig. 1. Box-Plot: First Semester Students Comparison.

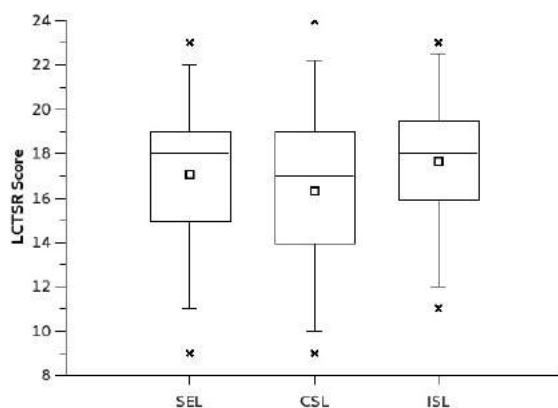


Fig. 2. Box-Plot: Last Semester Students Comparison.

3.4 The Outstanding Group versus Last Semester Groups

Until here, we cannot affirm that first and last semester students of the same course belongs to unrelated populations. Moreover, our results did not reveal cognitive differences among students of different courses. With these results, it is reasonable to question, for example, if the approach of treating students as a group is not very drastic, in the sense that it will be difficult to perform any intervention (conducted in a way that includes and influences all students from a group). The authors are aware of this great challenge, but Lawson [37] presents encouraging results where the interventions were statistically significant in a short period.

The comparison of the OG and the last semester students (SEL, CSL and ISL) shows that as a group, the OG belongs to a different population, with p-value = 0.0007 (KW chi-squared = 16.9068, df = 3). In an interview with members of the OG, it is evident that the vast majority of them assume to have a life story of particular taste for puzzle games and they also feel challenged by new knowledge. The authors believe that perhaps one of our greatest challenges is exactly to be able to provoke moments of new knowledge transmission in conjunction to the construction of cognitive skills. To carry out these interventions, the authors needed an instrument to measure the effectiveness of the intervention and LCTSR seems to be a strong candidate to this purpose.

3.5 Reasoning Patterns

The LCTSR consists of a sequence of questions, where each question or group of questions measure some kind of cognitive ability. Analyzing the mistakes of each researched group, a pattern was identified. Figure 3 shows the relative frequency of wrong answers (*rf*) given by two groups (SEF and SEL) on each question. The *rf* consists of the ratio between the total amount of errors of some question by the number of participants in the questionnaire (in percentage). Based on the *rf* profile it is possible to verify in which reasoning abilities the researched group has greater deficiencies and to devise approaches for an intervention.

The results reveal a pattern between the errors in the pair of questions 11-12 and 13-14. Surprisingly, the same pattern is also recognized in the results of all the other researched groups, including the OG (the pattern is highlighted in Fig. 4).

In order to confirm the unexpected relationship of the *rf* profiles among the researched groups, we computed the Spearman Rank Correlation Coefficient (r_s) [22] between all groups of results. Table 1 shows the values of r_s for each pair. All the table values are higher than the reference value of $r_s = 0.642$, for an alpha < 0.0005 (or 99.9995% of confidence) and $n = 24$. In other words, our results revealed a common error sequence. Based on Lawson cognitive level classification, it was expected that the *rf* would increase as the order of the question increases. This way, it was expected that the *rf* of the last items of the questionnaire would be greater than the *rf* of the first items.

After discovering this error pattern, double checking the translation quality and error counting, we conclude that this is really a significant result that can inspire and guide other researchers. Once LCTSR is proved to be able to identify the reasoning patterns in which students have more trouble, and their relation, the intervention possibilities are numerous.

Analyzing the cognitive skills that the pairs of questions 11-12 and 13-14 measure, according to Lawson [38], it is possible to verify that both pairs measure the "identification and control of variables and probabilistic thinking". Despite each of these skills

were separately assessed by the pairs of questions 9-10 and 15-16, respectively, apparently in these questions they were not so challenging for the most groups (this is confirmed by their lower *rf*). The results suggest that the need to use these two skills at the same time to solve question pairs 11-12 and 13-14, has been somehow a great challenge for our students (confirmed by the higher *rf*).

Once obtained these results, the authors started to develop educational materials in order to conduct a cognitive intervention. With this intervention we expect to help students to improve their cognitive levels. Acting specifically in the reasoning skills that LCTSR detected the higher *rf*, we hope to reduce the difficulties presented, improving the students success rates on courses.

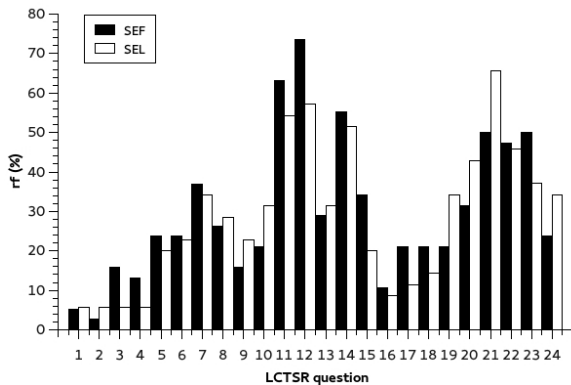


Fig. 3. Relative Frequency of Wrong Answers (SEF and SEL groups).

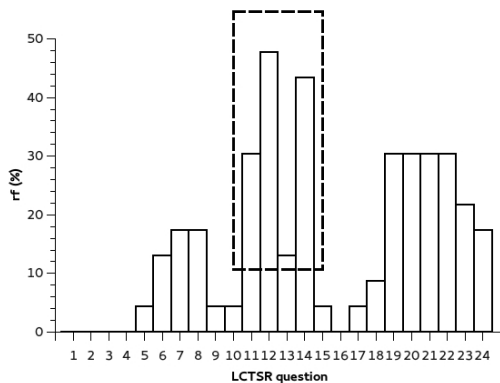


Fig. 4. Relative Frequency of Wrong Answers (OG group).

Table 1. Spearman Rank Correlation Coefficients

	CSF	CSL	SEF	SEL	ISF	ISL
OG	0.797	0.876	0.848	0.933	0.786	0.775
CSF	-	0.882	0.815	0.768	0.698	0.686
CSL		-	0.891	0.862	0.866	0.794
SEF			-	0.873	0.925	0.891
SEL				-	0.819	0.854
ISF					-	0.857

4. CONCLUSIONS

It is not difficult to find colleagues who, like the authors, believe that when students acquire new knowledge related to their expertise area, their cognitive levels change, in other words, increase with it. Lawson [39, 40] helps us with the differences between declarative knowledge and procedural knowledge:

“The word 'learning' is often used in conjunction with the acquisition of declarative knowledge. ...scientific reasoning consists of an overall pattern of reasoning, which can be characterized as hypothetic-deductive, as well as several sub-patterns the very process of generating and testing ideas results not only in the construction of ideas that work (i.e., the learning of useful declarative knowledge), but also in improved skill in learning (i.e., the development of improved procedural knowledge). Thus, to teach most effectively, teachers should allow their students to participate in the idea generation and testing.”

The present authors believe that the first step in order to develop strategies to improve teaching and learning processes would be the identification of instruments for cognitive processes assessment with a strong background theory.

In this work, it was found numerous methods, created through the years, in diverse countries and in many languages. As each one claims to be the best alternative, there was a need to organize them in a systematic way, and apply some criteria to help to decide which one to use. Of the many reviewed methods, ATFR, BLOT e LCTSR demonstrated to be strong candidates. None of them had Portuguese versions, this way, LCTSR was selected to be the first translated and administered to computing courses alumni (Information Systems, Computer Science and Software Engineering).

LCTSR was also administered to a special group in order to test the scale quality of the instrument and to verify if it would be sensible enough to distinguish students groups. The questionnaire and its translation proved to be valid for the initial objectives of the research, offering a good option to the scientific community.

The results were extremely encouraging, providing much more than just the classification of students in levels of cognitive ability. They allowed some initial work hypotheses to be tested, like the variation of cognitive abilities throughout the courses and the existence of differences among first and last semester students between courses. Based on our results, we cannot affirm that exist differences between first and last semester students of some course, neither of first and last semester among courses. The results also indicated which reasoning skills need more attention when conducting an intervention in order to be more efficient in the teaching process.

The present research brings uniqueness in two ways: 1. the first Brazilian Portuguese version of LCTSR is the one translated by the present authors; 2. the first application data and results of this test in our country.

Our future steps are: expand the application of the test, monitor over time the groups already tested and propose, according to the background theory, possible interventions.

5. ACKNOWLEDGMENTS

We would like to thank Professor Anton E. Lawson for providing the permission to translate and use his method. Professor Trevor G. Bond for his sympathy and professional cordiality on having shared part of his work. We would also like to thank Mr Fabrício Vieira Personi for his translation services and Miss Nayara Cristina Rodrigues de Andrade for her flawless revision.

6. REFERENCES

[1] Arlin, P. K. 1982. A multitrait-multimethod validity study of a test of formal Reasoning. *Educational and Psychological Measurement*. 42 (1982), 1077 – 1088.
 [2] Ates, S. Cataloglu, E. 2007. The effects of students' reasoning abilities on conceptual understandings and

- problem-solving skills in introductory mechanics. *Eur. J. Phys.* 28 (2007), 1161–1171.
- [3] Barrows, H. S. Problem-based learning in medicine and beyond: A brief overview. *New directions for teaching and learning*. 68 (1996), 3-12.
- [4] Bart, W. M. 1972. Construction and Validation of Formal Reasoning Instruments. *Psychological Reports*, 30 (1972), 663 – 670.
- [5] Beaton, D. E. et al. Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *SPINE* 25, 24. (2000), 3186 –3191.
- [6] Bergin, S. Reilly, R. 2005. Programming: factors that influence success. In *Proceedings of the 36th SIGCSE technical symposium on Computer science education (SIGCSE '05)*. ACM, New York, NY, USA, 411-415.
- [7] Bond, T. G. 1976. BLOT: Bond's Logical Operations Test. Townsville, Australia. James Cook University.
- [8] Bond, T. G. Piaget and measurement I: The twin really do meet. *Archives de Psychologie*. 63, 245 (1995), 71-87.
- [9] Briggs, M. I. McCaulley, M. H. Quenk, N. Hammer, A. 1998. *MBTI Handbook: A Guide to the development and use of the Myers-Briggs Type Indicator* Consulting (3rd edition). Psychologists Press.
- [10] Burney, G. M. 1974. *The construction and validation of an objective formal reasoning instrument*. Thesis. Division of Arts and Science., Uni. of Northern Colorado., Greeley., CO.
- [11] Byrne, J. H. Ph.D. 2014. Neuroscience Online. Department of Neurobiology and Anatomy, The UT Medical School at Houston. Retrieved September 4, 2014 from <http://neuroscience.uth.tmc.edu/>
- [12] Carlson, G. R. Streitberger, E. 1983. The Construction and Comparison of Three Related Tests of Formal Reasoning. *Science Education*. 67, 1 (1983), 133-140.
- [13] Carter, K. R. Ormrod, J. E. 1983. *The Social Sciences Piagetian Inventory (rev. edn.)*. Uni. of Northern Colorado, Dep. Ed. Psychology. Greeley, CO.
- [14] Crouch, C. H. Mazur, E. 2001. Peer instruction: Ten years of experience and results. *American Journal of Physics*. 69, 9 (2001), 970-977.
- [15] Digiampietri, A. L. 2012. Complementando o Aprendizado em Programação: Experiências no Curso de Sistemas de Informação da USP. In *VIII Simpósio Brasileiro de Sistemas de Informação* (São Paulo, SP, Brazil 2012). SBSI, 468-479.
- [16] Eckstein, S. G. and M. Shemesh. 1992. The rate of acquisition of formal operational schemata in adolescence: A secondary analysis. *Journal of Research in Science Teaching* 29, 5 (1992), 441-451.
- [17] Felder, R. M. Soloman, B. A. 2014. Learning styles and strategies. North Carolina State University. Retrieved September 4, 2014 from <http://www.ncsu.edu/felder-public/ILSdir/styles.htm>
- [18] Felder, R. M. Spurlin, J. 2005. Applications, reliability and validity of the index of learning styles. *International journal of engineering education* 21, 1 (2005), 103-112.
- [19] Furth, H. 1970. *An Inventory of Piaget's Developmental Tasks*. Thesis. Dept. of Psychology, Catholic Univ, Washington, DC.
- [20] Han, J. 2013. *Scientific Reasoning: Research, Development, and Assessment*. Doctoral Thesis. School of The Ohio State University. Columbus, OH.
- [21] Hmelo-Silver, C. E. 2004. Problem-based learning: What and how do students learn?. *Educational Psychology Review*. 16, 3 (2004), 235-266.
- [22] Hollander, M. Wolfe, D. A. Chicken, E. 2013. *Nonparametric statistical methods* (2nd. ed.). John Wiley & Sons, Danvers, MA.
- [23] Honzel, B. 2012. *The Effects of an Inquiry Based Data to Concept Curriculum*. Thesis. Montana State University.
- [24] Inhelder, B. Piaget, J. 1958. *The Growth of Logical Thinking From Childhood to Adolescence*. Basic Books Inc, New York, NY.
- [25] Ippolito, K. O. 2012. *Cognitive development and the attainment of critical thinking skills in associate degree nursing students*. Thesis. University of the Pacific Stockton, CA.
- [26] Kalat, J. W. 2008. Nature, Nurture, and Human Development. In *Introduction To Psychology* (8th. ed.). ch. 5, sec. 5.2, pp. 163 – 181. Thomson Wadsworth, Belmont, CA.
- [27] Kayes, D. C. 1999. Internal validity and reliability of Kolb's learning style inventory version 3. *Journal of Business and Psychology*. 20, 2 (Winter 2005), 249-257.
- [28] Kinnunen, P. Meisalo, V. Malmi, L. 2010. Have we missed something?: identifying missing types of research in computing education. In *Proceedings of the Sixth Int. Workshop on Computing education research* (New York, NY, USA 2010). ICER '10. ACM, New York, NY, 13-22.
- [29] Kitchenham, B. et al. 2007. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. Technical Report. School of Computer Science and Mathematics at Keele University. Department of Computer Science at University of Durham. Durham, UK.
- [30] Kolb, D. A. 1984. *Experiential learning: Experience as the source of learning and development*. (Vol. 1). Prentice-Hall, Englewood Cliffs, NJ.
- [31] Krathwohl, D. R. 2012. A revision of Bloom's taxonomy: An overview. *Theory Into Practice*. Routledge. 41, 4 (Aut. 2002), 212–218.
- [32] Kuhn, D. 2008. Formal Operations from a Twenty-First Century Perspective. *Human Dev.*, 51, 1(2008), 48–55.
- [33] Kurtz, B. L. 1980. Investigating the relationship between the development of abstract reasoning and performance in an introductory programming class. In *Proceedings of the eleventh SIGCSE technical symposium on Computer science education*. ACM. Kansas City, Missouri, USA. 110-117.
- [34] Lawson, A. 1978. The Development and Validation of a Classroom Test of Formal Reasoning. *Journal of Research in Science Teaching*. 15 (1978), 11 – 24.
- [35] Lawson, A. E. 2000. *Lawson Classroom Test of Scientific Reasoning*. Personal Communication. E-mail received from the author Feb. 2014.

- [36] Lawson, A. 2010. *Teaching Inquiry Science in Middle and Secondary Schools*. Sage, Publications, Inc. Thousand Oaks, CA.
- [37] Lawson, A. E., et al. 2000. What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *Journal of Research in Science Teaching*. 37, 9 (2000), 996-1018.
- [38] Lawson, A. E. *Keys to Lawson Classroom Test of Scientific Reasoning*, 2000 version. Personal Communication. E-mail received from the author Feb. 2014.
- [39] Lawson, A. E. 2004. The nature and development of scientific reasoning: A synthetic view. *Int. Journal of Science and Mathematics Education*. 2, 3 (2004), 307-338.
- [40] Lawson, Anton. E. 2003. *The neurological basis of learning, development and discovery: Implications for science and mathematics instruction*. Springer, Kluwer Academic Publishers, Dordrecht.
- [41] Longeot, F. 1966. L'échelle de développement de la pensée logique. BINOP, Paris.
- [42] Logan, R. A. 1991. *Relationships between the Arlin Test of Formal Reasoning and Classroom Multiple Choice*. Thesis. Emporia State University.
- [43] Lippman, J. 2012. *Improving and predicting novice reasoning about the evidentiary connection between studies and theories*. Doctoral Thesis. US, ProQuest Information & Learning. University of Illinois at Chicago, Chicago, IL.
- [44] Lunzer, E. A. 1965. Problems of Formal Reasoning in test situations. *Monographs of the Society for Research in Child Development*. 30, 2 (1965), 19 – 46.
- [45] Malmi, L. et al. 2014. Theoretical underpinnings of computing education research: what is the evidence?. In *Proceedings of the tenth annual conference on International computing education research*. (Glasgow, Scotland, Aug 2014). ICER '14. ACM, New York, NY, USA, 27-34.
- [46] Malmi, L. Sheard, J. Simon, Bednarik, R. Helminen, J. Korhonen, A. Myller, N. Sorva, J. Taherkhani, A. 2010. Characterizing research in computing education: a preliminary analysis of the literature. In *Proc. of the Sixth Int. workshop on Computing education research* (Aarhus, Denmark 2010). ICER '10. ACM, New York, NY, 3-12.
- [47] Manhães, M. B. L. et al. 2012. Identificação dos Fatores que Influenciam a Evasão em Cursos de Graduação Através de Sistemas Baseados em Mineração de Dados: Uma Abordagem Quantitativa. In *VIII Simpósio Brasileiro de Sis. de Informação SBSI* (São Paulo, SP 2012), SBC, 468-479.
- [48] Mazur, E. 1997. *Peer Instruction: A User's Manual*. Prentice Hall, Upper Saddle River, NJ.
- [49] Mozzachiodi, R. Byrne, J. H. 2010. More than synaptic plasticity: role of nonsynaptic plasticity in learning and memory. *Trends in neurosciences*. 33, 1 (Jan. 2010), 17-26.
- [50] Parham, J. R. 2003. An assessment and evaluation of computer science education. *J. Comput. Sci. Coll.* 19, 2 (December 2003), 115-127.
- [51] Pears, A. Malmi, L. 2009. Values and objectives in computing education research. *ACM Trans. Comput. Educ.* 9, 3, Article 15 (September 2009), 6 pages.
- [52] Pears, A. Seidman, S. Malmi, L., et al. 2007. A survey of literature on the teaching of introductory programming. In *Working group reports on ITiCSE on Innovation and technology in computer science education* (2007). ITiCSE-WGR '07. ACM, New York, NY, USA, 204-223.
- [53] Piaget, J. 1957. *Logic and Psychology*. New York: Basic Books Inc, New York, NY.
- [54] Randolph, J. J. 2007. *Computer Science Education Research at the Crossroads: A Methodological Review of Computer Science Education Research: 2000-2005*. Doctoral Thesis. Utah State University. Logan, UT.
- [55] Raven, R. J. 1973. The Development of a Test of Piaget's Logical Operations. *Science Education*, 57,3 (Jul/Sep 1973) 377 – 385.
- [56] Roadranga, V. 1985. *The construction and validation of the group assessment of logical thinking (GALT)*. Thesis, Uni. Of Georgia, Athens, GA.
- [57] Roberge, J. J. Flexer, B. K. 1982. The Formal Operational Reasoning Test. *The Journal of General Psychology*. 106, 1 (1982), 61-67.
- [58] Rowell, J. A. Hoffman, P. J. 1957. Group tests for distinguishing formal from concrete thinkers. *Journal of Research in Science Teaching*, 12 (1975), 157 – 164.
- [59] Russel, H. 1962. The Measurement of Reasoning Ability in Adolescents. *Ontario Journal of Educational Research*. 5, 1(1962), 33 – 40.
- [60] Shayer, M. Adey, P. Wylam, H. 1981. Group Tests of Cognitive Development Ideals and a Realization. *Journal of Research in Science Teaching*. 18, 2 (1981), 157 – 168.
- [61] Simons, D.J. Chabris, C.F. 2011. What people believe about how memory works: A representative survey of the US population. *PloS one*. 6,8, e22757 (Aug. 2011).
- [62] Sousa, V. D. Rojjanasrirat, W. 2011. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*. 17 (2011), 268–274.
- [63] Tisher, R. P. 1971. A Piagetian questionnaire applied to pupils in a secondary school. *Child Development*. 42, 5 (Nov 1971), 1633 – 1636.
- [64] Tobin, K. G. Capie, W. 1981. The Test of Logical Thinking (TOLT): The Development and Validation of a Group Test of Logical Thinking. *Educational and Psychological Measurement*, 41, 2 (Jul 1981), 413 – 423.
- [65] Tomlinson-Keasey, C. 1976. Can we develop abstract thought in college freshmen? *84th Annual American Psychological Association Conv.* (1976). Washington, D.C.
- [66] Veenman, M. V. J. Van Hout-Wolters, B. H. A. M. Afflerbach, P. 2006. Metacognition and learning: Conceptual and methodological considerations. *Metacognition and learning*. 1, 1 (Apr. 2006), 3–14.