Pattern Recognition in Computing Education: A Systematic Review *

Braz Araujo da Silva Junior[1], Júlia Veiga da Silva, Simone André da Costa Cavalheiro[1], Luciana Foss[1]

[1]Programa de Pós-Graduação em Computação - Universidade Federal de Pelotas
CEP 96.010-610 - Pelotas - RS - Brazil

{badsjunior,jvsilva,simone.costa,lfoss}@inf.ufpel.edu.br

Abstract. *This paper presents a systematic literature review to investigate how pattern recognition has been approached in computing education. Pattern Recognition has long been an important concept among various areas, from cognitive psychology and neuroscience to machine learning and computer vision. Recently, it gained the attention of education, being associated with computational thinking. An old concept being revisited in a new context raises important questions on how it is being approached, how it is being assessed and if it is making use of previous contributions from the other fields that have been studying the concept. This work systematically reviews the literature to answer these questions on pattern recognition. It is found that, as other concepts related with computational thinking, it is rarely treated alone and when among others, few studies have pattern recognition as the main theme. The results also show that no standardized assessment method is used and contributions from other fields are barely mentioned.*

1. Introduction

The concept of **Pattern Recognition (PR**) is studied across multiples areas of knowledge. Generally, PR refers to "a process of inputting stimulating (pattern) information and matching with the information in long-term memory, then recognizing the category which the stimulation belongs to" [Pi et al. 2008]. In its broadest sense, PR may be considered "the heart of all scientific inquiry, including understanding ourselves and the real-world around us" [Liu et al. 2006]. It is considered fundamental to both human cognition [Reed 1972] and machine intelligence [Bishop and Nasrabadi 2006], holding a huge variety of applications, such as: computer vision; computer aided diagnosis; character recognition; speech recognition; face recognition; fingertip identification; astronomical telescope image analysis; DNA sequence analysis; soil evaluation; earthquake analysis; and fault diagnosis [Liu et al. 2006].

Therefore, it has been explored especially by cognitive sciences [Eysenck and Keane 2015], machine learning [Bishop and Nasrabadi 2006] and mathematics [Mulligan and Mitchelmore 2009]. And it is now entering a new area: education, particularly computing education, where it appears as one of the Computational Thinking (CT) competences [Selby and Woollard 2013]. The idea of

*This work was carried out with the support of CAPES - Brazil - Financing Code 001, the SACCI Network, SMED/Pelotas, PREC and PRPPG / UFPel.

DOI: 10.5753/sbie.2022.225128

CT is that of a problem solving mental process involving several competences related to computing [Wing 2006]. Which has being one of the hottest trends in computing education lately [Saqr et al. 2021]. This means PR is now being revisited in this new context, what lead us to the following research questions that drives this paper:

1. How is Pattern Recognition being approached in Computing Education?
2. How is Pattern Recognition being assessed in Computing Education?
3. What contributions to Patter Recognition research from other fields are mentioned or discussed in Computing Education?

The rest of this paper is organized as follows. The Section 2 discusses related work, situating the paper within PR in CS education. The section 3 exposes the methods we followed to conduct the SLR. The section 4 presents our findings. The section 5 concludes the paper with suggestions for future research.

2. Related Work

As previously mentioned, PR has been fairly associated to CT and there are several recent SLR on CT. They explore the concept both in STEM [Wang et al. 2021] and by itself [Acevedo-Borrega et al. 2022]. Although, most of them doesn't mention PR beyond CT definitions, let alone explores how specifically PR is being approached and/or assessed. Some CT SLR [Tang et al. 2020] do discuss how CT is assessed, but it is not possible to retrieve from their results particularly how PR is assessed. Thus, we intend to fill this gap of the literature treating PR, not CT.

One study [Dasgupta and Purzer 2016] followed a similar path, i.e. targeted specifically PR, but found only two studies. At first, we assumed it was due to methodological limitations of their review, which was restricted to peer-reviewed journal articles found in one engineering education research database: the Engineering Village. Then, we intended to extend their review using other STEM and psychology related databases (such as ERIC) as they suggested for future research, as well as considering papers published in conferences. Noteworthy, their review was published 8 years ago, so an update to check if the literature advanced in this specific topic is necessary. Moreover, their findings brought a position paper [Tanimoto 1998] discussing how PR and computer vision are important for mathematics education in middle school; and a medical sciences paper [Kellman and Garrigan 2009] identifying PR as a relevant skill for medical students. No information about how PR is being approached nor assessed is given in such papers.

3. Methods

We followed a guideline proposed for systematic reviews in software engineering researches [Kitchenham 2004], which proposes three main phases: Planning the Review, including the identification of the need for a review and the development of a review protocol; Conducting the Review, including selection of studies, data extraction and synthesis; and Reporting the Review, which we followed PRISMA [Page et al. 2021] guidelines to do so.

The review protocol was built based on the research questions presented in the introduction. The chosen databases and search engines were aligned with the suggestion of previous literature, targeting those specialized in education/psychology (ERIC) and computing/STEM (IEEE Xplore). In order to capture a larger number of studies without going out of the scope of the work, we developed the following search string using synonyms for PR and words related to computing education: ("pattern recognition" OR "pattern generalization" OR "patterning" OR "pattern matching") AND ("computational thinking" OR "computing education" OR "computer science education").

The selection of the studies was done through an initial screening on titles and abstracts for meeting the following inclusion criteria:

1. Being a empirical study (experiments and applications or validations of tools) on PR;
2. Being a empirical study on CT while explicitly mentioning that PR is involved;
3. Being a conceptual study (reviews, frameworks, models and proposals) on PR;
4. Being a conceptual study on CT while explicitly mentioning that PR is involved;

The screening also excluded all papers who met the following exclusion criteria:

1. Not being in an educational context, i.e. not related to teaching-learning;
2. Having just mentioned PR, but not featuring it in the actual experiment/model;
3. Being a duplicate (already analysed from other search engine/databases);
4. Being written in any language other than English;
5. Not treating PR as the main theme;
6. Not being published in a journal or conference;

We sought for retrieval of the full paper of all those passing the screening phase or whose title and abstract were not enough to determine whether the study should be included or not. At last, we repeated the analysis upon the full texts for selecting the papers to be included in the review.

The data extraction process, guided by answering our research questions, consisted on coding the selected studies according to their approach (using games, digital tools, theoretical classes or robotics), tool (name of the game, activity or tool used, if any), form of assessment (pre/postest, surveys or rubrics) and if the study uses, from previous contributions of other fields: concepts/models to define PR; standardized/validated PR assessment methods; or discusses/compares PR in computing education to PR in other fields. The data synthesis process consisted on collectively analysing the extracted and coded data and individually discussing each selected study.

4. Results

Our search returned 264 documents, 2 of them were removed due to being a duplicate, 245 due to inclusion/exclusion criteria in the title/abstract screening and other 8 in the full text analysis, as well as 3 being removed because we couldn't get access to their full texts. Figure 1 details the search outcomes using the PRISMA [Page et al. 2021] flowchart. The selected studies are listed along with their respective approach, tool, form of assessment and contributions from other fields taken in consideration in Table 1.

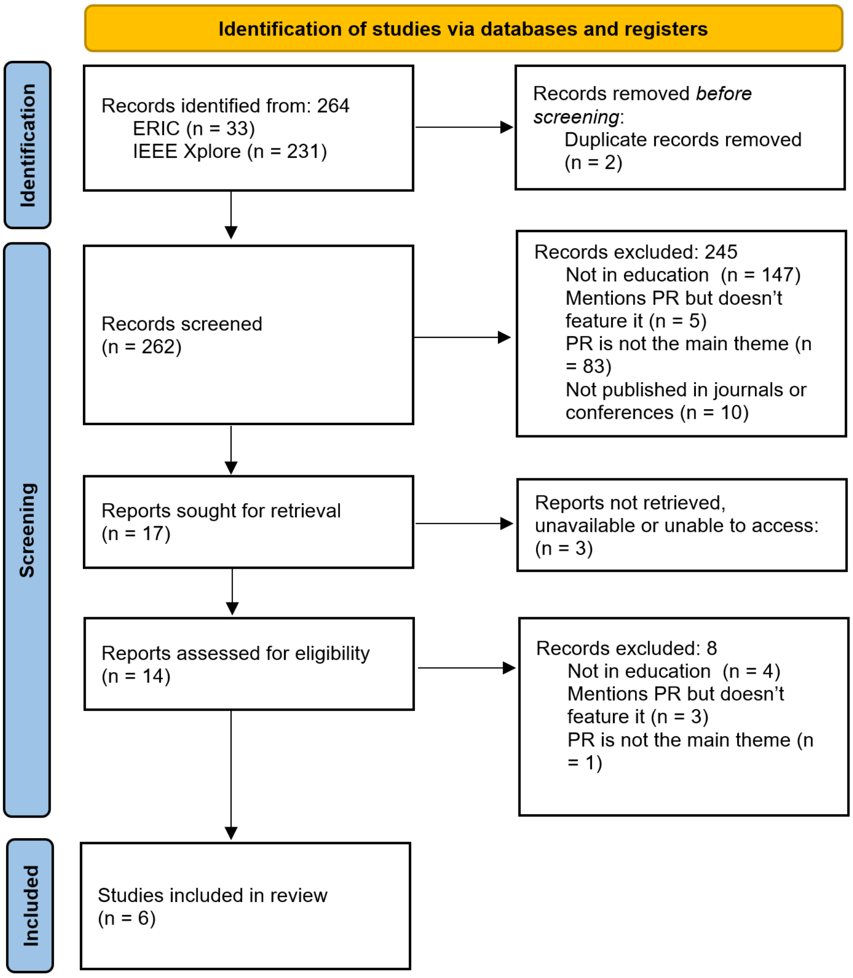


Figure 1. PRISMA flowchart of the review.

.

Table 1. Studies approaching Pattern Recognition in Computing Education.

Study	Approach	Tool	Assessment	From other fields
[Barrón-Estrada et al. 2022]	Mobile Application / Software	Patrony, presented by the study	Self-developed Pre and Post Tests	Uses psychology and cognitive neuroscience concepts (seriation-based PR)
[Guenaga 2021]	Digital Game / Software	Lempel, presented by the study	Self-developed Scoring System	None, everything is based on CT literature
[Saxena et al. 2020]	Unplugged Activities	LEGO patterns	Existing CT Rubric	None, everything is based on CT literature
[Miller 2019]	Programming	Scratch and unspecified robots	Self-developed Pre and Post Tests	Uses math and STEM concepts (mathematical patterns)
[Abdullah et al. 2019]	Mobile Application / Software	Unnamed app presented by the study	Self-developed Scoring System	Cites math concepts, but is based on CT literature
[Dasgupta and Purzer 2016]	Systematic Literature Review	None	None	None, everything is based on CT literature

The first study reports the development and experimentation of Patrony, a software application to promote the learning of CT, especially PR [Barrón-Estrada et al. 2022]. The software presents sequences of simple images and offers the user multiple choices for completing/continuing the sequence. Figure 2 shows an example of the Patrony activities.

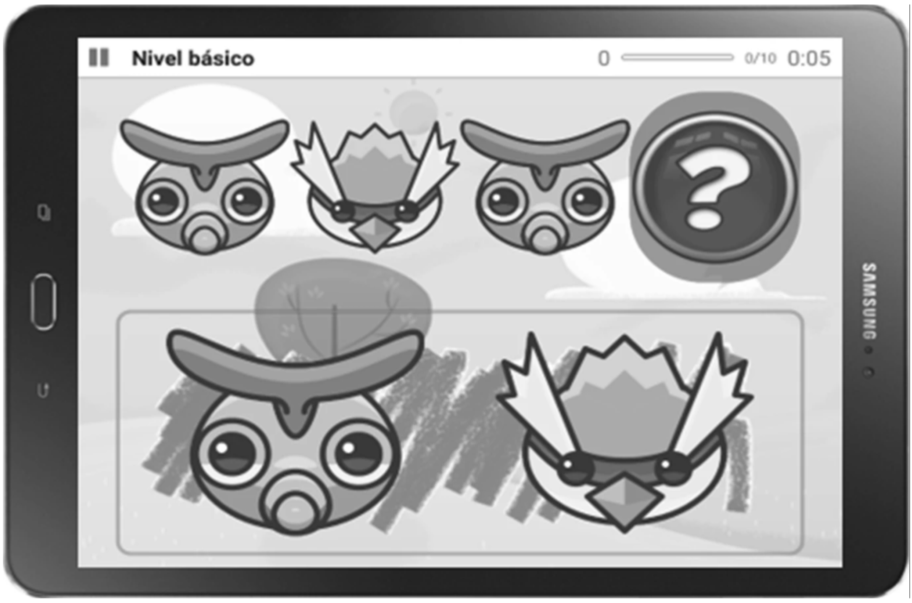


Figure 2. Patrony software. Source: [Barrón-Estrada et al. 2022]

.

The second study study reports the development and experimentation of Lempel, a game created to be an educational resource to develop CT in a classroom or stand-alone environment [Guenaga 2021]. The game consists in compressing text strings composed of different characters, where the players must recognize the pattern(s) in the string and insert them into containers called "registers". Figure 3 shows an example of the game.

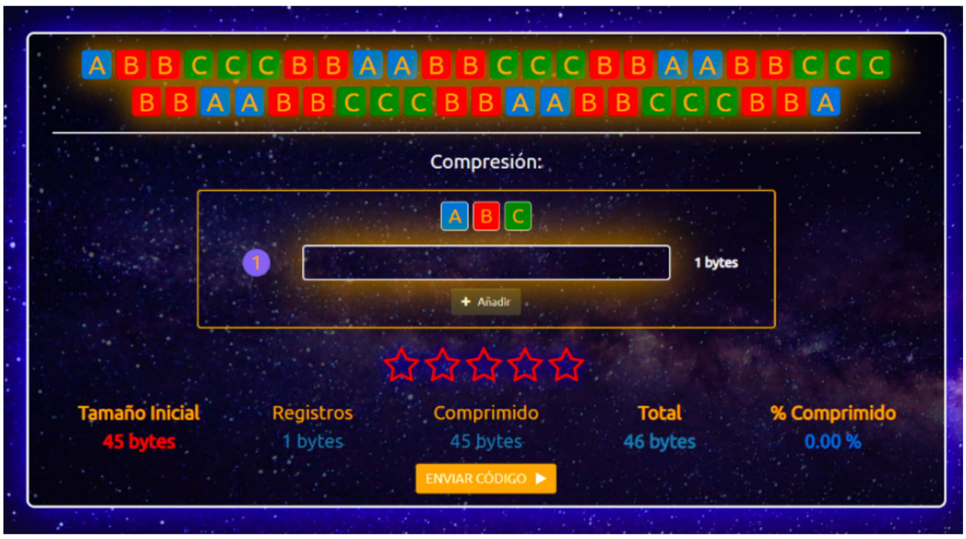


Figure 3. Lempel software. Source: [Guenaga 2021]

.

The third study reports the design and experimentation of a series of activities, involving both unplugged and plugged ones, to develop CT specifically in early childhood education. Their first proposed activity, LEGO pattern, targets PR [Saxena et al. 2020]. The activity consists in the teacher using LEGO bricks to create color and shape patterns, which are then asked to be continued by the students. Figure 4 shows the idea of the LEGO patterns activity.

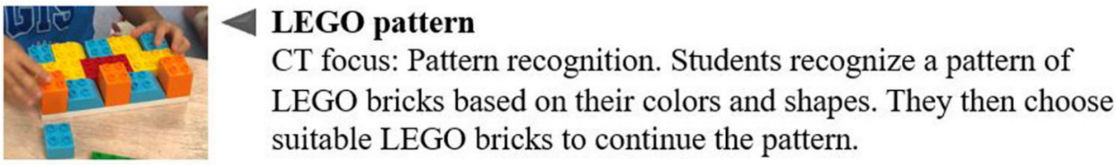


Figure 4. LEGO patterns activity. Source: [Saxena et al. 2020]

.

The fourth study reports experimentation of six coding lessons using Scratch [Resnick et al. 2009] and robots to understand how students build mathematical knowledge, specifically the identification of mathematical patterns, through coding [Miller 2019]. The lessons involved tasks that can be accomplished by the repetition of pattern, as drawing squares, drawing spirolaterals and moving robots through particular paths. Figure 5 shows the idea of the coding tasks and their relation with PR.

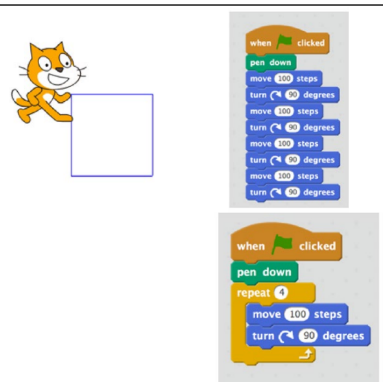
I made a square. Students were able to write a code to make the Scratch cat draw a square parallel to the bottom of the screen. Five students identified that they can see a repeating pattern and used the repeat coding block to draw a square.		

Figure 5. Coding task involving patterns. Source (Adapted from): [Miller 2019]

.

The fifth study reports the development and experimentation of an unnamed mobile application for fostering, in the Mathematics subject, the the concept of PR from CT [Abdullah et al. 2019]. The mobile app offers multiple choice quizzes in different levels of difficulty on PR, as well as mini-games of finding pairs through cards showing different patterns. Figure 6 shows three sample questions of the app's quizzes.

The sixth study is the SLR we mentioned and discussed in the Related Work section. It does not present approaches, tools or forms of assessment.

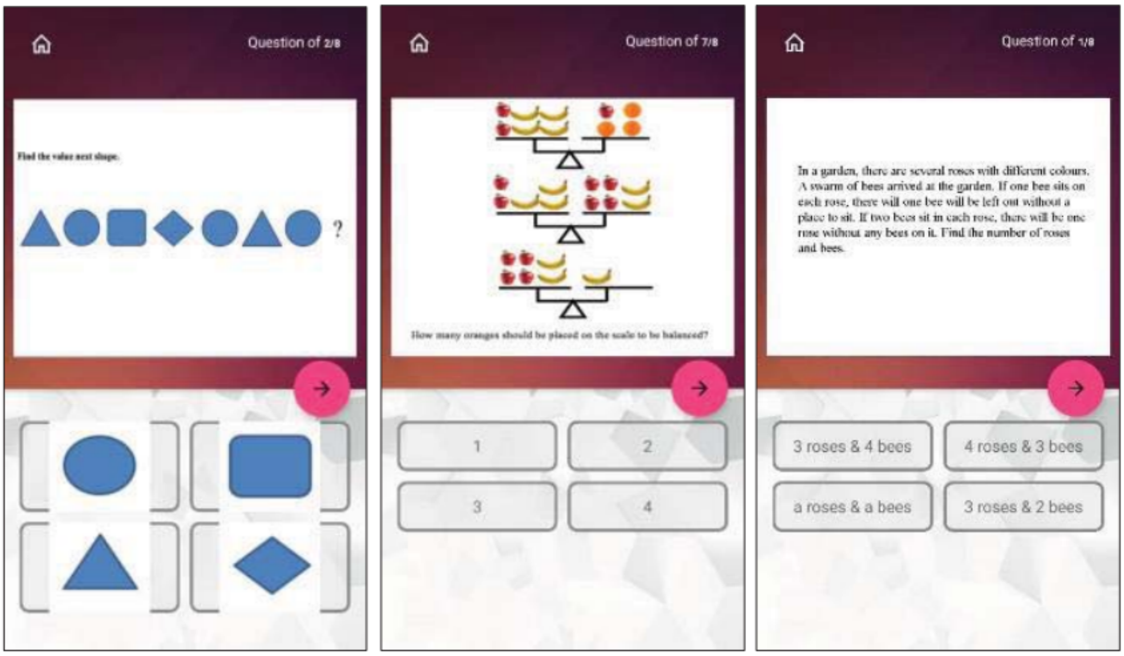


Figure 6. Unnamed PR quiz app. Source: [Abdullah et al. 2019]

.

4.1. How is Pattern Recognition being approached in Computing Education?

The results show that, in computing education, PR is being approached predominantly using Softwares, featuring games [Guenaga 2021] or gamified quizzes [Abdullah et al. 2019]/series of questions [Barrón-Estrada et al. 2022]. Alternatively, there are efforts to approach PR using unplugged activities [Saxena et al. 2020] and coding [Miller 2019]. The tools being used are mostly original tools that are presented in the study, as Patrony [Barrón-Estrada et al. 2022], Lempel [Guenaga 2021] and an unnamed PR quizz app [Abdullah et al. 2019]. Alternatively, there are efforts using LEGO bricks [Saxena et al. 2020] and Scratch allied to robots [Miller 2019].

4.2. How is Pattern Recognition being assessed in Computing Education?

The assessment of PR in computing education is being done mostly through self-developed resources [Barrón-Estrada et al. 2022, Guenaga 2021, Abdullah et al. 2019, Miller 2019], i.e. the authors of the papers are also the authors of the assessment resources used, which were developed specifically for the given experiment. However, there is a case using existing assessment resources on CT skills [Saxena et al. 2020]. The identified types of assessment are pre and post tests [Barrón-Estrada et al. 2022, Miller 2019], scoring systems [Guenaga 2021, Abdullah et al. 2019] and rubrics [Saxena et al. 2020].

The authors of Patrony developed tests consisting of 10 multiple choice questions asking things like correct form of completing sequences, as illustrated in Figure 7. The authors of Lempel developed a logging system to enable data analysis upon playing their game. The data analyzed for the assessment include: permanence rate through levels; success rate on 1-40 levels; average quality of the solution ("the quality is the percentage of compression being 0% the worst correct solution and 100% the best solution" [Guenaga 2021]); average interactions per user; and average time taken to resolve the level. Figure 8 shows an example of two statistics generated from the logging system.

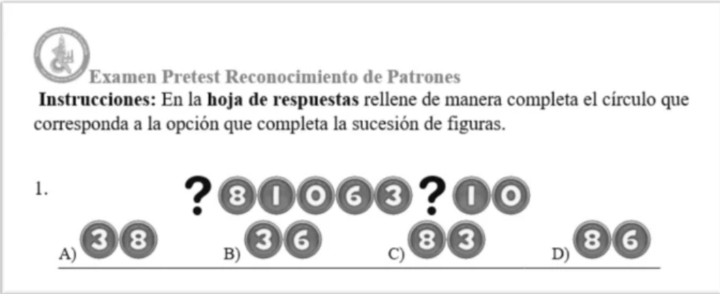
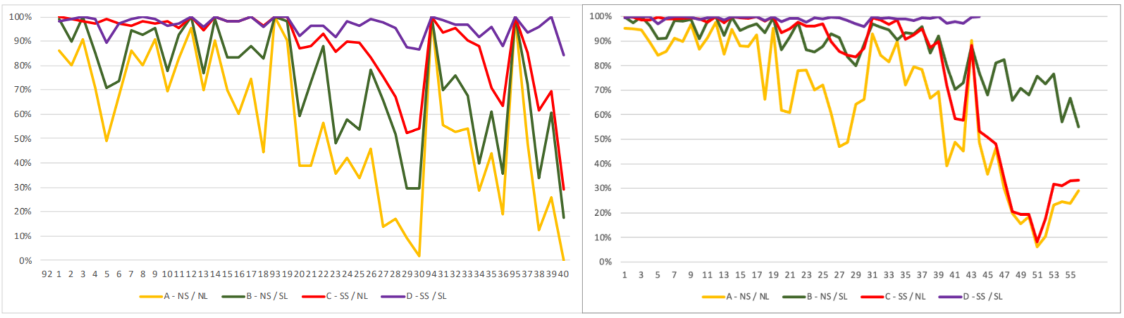


Figure 7. Patrony pre test sample question. Source: [Barrón-Estrada et al. 2022]

.



Success rate on 1-40 levels   Quality of the solution (avg.)

Figure 8. Success rate and quality of solution statistics of a Lempel application. Source (Adapted from): [Guenaga 2021]

.

The authors of the LEGO patterns activity used the rubric of performance assessments for CT activities from [Bers et al. 2014], showed in Figure 9.

Score	Description
5	Complete achievement of the goal, task, or understanding
4	Mostly complete achievement of the goal, task, or understanding
3	Partially complete achievement of the goal, task, or understanding
2	Very incomplete achievement of the goal, task, or understanding
1	Did not complete the goal, task, or understanding
0	Did not attempt/Other

Figure 9. Rubric of performance assessments for CT activities from [Bers et al. 2014]. Source: [Saxena et al. 2020]

.

The authors of the six coding lessons developed tests they report that considers prior research on patterning and structures. “The items included students identifying repeating patterns, making predictions about patterns, generalising patterns, and creating patterns” [Miller 2019]. There is one test with 10 patterning tasks (maximum score of 30) and another with 10 coding tasks (maximum score of 10). Figure 10 shows sample patterning and coding tasks from the developed tests.

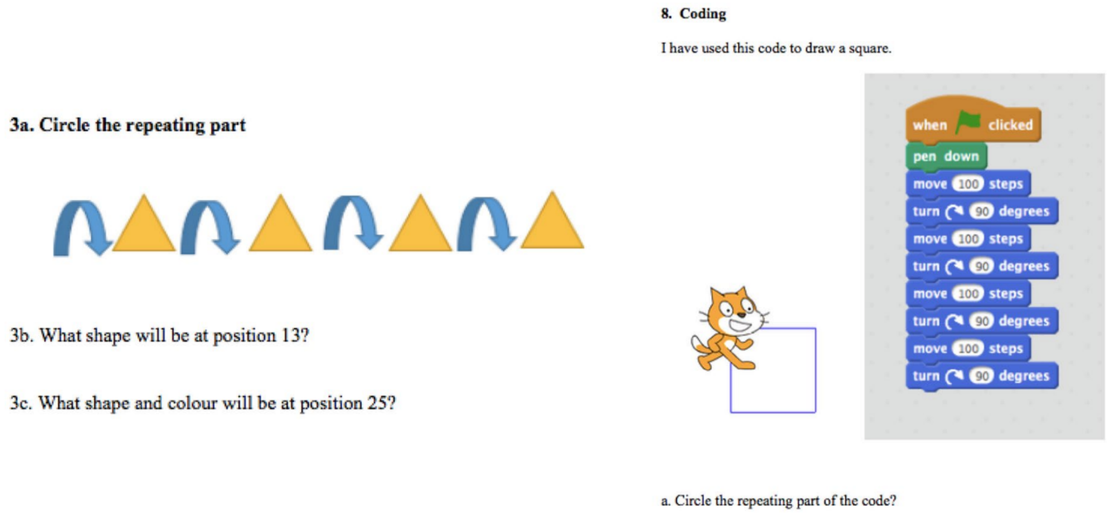


Figure 10. Sample tasks of the patterning and coding test. Source (adapted from): [Miller 2019]

.

The authors of the PR quizz app developed a scoring system based on time spent in each mini game and quiz question, as well as the selected difficulty level. A total of 23 experts from a secondary school were consulted to evaluate the app, in addition to several Lickert-scale questionnaires, on: facilitation, usefulness, design, knowledge provided and suitability of the app [Abdullah et al. 2019]. However, nothing explicitly targeted the validity of the scoring system in particular.

4.3. What contributions to Patter Recognition research from other fields are mentioned or discussed in Computing Education?

The CT background was unanimous amongst the selected study. However two emphasized STEM and Mathematics [Abdullah et al. 2019, Miller 2019]. While one brought psychology and cognitive neuroscience as their theoretical backgrounds [Barrón-Estrada et al. 2022].

The Patrony work bases their definitions of PR on cognitive psychology references [Kidd et al. 2012, Eysenck and Keane 2015, Shugen 2002], seeing PR as cognitive process to combine information from stimulus to those retrieve from memory involving short-term, long-term and semantic memory. They particularly explore the concept of seriation, “the ability to organize elements in a logical order along with a quantitative dimension, such as length, weight, and age” [Kidd et al. 2012].

The six coding lessons work bases their definitions of PR on mathematics references [Zazkis and Liljedahl 2002, Mulligan and Mitchelmore 2009, Warren et al. 2012], seeing PR (or patterning, as they call) as what forms the basis for young students to recognise mathematical structures and engage with early algebraic thinking [Miller 2019]. They particularly explore the concept of mathematical patterns, “predictable regularities usually involving spatial, numerical or logical relationships” [Mulligan and Mitchelmore 2009].

The PR quizz app work uses some mathematics references to define PR, but explicitly generalize them for CT: "It should be noted that the concept of pattern recognition in computational thinking has a free context and application which means that when students have mastered the concept of this pattern conceptually, they will be able to integrate it in any area of study" [Abdullah et al. 2019].

5. Conclusion

Thiw work reviewed how PR is being held in computing education. We found only 6 studies, being one of them a SLR from 2016 exposing how scarce studies exploring PR are. While the other 5 were experiments showing that PR in computing education is unanimously related to CT, mostly approached using digital games or gamified environments, and assessed using self-developed resources, usually pre and post tests. It was also revealed that despite PR being a topic and keyword vastly used in other fields since a long time ago, few studies consider any contribution from other fields to PR. When doing so, it is restricted to conceptual foundations, as it was found for mathematics and cognitive psychology. Our suggestions for future research are: first, run more experiments on PR, since the literature has been scarce for a long time now and barely evolved in the recent years; and second, make use of the knowledge other fields already built on PR, as did [Barrón-Estrada et al. 2022] considering the seriation phenomenon from psychology.

References

Abdullah, A. H., Othman, M. A., Ismail, N., Abd Rahman, S. N. S., Mokhtar, M., and Zaid, N. M. (2019). Development of mobile application for the concept of pattern recognition in computational thinking for mathematics subject. In *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*, pages 1–9. IEEE.

Acevedo-Borrega, J., Valverde-Berrocoso, J., and Garrido-Arroyo, M. d. C. (2022). Computational thinking and educational technology: A scoping review of the literature. *Education Sciences*, 12(1):39.

Barrón-Estrada, M. L., Zatarain-Cabada, R., Romero-Polo, J. A., and Monroy, J. N. (2022). Patrony: A mobile application for pattern recognition learning. *Education and Information Technologies*, 27(1):1237–1260.

Bers, M. U., Flannery, L., Kazakoff, E. R., and Sullivan, A. (2014). Computational thinking and tinkering: Exploration of an early childhood robotics curriculum. *Computers & Education*, 72:145–157.

Bishop, C. M. and Nasrabadi, N. M. (2006). *Pattern recognition and machine learning*, volume 4. Springer.

Dasgupta, A. and Purzer, S. (2016). No patterns in pattern recognition: A systematic literature review. In *2016 IEEE Frontiers in Education Conference (FIE)*, pages 1–3. IEEE.

Eysenck, M. W. and Keane, M. T. (2015). *Cognitive psychology: A student'handbook*. Psychology press.

Guenaga, M. (2021). Lempel: Developing the pattern recognition skill in computational thinking through an online educational game. In *2021 Learning Analytics Summer Institute (LASI)*.

Kellman, P. J. and Garrigan, P. (2009). Perceptual learning and human expertise. *Physics of life reviews*, 6(2):53–84.

Kidd, J. K., Curby, T. W., Boyer, C. E., Gadzichowski, K. M., Gallington, D. A., Machado, J. A., and Pasnak, R. (2012). Benefits of an intervention focused on oddity and seriation. *Early Education & Development*, 23(6):900–918.

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004):1–26.

Liu, J., Sun, J., and Wang, S. (2006). Pattern recognition: An overview. *IJCSNS International Journal of Computer Science and Network Security*, 6(6):57–61.

Miller, J. (2019). Stem education in the primary years to support mathematical thinking: Using coding to identify mathematical structures and patterns. *Zdm*, 51(6):915–927.

Mulligan, J. and Mitchelmore, M. (2009). Awareness of pattern and structure in early mathematical development. *Mathematics Education Research Journal*, 21(2):33–49.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., et al. (2021). The prisma 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88:105906.

Pi, Y., Liao, W., Liu, M., and Lu, J. (2008). Theory of cognitive pattern recognition. *Pattern recognition techniques, technology and applications*, pages 433–463.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive psychology*, 3(3):382–407.

Resnick, M. et al. (2009). Scratch: Programming for all. *Communications of the ACM*, 52(11):60–67.

Saqr, M., Ng, K., Oyelere, S. S., and Tedre, M. (2021). People, ideas, milestones: a scientometric study of computational thinking. *ACM Transactions on Computing Education (TOCE)*, 21(3):1–17.

Saxena, A., Lo, C. K., Hew, K. F., and Wong, G. K. W. (2020). Designing unplugged and plugged activities to cultivate computational thinking: An exploratory study in early childhood education. *The Asia-Pacific Education Researcher*, 29(1):55–66.

Selby, C. and Woollard, J. (2013). Computational thinking the developing definition. *University of Southampton (E-prints)*.

Shugen, W. (2002). Framework of pattern recognition model based on the cognitive psychology. *Geo-spatial Information Science*, 5(2):74–78.

Tang, X., Yin, Y., Lin, Q., Hadad, R., and Zhai, X. (2020). Assessing computational thinking: A systematic review of empirical studies. *Computers & Education*, 148:103798.

Tanimoto, S. L. (1998). Connecting middle school mathematics to computer vision and pattern recognition. *International journal of pattern recognition and artificial intelligence*, 12(08):1053–1070.

Wang, C., Shen, J., and Chao, J. (2021). Integrating computational thinking in stem education: A literature review. *International Journal of Science and Mathematics Education*, pages 1–24.

Warren, E., Miller, J., and Cooper, T. (2012). Repeating patterns: Strategies to assist young students to generalise the mathematical structure. *Australasian Journal of Early Childhood*, 37(3):111–120.

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

Zazkis, R. and Liljedahl, P. (2002). Generalization of patterns: The tension between algebraic thinking and algebraic notation. *Educational studies in mathematics*, 49(3):379–402.