Measuring Cognitive Loads while Learning Computational Statistics

1st Roxana Quintero-Manes
Departament of Education
Universidad del Norte
Barranquilla,Colombia
rmanes@uninorte.edu.co

2nd Camilo Vieira
Departament of Education
Universidad del Norte
Barranquilla, Colombia
cvieira@uninorte.edu.co

3rd Natalia Hernandez-Vargas
Mathematics and Statistics
Universidad del Norte
Barranquilla, Colombia
nataliah@uninorte.edu.co

***Abstract*— This work-in-progress paper is part of a project that aims to analyze how the implementation of different instructional designs through complete and incomplete examples influence the cognitive load that students perceive when learning statistics and computer programming. For this first stage, we conducted a pilot study aimed at conducting a differentiated measurement of cognitive loads through secondary tasks and a subjective scale. The pilot study consisted of three moments. The first one focused on measuring their working memory capacity and their prior knowledge of statistics. Then, students worked on a learning activity focused on using Chi-square and Pearson correlation to conduct inferential analyses. The study included two different instructional designs in the form of computational notebooks. Design 1 consisted of correct and incomplete examples, while Design 2 included correct, incomplete, and incorrect examples. Simultaneously, students had to attend to a secondary task, which consisted of responding to an auditory stimulus, after which they had to press a key on their keyboard. Finally, students completed a naïve rating scale of cognitive loads involved in the task. The results highlight the important role of prior knowledge in different instructional designs and how it connects to students' perceived cognitive effort. Likewise, the differences in the instructional designs influenced students' cognitive loads, suggesting a differentiated measure of extraneous loads. Using incorrect examples in learning statistics and programming may affect students' perceived difficulty without resulting in a better learning outcome.**

Keywords—Cognitive load, Computational Statistics, Computational Thinking, Subjective Scale, Secondary Task, Working Memory.

I. INTRODUCTION

Multiple factors can influence the quality of the learning process. The How People Learn framework discussed that learning environments should be learner-centered, assessment-centered, knowledge-centered and community-centered [1]. Using learning activities and materials that do not consider the unique characteristics of the knowledge to be learned, may expose students to overwhelming experiences.

For instance, computational statistics can become a challenge for students who are not used to programming or working with numbers. Students in the social sciences do not often find courses where they must become familiar with quantitative data analysis or computer programming. However, these programs are starting to realize the need to prepare students with the numeracy and computational thinking skills to understand social phenomena. The learning of computational statistics is increasingly present in the classrooms, including in engineering, psychology, and political science programs.

Hence, it is important to understand how learning occurs in these contexts and what cognitive loads are involved in this process. This understanding may enable researchers and educators to design and implement effective learning environments for these complex topics. This work-in-progress paper will explore students' cognitive loads while learning computational statistics using two instructional approaches: (1) correct and incomplete examples; and (2) correct, incomplete, and incorrect examples. The guiding research question for this study is:

How do a secondary task and a subjective scale measure the extraneous, germane, and intrinsic cognitive loads, and how do these relate to students' prior knowledge and learning outcomes?

II. THEORETICAL FRAMEWORK

A. Cognitive load theory

Initially developed by Sweller, the cognitive load theory describes a cognitive architecture where the working memory capacity is limited. This theory defines the cognitive architecture as the structure of cognitive functioning that allows humans to develop different mental processes, including learning [2, 3]

According to this theory, the cognitive architecture comprises a working memory and a long-term memory, which process incoming information using existing knowledge when learning occurs. The long-term memory is organized as schemata consisting of groups of concepts that are related to each other. These schemata are strengthened to the extent that people make sense of and practice new knowledge. In this process, the learner must attend to the new information, manipulate it mentally and give it meaning; when this happens, the received information for the first time is retained and can begin to be stored in the long-term memory [2, 3, 4, 5].

When a human is processing new information and manipulating it mentally, s/he can hold about four to seven chunks of information in the working memory. Thus, if s/he needs to attend and remember more information, s/he must use strategies such as elaboration that allow them to group or associate these chunks of information [6].

Cognitive loads refer to a person's mental effort in learning. A person will have a higher or lower cognitive load according to how challenging the task is, their prior knowledge, and the learning materials. For Sweller, there are three types of cognitive loads: intrinsic, extraneous, and germane [5].

The intrinsic cognitive load is related to the complexity of the subject, concepts, or skills. If the content itself has several interactive elements that need to be learned simultaneously, it represents complex learning. For instance, learning to conduct and interpret a statistical analysis using R programming requires the learner to recognize whether specific assumptions are met, decide which statistical analysis to use, and how to implement it in R. The extraneous cognitive load is related to the learning design and how much effort a student needs to devote to figure it out. Thus, the extraneous load is affected by the amount of information, representations, examples in the learning activity, the organization of the material, etc. Therefore, the extraneous load depends more on the instructional design and does not contribute to learning. Finally, the germane or relevant load is beneficial to learning, where the learners load existing knowledge to make sense of new information [2, 3, 4, 5].

Consequently, instructional designers must consider the characteristics of the cognitive architecture and how it influences the amount of information that students can process. If such an amount is exceeded, the learning task may overwhelm students and negatively affect their learning. Thus, it is becoming increasingly relevant that educators, when designing learning environments, should be able to recognize whether the selected structure, resources, and strategies are adequate and adjusted to the subject matter.

B. Measurement of cognitive loads

The cognitive loads a student experiences may be measured indirectly through three variables: mental load, performance, and mental effort. The mental load refers to the difficulty and complexity of a task. This complexity relates to the number of elements that make up the task. The task is difficult if there are many elements, but these can be learned independently from each other (e.g., a new vocabulary in second-language learning). The task is complex if these elements interact with each other during learning (e.g., learning programming). These two aspects (i.e., difficulty and complexity) influence students' mental load while engaged in the task. Thus, increasing the number of interrelated components increases the required mental load to solve the activity [7]

The mental effort refers to the relationship between the task being executed and the characteristics of the individual. Therefore, the mental effort can be differentiated among different people since it depends on the prior knowledge of the subject. The prior knowledge enables them to successfully complete the task more or less easily [7]

Finally, performance refers to the measurement carried out to identify the number of errors and successes in a learning activity. Thus, performance is similar to mental effort as the learner's prior knowledge and skills may influence the outcome.

The importance of understanding the cognitive loads involved in a learning process has led to the study of their measurement. Over the last thirty years, several researchers have developed different techniques to measure them. Some of the most widely used include self-reports or subjective scales and secondary tasks. Subjective scales seek to identify the learner's perception of the mental load devoted while working on a given task [7]. The secondary tasks involve an additional task to the learning activity and can be implemented using two different approaches. The first approach applies a secondary task immediately after the primary task and measures whether the reaction time is much higher than regular, considering the previous effort. The second approach implements the secondary task at the same time as the primary task, and measures either the number of times the student responds to the secondary task or the reaction time [8, 9, 10]. Yet another approach for cognitive load measurement involves assessing the learner's physiological responses, such as pupil dilation [7].

C. The Learning Design

The development of the cognitive load theory has also allowed studying the effects of instructional designs on students' cognitive load. These advances have allowed recognition of the effect of cognitive loads on student learning and the differences they can have if implemented with expert or novice learners. Some of the described effects include the goal-free effect, the worked example effect, the fading effect, the redundancy principle, the modality effect, the self-explanation effect, etc. [3, 4].

For this work-in-progress paper, we used complete worked-examples (i.e., guided by the worked-example effect), incomplete worked-examples, and incorrect worked-examples (i.e., guided by the fading effect). The complete worked-examples refer to examples with an expert's complete solution to an exercise or problem, accompanied by annotations, or auxiliary representations that support student understanding of the learning activity. The incomplete worked-example shows a partial solution of the example, and the student needs to identify what to do in order to complete the solution. Finally, the incorrect worked-example shows a solution that includes one or more bugs or mistakes, and the learner needs to identify the issue and figure out how to solve it. [3, 4].

Integrating computational science practices into disciplinary contexts, such as the social sciences, generates a cognitive overload in novice learners[11]. Thus, it is essential to identify which learning activities may reduce students' cognitive loads and support their learning process. This study aims to assess student cognitive loads during a computational statistics activity and identify to what extent the integration of different worked-examples with self-explanation activities influences the students' cognitive load.

Identify applicable funding agency here. If none, delete this text box.

III. METHODS

A. Participants

Four students enrolled in an inferential statistics course participated in this study. In this course, students learn statistical data analysis techniques and implement these using basic elements of R programming. Participation in this pilot was voluntary, and all four participants were women.

B. Procedures

The procedures for this pilot study comprised three stages. The first stage focused on measuring their working memory capacity. For this purpose, we used a subscale of the intelligence test WAIS III [12]. We also measured their prior knowledge of statistics using a test that the first author designed, and authors two and three (experts in statistics and education) validated. After completing these tests, the participating students worked on a learning activity focused on using Chi-square and Pearson correlation for conducting data analysis. The study included two different instructional designs in the form of computational notebooks. Design 1 included both correct and incomplete worked-examples. Design 2 included correct, incomplete and incorrect worked-examples. The students identified as S1 and S3 worked on the activities of Design 1, while students S2 and S4 completed the learning materials from Design 2. While working on these learning activities, the students simultaneously attended to a secondary task. The secondary task required them to respond to an auditory stimulus, after which they had to press a key on their keyboard. We used the number of times the sound was emitted and the number of times they responded to it to compute their score. After completing the learning activity, we measured students' cognitive load using a self-report scale. This scale was translated into Spanish, adapted from English, and empirically validated with a large sample size. This scale consists of three subscales, each assessing a type of cognitive load (i.e., intrinsic, extraneous, and germane) [7].

IV. RESULTS AND DISCUSSION

Table I presents a summary of the results for each of the participating students, including: their working memory capacity, their level of prior knowledge before the task, their performance on the secondary task, their intrinsic, extraneous, and germane cognitive loads in task, and the assessment of their understanding of the concepts at the end of the activity.

The results suggest a similar working memory capacity among students S2, S3, and S4, while student S1 obtained a slightly lower score. Students S1 and S3 answered all the questions correctly when assessing their prior knowledge, while student S4 had the lowest score. As for their performance in the secondary task, student S1 obtained the lowest score. While S1 was working on the task, the interviewer identified that, after a few minutes, she stopped paying attention to the secondary task and focused instead exclusively on the main task; she stopped responding to the auditory stimulus. When asked about the reasons for this behavior, she expressed that she preferred to have all her attention on the main task, especially when performing the activities in R.

For the cognitive load scales, the evaluation of the perceived extraneous load was low for all four students, while the germane load was high for students S1 and S4 (i..e., above 90/100) and slightly lower for S3 compared to the others. The intrinsic cognitive load was very low for student S1, while the rest of the participants rated it with similar scores. Finally, in the performance task at the end of the activity, S1 had the lowest score (50), and S2 showed the highest score (100).

TABLE I.

	W.M[a]	P. K[b]	S. T[c]	E. L[d]	G. L[e]	I. L[f]	E.K[g]
S 1	8	90	16,13	0,00	94,58	8,54	50
S 2	13	66	89,00	0,00	89,55	49,10	100
S 3	12	100	70,83	33,27	78,31	40,83	75
S 4	11	33	80,46	24,90	94,58	55,02	75

[a]Working Memory. [b]Prior knowlege [c]Secondary Task, [d]Extraneous load, [e]Germane load [f]Intrinsic load and [g]Evaluation knowlege

When analyzing the results as a whole, it was noted that participant S1 had the lowest score in working memory, "abandoned" the secondary task, perceived the lowest intrinsic load compared to the other participants, and obtained one of the highest scores in germane cognitive load. These results are consistent given that having a lower memory capacity and trying to respond to both tasks simultaneously could generate a cognitive overload. This process ended up affecting the participant's ability and intention to continue responding to the secondary task. Instead, she focused on the learning strategies needed to understand and learn about Chi-square and Pearson correlation. Still, S1 had the lowest performance at the end of the activity compared to the rest of the participants.

In the case of student S2, she obtained the highest score in working memory capacity among the participants and was the one who got the highest score in the secondary task. She also showed a perfect score in the performance activity after engaging with the examples. Meanwhile, student S3 demonstrated a clear understanding during the prior knowledge task. She is the one who perceives a lower germane cognitive load, an indicator of how previous knowledge serves as a basis for the construction of new knowledge and the implementation of learning strategies. She also rated the task to have a higher extraneous cognitive load than the rest of the participants. These results are aligned to the cognitive load theory, where the types of loads are associated with each other: the higher the perceived extraneous load, the lower the cognitive capacity can be directed to the relevant cognitive load; thus, a cognitive overload may affect learning [2, 3, 4, 5].

Finally, student S4 obtained the lowest performance in prior knowledge, and her perception of the cognitive load was high in the germane and intrinsic loads. This shows how prior knowledge may influence the perceived effort of students associated with the intrinsic load when learning a new topic.

Furthermore, when considering the type of instructional design each participant engaged with, the germane and intrinsic cognitive loads were greater in those who engaged with complete, incomplete, and incorrect examples. This result could indicate a potential cognitive overload when having to identify errors, as this requires higher cognitive and metacognitive skills.

The results reiterate the relevance of prior knowledge for the learning strategies and students' perceived cognitive effort. Likewise, some of the differences in the instructional designs influenced students' cognitive loads. Using incorrect examples in learning statistics and programming may affect the student's perceived difficulty without implying a better learning outcome [13, 14].

V. CONCLUSIONS AND NEXT STEPS

This study explored a differentiated measurement of cognitive loads through secondary tasks and a subjective scale when learning statistics and computer programming with two instructional designs. Sometimes it is considered that more challenging learning will make students work harder to understand, and will then allow them better results, greater understanding, and higher quality learning. However, a greater mental effort will not necessarily indicate better results, especially for novice learners.

This pilot study shows the relationship of work memory capacity with the cognitive load while learning and how it influences the perception of difficulty and cognitive effort that students experience. However, this small sample size only shows some emergent patterns in terms of cognitive loads and performance. Additional work is needed to better understand how computational notebooks for statistics interact with students' cognitive loads. In the context of computational statistics, some elements to be considered include students' interest in data analysis and the characteristics of the programming language they use. The next steps of this process involve implementing the learning activity in a classroom environment with a larger sample size, and understanding how this works in real-world settings. Moreover, one of the questions for future work includes how novices and experts experience these tasks in terms of cognitive loads.

REFERENCES

[1] J. Bransford, A. Brown, and R. Cocking. *How people learn*. Vol. 11. Washington, DC: National academy press, 2000.

[2] J. Sweller, R. Mawer and W. Howe. Consequences of History-Cued and Means-End Strategies in Problem. In The American Journal of Psychology. Vol. XCV, 1982, pp. 455-483.

[3] J. Sweller, R. Mawer and M. Ward, (1983). Development of Expertise in Mathematical Problem Solving. In Journal of Experimental Psychology, vol. CXII, 1983, pp. 639-661.

[4] J. Sweller, J. Merriënboer, and F. Paas. Cognitive Architecture and Instructional Design: 20 Years Later. In Educational Psychology Review, vol. XXI, 2019, pp. 261-292.

[5] J. Leppink, F. Paas, C., T. van Gog and C. van der Vleuten. J. Merriënboer. Effects of pairs of problems and examples on task performance and different types of cognitive load. In Learning and Instruction , vol. XXX, 2014, pp. 32-42

[6] N. Cowan, The Magical Mystery Four: ¿How Is Working Memory Capacity Limited, and Why?, in Current directions in psychological science, vol. XIX, 2010, pp. 51–57.

[7] M. Klepsch, F. Schmitz, and T. Seufert. Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. In Front. Psychol, vol. XVI, 2017, pp. 1 – 18.

[8] B. Park, and R. Brunker. The rhythm method: a new method for measuring cognitive load—an experimental dual task study. In Applied Cognitive Psychology, vol. XXIX, 2015, pp. 232-243.

[9] J. Leppink, F. Paas, C. van der Vleuten. T, van Gog and J. Merriënboer. Development of an instrument for measuring different types of cognitive load. In Behav. Res. Methods, vol. XLV, 2013, pp. 1058-1072.

[10] F. Pass and J. Van Merrinboer. Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. In Journal of Educational Psychology, vol. LXXXVI, 1994, pp. 122-133

[11] C. Vieira, A. J. Magana, M. L. Falk, and R. E. Garcia. Writing in-code comments to self-explain in computational science and engineering education. ACM Transactions on Computing Education (TOCE), vol, XVII, 2017, pp. 1-21.

[12] D. Wechsler. Adult Intelligence Scale-Third Edition (WAIS-III)administration and scoring manual. San Antonio: The Psychological Corporation, 1997.

[13] D. Shaffer, W. Doube and J. Tuovinen. Applying Cognitive Load Theory to Computer Science Education. In the 15th Workshop of the Psychology of Programming Interest Group.

[14] J. Van Merriënboer, J. Kester, J. and F. Pass. Teaching complex rather than simple tasks: balancing intrinsic and germane load to enhance transfer of learning. In Applied Cognitive Psychology, vol. XX, 2006, pp. 343-352.