

ORIGINAL RESEARCH ARTICLE

Empowered learning through microworlds and teaching methods: a text mining and meta-analysis-based systematic review

Joana Martinho Costa^{a*}, Sérgio Moro^a, Guilhermina Miranda^b, and Taylor Arnold^c

^aInstituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, Lisboa, Portugal; ^bInstituto de Educação, Universidade de Lisboa, Lisboa, Portugal; ^cDepartment of Mathematics and Computer Science, University of Richmond, Richmond, VA, USA

(Received: 27 May 2020; Revised: 7 September 2020; Accepted: 8 September 2020; Published: 12 October 2020)

Microworlds are simulations in computational environments where the student can manipulate objects and learn from those manipulations. Since their creation, they have been used in a wide range of academic areas to improve students learning from elementary school to college. However, their effectiveness is unclear since many studies do not measure the acquired knowledge after the use of microworlds but instead they focus on self-evaluation. Furthermore, it has not been clear whether its effect on learning is related to the teaching method. In this study, we perform a meta-analysis to ascertain the impact of microworlds combined with different teaching methods on students' knowledge acquisition. We applied a selection criterion to a collection of 668 studies and were left with 10 microworld applications relevant to our learning context. These studies were then assessed through a meta-analysis using effect size with Cohen's d and p -value. Our analysis shows that the cognitive methods combined with microworlds have a great impact on the knowledge acquisition ($d = 1.03$; $p < 0.001$) but failed to show a significant effect ($d = 0.21$) for expository methods.

Keywords: meta-analysis; microworlds; simulations; teaching methods; technology-enhanced learning

Introduction

Since the creation of LOGO (Logic Oriented Graphic Oriented) (Papert 1980; Papert *et al.* 1979), microworlds have been used by teachers in a wide range of areas to enhance students' knowledge acquisition and to promote the development of higher order cognitive skills, such as planning a course of actions or heuristics of problem-solving (e.g. Pea and Kurland 1984). However, the development of cognitive skills and the acquisition of disciplinary knowledge associated with microworlds in schools are influenced by a wide range of variables that many studies do not discriminate. One of these variables is the teaching method, not always measured, because some studies use microworlds in the classroom to study other variables, such as

*Corresponding author. Email: joana.martinho.costa@iscte-iul.pt

improving students' learning just by using the microworld (e.g. Fargas-Marques and Costa-Castelló 2006; Yurtseven and Buchanan 2012).

We begin our paper with the definition of microworld and the main teaching methods that guides the teaching strategies with microworlds in the classroom. Next, we conduct a systematic review to understand the characteristics of the studies that used microworlds to achieve knowledge and pointed out the teaching method in their work. Finally, we present a meta-analysis with effect size and the study of the p-curve to understand the effectiveness of the teaching methods associated with the use of microworlds on students' knowledge acquisition.

Theoretical framework

Microworlds

Papert (1980) has defined microworlds as simulations in computational environments where students can manipulate their objects and learn from those manipulations. A microworld is also a framework of learning and an incubator of knowledge as the entire learning experience must be able to take place exclusively within the microworld (Papert 1980; Somekh 1996). Sometimes, students also must learn the language of the microworld itself as in cases where its software only accepts a specific syntax (Somekh 1996). In general, semantics of the content, simulations and the interface of microworlds contribute to the learning experience of students (Somekh 1996).

Although there are divergences between some authors regarding the concept and characteristics that should be present in a microworld (e.g. Hoyles, Noss, and Adamson 2002; Rieber 1996; Sarama and Clements 2002), we will ground our study on the original concept of Papert (1980) because it reflects the initial concept of microworld, still valid, and that led to different views.

Not all digital learning environments can be considered as microworlds. According to Papert (1980), a microworld must have three main characteristics: (1) it should allow the creation of simple examples related to the knowledge that is expected to be acquired; (2) there should be no obstacles in the manipulation of the objects within the microworld and (3) the required concepts for learning must be definable within the microworld.

The first microworld, the Geometric Tortoise from LOGO (Papert 1980), was created a few decades ago, but nowadays there are a large diversity of microworlds that can be applied in several academic areas such as computer programming with Alice (Alice Project 2020; Scratch 2020), neurosciences with BrainExplorer (Schneider *et al.* 2013) or mathematics with Speedy World (Wang *et al.* 2018), having in common the three characteristics pointed out by Papert (1980).

Taking the example of Alice software (Alice Project 2020), we can assume that it is a microworld for programming learning since the construction of simple instructions allows the opportunity of learning the introductory concepts of computer science (criterion (1)), the manipulation of characters and space does not have any limitation and students can create their own instructions and relate them to the chosen characters as well (criterion (2)) and students can define a large set of variables, data types and instruction for each character according to the learning objectives to be achieved (criterion (3)).

Teaching methods

There are several ways to categorize the different teaching methods and instructional design models. The teaching methods, namely, the traditional method, are more comprehensive and older than the instructional design models, since the latter appeared in the 1950s, with Skinner's programmed teaching (Skinner 1954), and are closely related to the scientific theories of learning and human development (cf. Bruner 1966; Gagné 1985). Teaching methods aim to promote knowledge acquisition and optimize student learning. Instructional design also has the same goal, although it is mainly concerned with optimizing learning and performance (cf. Merrill 2002), has a more prescriptive and normative bent and is based on scientific theories and models of human learning.

Regarding the teaching methods, we chose a categorization that can be easily grasped: the direct and indirect teaching methods. The former comprises some teaching strategies with a long history, such as the teacher's exposure or lecture, included in the so-called conventional or traditional methods of teaching. While much of the time spent in the classroom is used by the teacher to expose the matter, as Bruner (1965) refers to as the main feature of the educated societies, each discipline has a typical mode to develop and sequence the classroom activities, carried out by the teacher and students (Schofield 1995). These methods can also be defined as teacher centred (Bransford, Brown, and Cocking 2000). The instructional models that may be included in this category are the so-called behaviourist or instructivist models, such as Skinner's model (Skinner 1954) referred earlier, with the difference that instead of being the teacher who teaches the student, the program takes the same role by leading the student to the desired goal by successive stages of increasing difficulty.

The indirect teaching methods include the project method (Kilpatrick 2007), problem-based learning, guided discovery learning (Bruner 1966) and other related methodologies. Constructivist instructional models may also be included in this category. These methods are also called student centred (Bransford, Brown, and Cocking 2000). They advocate that learning is more consistent when it is built by the students, individually but preferably in collaboration with peers or someone more competent, as opposed to that which is transmitted by the teacher. In other words, 'the rationale underlying these so-called discovery approaches is that the material that is generated is better learned than the material that is only received' (De Jong and Lazonder 2014, p. 371).

Cognitive instructional models base their proposals on experimental results on human cognitive architecture, where memory is the mechanism that allows human beings to learn (e.g. Anderson 1993; Baddeley 1997). Examples of these models are Robert Gagné's conditions of learning theory (Gagné 1985) and van Merriënboer's four components instructional design model (van Merriënboer 1997). Regarding the teaching methods, we can include the meaningful learning theory of David Ausubel (2000) and the mastery learning of Benjamin Bloom (1956).

Relationship between microworlds and teaching methods

The integration of microworlds in education in the mid-1980s, after the publication of the book 'Mindstorms: Computers and powerful ideas' (Papert 1980), leads to many teachers and researchers, influenced by Papert's ideas, thought that it was

enough to get children to program without having to teach them, as the ‘nuggets of knowledge’ integrated into microworlds would be easily learned and transferred. Experimental research has contradicted this idea (e.g. De Corte 1993; Pea and Kurland 1984). Thus, together with the LOGO microworld, the indirect teaching methods were privileged, especially the guided discovery learning, which produced no statistically significant effects on knowledge acquisition and development of cognitive skills by students (e.g. Yuen-Kuang and Brighth 1991). Positive and significant results were found when teachers and researchers used teaching strategies that guided students (cf. Mayer 2014; Mendelsohn 1991), supporting researchers who criticize constructivist methods for learning complex skills (Kirsner 2019; Littlefield *et al.* 1988).

The most important conclusion that can be drawn from all the research efforts associated with microworlds in the learning of children and young people is that it has no intrinsic virtues unless teachers and researchers associate it with certain teaching methods. Microworlds can be combined with distinct teaching approaches (McDougall 2002). In cognitive approaches, students have access to correct solutions of problems to be solved inside the microworld, aiming to change their cognitive structures to match scientific understanding. The constructivist or constructionist approach advocated by Papert (1980) has the main intention of guiding the student to build or modify his or her own knowledge even if it involves more time to explore the microworld and the absence of scientific theories in the manipulation of objects (McDougall 2002).

The most promising seems to be the cognitivist or mixed teaching models due to their good balance between direct and indirect teaching strategies (De Corte 1993). However, remains unclear which are the most efficient combinations. We propose the development of a meta-analysis to answer this question.

Method

The systematic review

We conducted our systematic review through three steps (Cooper and Hedges 2009): (1) by searching on databases for specific term related to our research question; (2) application of inclusion and exclusion criteria to the resulted studies from (1); and (3) codification of the studies and meta-analysis.

The teaching method is an important factor in the use of technology in the schools as we mentioned before and recent research has confirmed (Costa and Miranda 2019; Vosinakis, Anastassakis, and Koutsabasis 2018). However, not all studies highlight the association between the use of microworlds and the teaching methods. Therefore, in our systematic review, we focus only on the studies that considered this variable combined with microworlds, even if it was only for control purposes.

Search criteria

During June 2019, we collected studies on the major databases: ACM (Association for Computing Machinery) Digital Library, ERIC (Education Resources Information Center), IEEE (Institute of Electrical and Electronics Engineers) Digital Library, ISI (Institute for Scientific Information) Web of Science, Science Direct, Springer Open and DOAJ (Directory of Open Access Journals). We only searched for the term

‘microworld’ and did not select any time range because we wanted to reach all the studies about the theme, with no time restrictions.

The search criteria returned a total of 668 articles. Through a subsequent analysis of their corresponding titles, 395 articles were identified as related to the educational field and were selected for a deeper analysis.

Inclusion criteria

The total number of selected articles, 395, is still a large number of manuscripts, which would require careful reading to validate if each one constitutes a valid contribution from the microworld in learning context. Furthermore, at least two experts would need to validate each article to mitigate the known human assessment subjectivity (Santos, Laureano, and Moro 2019). Therefore, an approach based on standard text mining techniques can be an alternative to automatically select those articles that are relevant to the studied subject (Moro *et al.* 2019), and help in defining and executing inclusion criteria. This approach consists in identifying a lexicon that enables to correctly categorize each article, and then in running a text mining script that builds a document-term matrix that quantifies how many times each of the terms from the lexicon occurs within each article (Cortez *et al.* 2018). Although the lexicon definition is subjective, after it is completed, the computational parsing process of each text is the same regardless of the article. Also, the negligible computational time taken and the fact that the script can be executed any number of times to tune the lexicon are two other advantages of this approach. The lexicon was defined under two key semantic concepts: (1) quantitative research design and (2) teaching strategies, according to the main concepts presented in ‘Teaching methods’ section. It should be stated that the list of terms was developed by considering a broader perspective, that is, we prefer to select a few irrelevant articles to discarding relevant ones. Tables 1 and 2 exemplify the main terms (reduced terms) and some of the corresponding related terms for each case (the full list can be consulted at <https://fenix.iscte-iul.pt/homepage/smcmo@iscte.pt/microworlds>). An article would be included if it met at least one term in each dictionary. For example, if an article contained the ‘pre-experimental’ term, stated in Table 1, and ‘cognitivism’, stated in Table 2, it would be added to the results from the application of the inclusion criteria. The application of the defined inclusion criteria returned 105 articles.

Table 1. Semantic concepts on quantitative research design.

Reduced term	Similar terms or from the same domain
Experimental	Experimental
Quasi-experimental	Time series, equivalent time samples, non-equivalent control groups (CGs), systematically selected CGs, remediation process, combination process, single subjects
Non-experimental	Pre-experimental, intact group
Quantitative	Quantitative, empirical
Mixed method	Mixed method
Factorial	Factorial
Ex post facto	Correlational, criterion group, longitudinal

Table 2. Semantic concepts on teaching strategies.

Reduced term	Similar terms or from the same domain
Instructional design	Instructional design, student centred, teacher centred, teaching strategy, teaching strategies, teaching strategy, differentiated instruction, scaffolding
Learning design	Flipped classroom, personalized learning, game based, inquiry based, problem based, project based, differentiated, direct instruction, teaching strategy, expository, gamification, measurable objectives, directivity, competency-based education, training, reinforcement, cognitive development, multimedia learning, andragogy, self-directed learning
Teaching method	Behaviourism, behaviorism, social cognitive, connectivism, constructivism, cognitivism, kinaesthetic, universal design, cognitive theory, humanism, socioconstructivism

Exclusion criteria

The application of text mining did not distinguish whether the studies were only descriptive or the words of the dictionaries were introduced in the right context of our work. So, at this phase, we decided to apply the exclusion criteria manually to introduce a semantical analysis. We defined a set of criteria related to the study characteristics that could undermine the validity of the meta-analysis and the bias of results, namely:

- (1) No presentation of results/descriptive article;
- (2) Document compilation, not entirely related to our research scope;
- (3) Study does not use microworlds;
- (4) Results provided from self-reports;
- (5) Results provided from time counting or items counting;
- (6) Results not related to achievement or learning;
- (7) Results do not allow the calculus of Cohen's *d*.

The exclusion criteria were applied by the indicated order. If one study did not accomplish one exclusion criterion, it was automatically excluded since it was not possible to extract all the required information for subsequent analysis.

After the application of the exclusion criteria, a total of 10 articles remained.

Codification of the studies

All descriptors that we used to develop this systematic review are available in Appendix A. Following Costa and Miranda's 2017 approach, we have followed the same guidelines of Hedges, Shymansky, and Woodworth (1989) for codification of the studies in their meta-analysis and then adapted to our work context.

The meta-analysis

As most of the systematic reviews, our 10 studies contain a set of substantial differences that could make it difficult to define the variance between them and the

implications for understanding the effect of microworlds with different teaching methods. However, the meta-analysis process can mitigate those issues (Borenstein *et al.* 2009).

In a first approach, we combined the studies through the effect sizes of each study using Cohen's *d*, and we converted the studies that did not present this measure. Conducting a meta-analysis with effect sizes to synthesize our studies has several advantages to accurate the effect of microworlds on learning for its precision and clear method (Borenstein *et al.* 2009; Hedges and Pigott 2004). This method also enables us to understand the factors and characteristics of the studies that influence the effect size and to explain the differences of the results between them (Coe 2002).

In a second approach, we combined the studies through the *p*-values. The *p*-curve method provides a meta-analysis method for determining the likelihood that an effect can be explained as a result of publication bias or *p*-hacking (Simonsohn, Nelson, and Simmons 2014a). A *p*-curve looks at the distribution of significant *p*-values and attempts to determine whether the distribution has a right-skew, which are diagnostic of evidentiary value (Simonsohn, Nelson, and Simmons 2014a). We followed the more specific and robust procedure described in Simonsohn, Nelson, and Simmons (2014b), and as applies recent meta-analyses in psychological research (Nelson, Simmons, and Simonsohn 2018).

Finally, we analyse the results, their limitations and conclusions.

Results

The main characteristics of the 10 selected studies are available in Appendix B. The first five studies were performed with an expository teaching method and the last five studies with a cognitive method. We present the results of all studies after the experimental treatment.

These studies were conducted in three different continents represented by six countries, namely, USA, Brazil, Israel, Belgium, Germany and Portugal. Most of them were conducted in the area of computer science, but there is some expression of the natural sciences and the arts. All of the studies measured the achieved knowledge in an academic field and most of them used tests as data collection instruments, except for Pfahl, Koval, and Ruhe (2001) study, which used projects.

From this combination, we verify in both groups a large effect size in studies conducted in K-12. These studies were implemented in different areas of knowledge (computer science, drama and biology), so the area does not seem to be the cause. On the other hand, the studies implemented in high school present low or negative effect size. These studies were only conducted with the expository method. Studies in this age range that use cognitive methods must be implemented in further research.

Effect size analysis

The studies with more than one outcome and treatment have the results correlated (Borenstein *et al.* 2009). Borenstein *et al.* (2009) suggest two solutions to deal with this issue. The first consists of the study of the effect of each experimental group (EG) on the effect of the CG separately, which would result in an increase in the power of that study relative to others and bias the estimate (Van den Noortgate *et al.* 2013). The alternative solution is to analyse the difference of the effect size between

the two EGs. However, our data could not follow this option due to some differences in measurements between studies. For these cases, we followed the Van den Noortgate *et al.*'s (2013) suggestion and selected only one effect size per study. The same authors also recommend the use of random-effect model for combine the studies with these characteristics. Table 3 presents the Cohen's *d* of each study (column *d*), their confidence interval to 95% (column 95%-CI), the weight of each study using random-effect model (column Weight) and the teaching method (column Method).

Table 4 synthesizes the Cohen's *d* and the heterogeneity test by method. The results revealed a significant difference between the subgroups ($Q = 12.51; p < 0.01$).

Figure 1 illustrates the forest plot with all studies. TE means the effect size and seTE means the standard error of the effect size. From the forest plot, we verify that there is a difference of less than 10 percentage points between the subgroups and both have five studies. These characteristics allow us to compare fairly the use of the expository and cognitive methods because the distribution is similar.

Table 4 and Figure 1 evidence that the combination of the subgroup with the expository methods contains a very low heterogeneity (only 14.4%) with a combined effect size ($d = 0.24$) considered small in educational field (Hattie 2009). On the other hand, the combination of the subgroup with the cognitive methods contains a moderate heterogeneity (53.8%) with a combined effect size ($d = 1.03$) considered strong in educational field (Hattie 2009). These results point out to a strategy in classroom that use microworlds combined with cognitive methods not only is it better than with the expository methods, but it also has a strong impact on knowledge acquisition since it is expected that 84% of the subjects of the CG would be below the average subjects of the EG (Coe 2002). Meanwhile, microworlds combined with expository methods do not indicate a significant effect since only 58% of the subjects of the CG would be below the average subjects of the EG (Coe 2002).

Table 3. Effect sizes.

Study	<i>d</i>	95%-CI	Weight	Method
Jenkins (2015)	0.62	[0.07; 1.17]	10.8	Expository
Oliveira, Monteiro, and Roman (2013)	0.08	[-0.72; 0.89]	8.4	Expository
DiCerbo <i>et al.</i> (2010)	0.30	[-0.1; 0.69]	12.3	Expository
Costa (2019)	-0.43	[-1.24; 0.39]	8.3	Expository
Pfahl, Koval, and Ruhe (2001)	0.45	[-0.75; 1.65]	5.6	Expository
Schneider <i>et al.</i> (2013)	1.51	[0.7; 2.33]	8.3	Cognitive
Lehrer, Lee, and Jeong (1999)	1.24	[0.64; 1.83]	10.4	Cognitive
Moons and Backer (2013)	0.89	[0.24; 1.53]	9.9	Cognitive
Kluge (2008)	0.66	[0.37; 0.96]	13.1	Cognitive
Zohar and Peled (2008)	1.20	[0.88; 1.51]	13.0	Cognitive

Table 4. Overall results for subgroups using random-effect model.

Method	Studies	<i>d</i>	95%-CI	<i>Q</i>	tau2
Expository	5	0.24	[-0.23; 0.72]	4.68	0.0768
Cognitive	5	1.03	[0.63; 1.43]	8.67	0.059

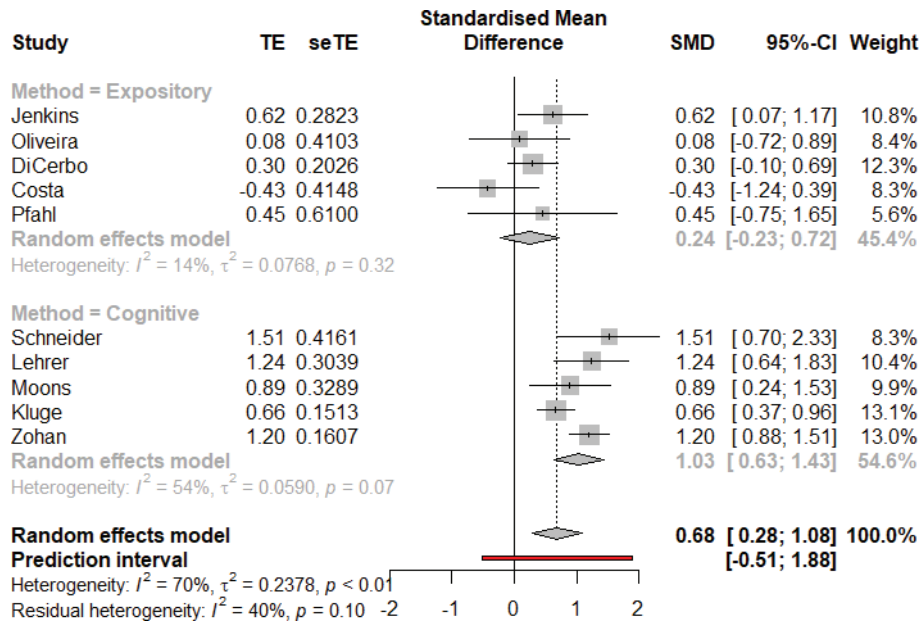


Figure 1. Forest plot with the studies.

P-curve analysis

Of the 10 studies selected in the analysis, only 6 had significant *p*-values and were therefore included in the p-curve analysis. The resultant p-curve can be seen in Figure 2. The results of the associated full p-curve and half p-curve statistical hypothesis tests indicate that there is evidential value ($p < 0.001$ and $p < 0.001$, respectively); a power analysis gives an estimated power of 92% (CI: 69%–99%). Together, these results indicate that it is likely that the observed *p*-values cannot be explained as result of publication bias or p-hacking and therefore indicate true evidentiary value.

We note that all but one of the *p*-values for the expository analyses are not significant at the 0.05 level, whereas all but one of the *p*-values for the cognitivism analyses are significant. Running a p-curve analysis for the cognitivism analysis returns similar results, with an indication of evidential value ($p < 0.001$) and an estimated power of 96% (CI: 82%–99%).

Conclusions

Considering the used methodology in this study, we seem to emphasize the combination of the traditional meta-analysis, with the measurement of the traditional effect size and the p-curve analysis as main criterium, an emerging methodology that reinforces the results of the traditional meta-analysis. We suggest that, in future meta-analysis, researchers should focus on combining these two statistical analysis methodologies, as they are complementary. We also point out the use of text mining technique to select the articles that would be included in the meta-analysis as a great technique whether the dictionaries are appropriate. In our study, to test the robustness of our dictionaries, we applied the inclusion and exclusion criteria manually to

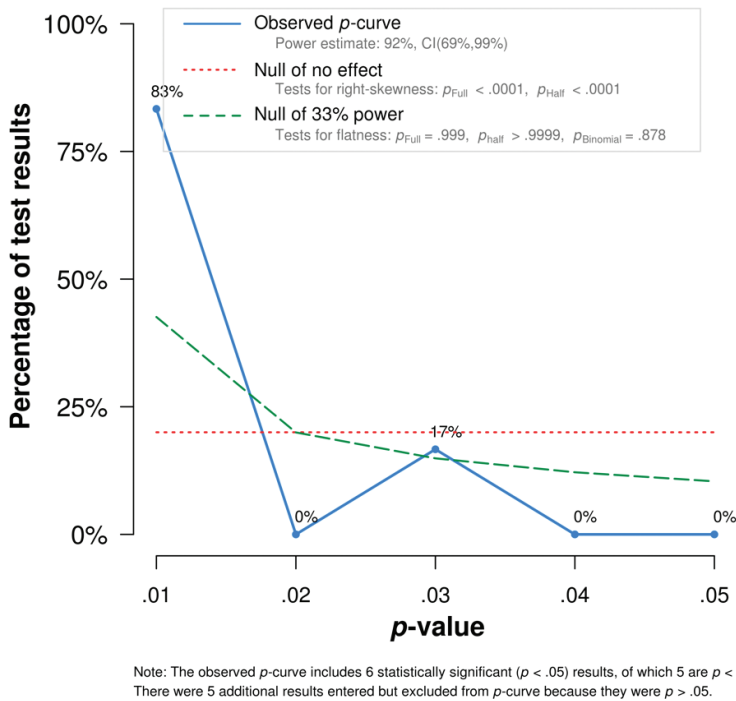


Figure 2. P-curve showing the distribution of significant results in the meta-analysis.

30 random studies from the search criteria and they matched the exact result of the process with text mining.

Regarding the results obtained in this work, the most salient is that cognitive methods associated with the microworlds have evident and significant effects on the acquisition of knowledge by students. The traditional method, namely, the expository method, has little impact on student learning, without statistical significance. One of the characteristics of the expository method is that it is centred on the teacher's strategies, reserving a less active role for the students. Today, we know that student activity is important in knowledge acquisition. Cognitive methods attach great importance to student cognitive activity and base their strategies on the way humans' process information (Ausubel 2000; Gagné 1985; van Merriënboer 1997). They also consider the results obtained from experimental research on cognitive architecture and the functioning of human memory (Anderson 1993; Baddeley 1997; Mayer 2014; Sweller 2011). We think that it is these characteristics of cognitive methods associated with the characteristics of the microworlds that facilitate and optimize student learning. Constructivist teaching methods associated with the microworlds, although widespread, did not enter the meta-analysis because they did not meet the inclusion and exclusion criteria. Many of these studies have no empirical results, and when they do, it is primarily about describing experiences and subjective analysis of student motivation. Another point to note is that the instruments used to collect empirical data are not always reliable. For example, simply counting programming keywords in codes developed by students.

As a suggestion for future work, we think it would be of all convenience and useful to develop experimental work that meets the inclusion and exclusion criteria used in this meta-analysis, associated with constructivist teaching methods. Another idea that we take from this work is the reuse of dictionaries used as inclusion criteria to study other educational resources, even without being associated with technological tools. We also suggest that studies with a longer duration may be developed in the future, for example, a semester or an academic year, as most of those we find are of short and/or medium duration.

Funding

The work by Joana Martinho Costa and Sérgio Moro was supported by the Fundação para a Ciência e a Tecnologia (FCT) within the following [Projects: UIDB/04466/2020 and UIDP/04466/2020].

References

- Alice Project. (2020) *Alice – Tell Stories. Build Games. Learn to Program*, [online] Available at: <http://www.alice.org>
- Anderson, J. R. (1993) *Rules of the Mind*. Hillsdale, Erlbaum, NJ.
- Ausubel, D. (2000) *The Acquisition and Retention of Knowledge: A Cognitive View*, Dordrecht: Kluwer Academic Publishers.
- Baddeley, A. (1997) *Human Memory. Theory and Practice*, Psychology Press, East Sussex, UK.
- Bloom, B. (1956) *Taxonomy of Educational Objectives: The Classification of Educational Goals*, Allyn and Bacon, Boston, MA.
- Borenstein, M., et al., (2009) *Introduction to Meta-Analysis*, Wiley, Chichester, UK.
- Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000) *How People Learn: Brain, Mind, Experience and School*, Academic Press, Washington, DC.
- Bruner, J.S. (1965) *The growth of mind*. *American Psychologist*, 20(12), pp. 1007–1017
- Bruner, J. (1966) *Toward a Theory of Instruction*, The Belknap Press of Harvard University Press, Cambridge, MA.
- Coe, R. (2002) *It's the Effect Size, Stupid: What Effect Size Is and Why It Is Important*, [online] Available at: <http://www.leeds.ac.uk/educol/documents/00002182.htm>
- Cooper, H. & Hedges, L. (2009) 'Research synthesis as a scientific process', in *The Handbook of Research Synthesis and Meta-Analysis*, eds H. Cooper, L. Hedges & J. Valentine, Russel Sage Foundation, New York, pp. 4–15.
- Cortez, P., et al., (2018) 'Insights from a text mining survey on expert systems research from 2000 to 2016', *Expert Systems*, vol. 35, no. 3, p. e12280. doi: 10.1111/exsy.12280
- Costa, J. M. (2019) 'Microworlds with different pedagogical approaches in introductory programming learning: effects in programming knowledge and logical reasoning', *Informatica*, vol. 43, pp. 145–174. doi: 10.31449/inf.v43i1.2657
- Costa, J. M. & Miranda, G. L. (2017) 'Relation between Alice software and programming learning: a systematic review of the literature and meta-analysis', *British Journal of Educational Technology*, vol. 48, no. 6, pp. 1464–1474. doi: 10.1111/bjet.12496
- Costa, J. M. & Miranda, G. L. (2019) 'Using Alice software with 4C-ID model: effects in programming knowledge and logical reasoning', *Informatics in Education*, vol. 18, no. 1, pp. 1–15. doi: 10.15388/infedu.2019.01
- De Corte, E. (1993) 'Toward embedding enriched LOGO-based learning environments in the school curriculum: retrospect and prospect', *Proceedings of the Fourth European LOGO Conference Greece*, Athens: University of Athens, pp. 28–31.

- De Jong, T. & Lazonder, A. W. (2014) 'The guided discovery learning principle in multimedia learning', in *The Cambridge Handbook of Multimedia Learning*, ed R. E. Mayer, 2nd edn, Cambridge University Press, Cambridge, pp. 371–390.
- DiCerbo, K. E., et al., (2010) 'Individual practice and collaborative inquiry: instructional configurations in a simulation environment', *Sixth International Conference on Networking and Services*, Cancun, Mexico: IEEE, pp. 335–339.
- Fargas-Marques, A. & Costa-Castelló, R. (2006) 'Using interactive tools to teach and understand MEMS', *IFAC Proceedings Volumes*, vol. 39, no. 6, pp. 589–594. doi: 10.3182/20060621-3-ES-2905.00101
- Gagné, R. (1985) *The Conditions of Learning and Theory of Instruction*, Holt, Rinehart and Winston, New York.
- Hattie, J. (2009) *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*, Routledge, London, UK.
- Hedges, L. & Pigott, T. (2004) 'The power of statistical tests for moderators in meta-analysis', *Psychological Methods*, vol. 4, pp. 426–445. doi: 10.1037/1082-989X.9.4.426
- Hedges, L., Shymansky, J. & Woodworth, G. (1989) *Modern Methods of Meta-Analysis: A Practical Guide*, National Science Teachers Association, Washington, DC.
- Hoyle, C., Noss, R. & Adamson, R. (2002) 'Rethinking the microworld idea', *Journal of Educational Computing Research*, vol. 27, no. 1, pp. 29–53. doi: 10.2190/U6X9-0M6H-MU1Q-V36X
- Jenkins, C. (2015) 'A work in progress paper: evaluating a microworlds-based learning approach for developing literacy and computational thinking in cross-curricular contexts', *Proceedings of the Workshop in Primary and Secondary Computing Education*, Potsdam Germany: ACM, pp. 61–64.
- Kilpatrick, W. (2007) *O método do projeto [The Project Method]*, Pedago Editions, Viseu.
- Kirsner, P. (2019) *Constructivist Pedagogy Is Like a Zombie that Refuses to Die. An Interview with Professor Paul A. Kirschner by Isak Skogstad*, [online] Available at: <https://3starlearningexperiences.wordpress.com/2019/03/26/constructivist-pedagogy-is-like-a-zombie-that-refuses-to-die/>
- Kluge, A. (2008) 'What you train is what you get? Task requirements and training methods in complex problem-solving', *Computers in Human Behavior*, vol. 24, no. 2, pp. 284–308. doi: 10.1016/j.chb.2007.01.013
- Lehrer, R., Lee, M. & Jeong, A. (1999) 'Reflective teaching of logo', *The Journal of the Learning Sciences*, vol. 8, no. 2, pp. 245–289. doi: 10.1207/s15327809jls0802_3
- Littlefield, J., et al., (1988) 'Learning LOGO: method of teaching, transfer of general skills, and attitudes toward school and computers', in *Teaching and Learning Computer Programming. Multiple Research Perspectives*, ed R. E. Mayer, Erlbaum, Hillsdale, NJ, pp. 111–135.
- Mayer, R. (2014) *The Cambridge Handbook of Multimedia Learning*, 2nd edn, Cambridge University Press, Cambridge, MA.
- McDougall, A. (2002) 'Technology-supported environments for learning through cognitive conflict', *Research in Learning Technology*, vol. 10, no. 3, pp. 83–91. doi: 10.1080/0968776020100307
- Mendelsohn, P. (1991) 'LOGO: Quest-ce qui se développe', in *LOGO et apprentissages*, eds J.-L. Gurtner & J. Retschitzki, Delachaux et Niestlé, Neuchâtel, pp. 29–37.
- Merrill, M. D. (2002) 'First principles of instruction', *Educational Technology Research and Development*, vol. 50, no. 3, pp. 43–59. doi: 10.1007/BF02505024
- Moons, J. & De Backer, C. (2013) 'The design and pilot evaluation of an interactive learning environment for introductory programming influenced by cognitive load theory and constructivism', *Computers & Education*, vol. 60, no. 1, pp. 368–384. doi: 10.1016/j.compedu.2012.08.009
- Moro, S., et al., (2019) 'A text mining and topic modelling perspective of ethnic marketing research', *Journal of Business Research*, vol. 103, pp. 275–285. doi: 10.1016/j.jbusres.2019.01.053

- Nelson, L. D., Simmons, J. & Simonsohn, U. (2018) 'Psychology's renaissance', *Annual Review of Psychology*, vol. 69, pp. 511–534. doi: 10.1146/annurev-psych-122216-011836
- Oliveira, O. L., Monteiro, A. M. & Roman, N. T. (2013) 'Can natural language be utilized in the learning of programming fundamentals?', *IEEE Frontiers in Education Conference*, Oklahoma: IEEE, pp. 1851–1856. doi: 10.1109/FIE.2013.6685157
- Papert, S. (1980) *Mindstorms: Children, Computers and Powerful Ideas*, Basic Books, Inc., New York.
- Papert, S., et al., (1979) *Final Report of the Brookline LOGO Project – Part II: Project Summary and Data Analysis*, LOGO Memo N.º 53, MIT – AIL, Cambridge, MA.
- Pea, R. & Kurland, D. M. (1984) 'On the cognitive effects of learning computer programming', *New Ideas in Psychology*, vol. 2, no. 2, pp. 137–168.
- Pfahl, D., Koval, N. & Ruhe, G. (2001) 'An experiment for evaluating the effectiveness of using a system dynamics simulation model in software project management education', *Seventh International Software Metrics Symposium*, London, England: IEEE, pp. 97–109. doi: 10.1109/METRIC.2001.915519
- Rieber, L. P. (1996) 'Seriously considering play: designing interactive learning environments based on the blending of microworlds, simulations, and games', *Educational Technology Research & Development*, vol. 44, no. 2, pp. 43–58. doi: 10.1007/BF02300540
- Santos, M. R., Laureano, R. M. & Moro, S. (2019) 'Unveiling research trends for organizational reputation in the nonprofit sector', vol. 31, *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, pp. 56–70. doi: 10.1007/s11266-018-00055-7
- Sarama, J. & Clements, D. H. (2002) 'Building blocks for young children's mathematical development', *Journal of Educational Computing Research*, vol. 27, no. 1, pp. 93–110. doi: 10.2190/F85E-QQXB-UAX4-BMBJ
- Schneider, B., et al., (2013) 'Preparing for future learning with a tangible user interface: the case of neuroscience', *IEEE Transactions on Learning*, vol. 6, no. 2, pp. 117–129. doi: 10.1109/TLT.2013.15
- Schofield, J. W. (1995) *Computers and Classroom Culture*, Cambridge University Press, Cambridge.
- Scratch. (2020) *Scratch – Imagine, Program, Share*, [online] Available at: <https://scratch.mit.edu/>
- Simonsohn, U., Nelson, L. D. & Simmons, J. P. (2014a) 'P-curve: a key to the file-drawer', *Journal of Experimental Psychology: General*, vol. 143, no. 2, p. 534. doi: 10.1037/a0033242
- Simonsohn, U., Nelson, L. D. & Simmons, J. P. (2014b) 'P-curve and effect size: correcting for publication bias using only significant results', *Perspectives on Psychological Science*, vol. 9, no. 6, pp. 666–681. doi: 10.1177/1745691614553988
- Skinner, B. (1954) 'The science of learning and the art of teaching', *Harvard Educational Review*, vol. 24, no. 2, pp. 86–97.
- Somekh, B. (1996) 'Designing software to maximize learning', *Research in Learning Technology*, vol. 4, no. 3, pp. 4–16. doi: 10.3402/rlt.v4i3.9974
- Sweller, J., Ayres, P. & Kalyuga, S. (2011) *Cognitive Load Theory*, Springer, New York.
- Van den Noortgate, W., et al., (2013) 'Three-level meta-analysis of dependent effect sizes', *Behavior Research Methods*, vol. 45, no. 2, pp. 576–594. doi: 10.3758/s13428-012-0261-6
- van Merriënboer, J. (1997) *Training Complex Cognitive Skills: A Four-Component Instructional Design Model for Technical Training*, Educational Technology Publications, Englewood Cliffs, NJ.
- Vosinakis, S., Anastassakis, G. & Koutsabasis, P. (2018) 'Teaching and learning logic programming in virtual worlds using interactive microworld representations', *British Journal of Educational Technology*, vol. 49, no. 1, pp.30–44. doi: 10.1111/bjet.12531
- Wang, S. Y., et al., (2018) 'A microworld-based role-playing game development approach to engaging students in interactive, enjoyable, and effective mathematics learning', *Interactive Learning Environments*, vol. 26, no. 3, pp. 411–423. doi: 10.1080/10494820.2017.1337038

- Yuen-Kuang, C. & Brigh, G. (1991) 'Effects of computer programming on cognitive outcomes: a meta-analysis', *Journal of Educational Computing Research*, vol. 7, no. 3, pp. 251–268. doi: 10.2190/E53G-HH8K-AJRR-K69M
- Yurtseven, M. K. & Buchanan, W. W. (2012) 'Educating undergraduate students on systems thinking and system dynamics', *Proceedings of PICMET'12: Technology Management for Emerging Technologies*, Vancouver, British Columbia, Canada: IEEE, pp. 1837–1844.
- Zohar, A. & Peled, B. (2008) 'The effects of explicit teaching of metastrategic knowledge on low-and high-achieving students', *Learning and Instruction*, vol. 18, no. 4, pp. 337–353. doi: 10.1016/j.learninstruc.2007.07.001

Appendix A. Codification of the data.

Type of data	Description
Study	Study authors and release date
Country	Country in which the study was done
<i>N</i>	Total number of participants
Area	Curricular area in which the study was conducted (law, informatics, arts, mathematics, natural sciences, medicine)
Design	Experimental, quasi-experimental or non-experimental
Measure criteria	Analysed dependent variables
Teaching method	Teaching method used during the study
Group criteria	Grouping criteria of the CG and EG
Grade	Grade where the study was conducted
DCI	Data collection instruments
Time	Duration of the study
<i>d</i>	Effect size using the Cohen's <i>d</i>

Appendix B. Characteristics of the selected studies.

Study	Country	N	Area	Design	Measure criteria	Teaching method	Grade	DCI	Time	d
Jenkins (2015)	USA	51	Drama	Quasi-experimental (pre/post-test) with non-equivalent comparison group design	Literacy knowledge	Expository	K-12	Literacy test	3 months	0.62
Oliveira, Monteiro, and Roman (2013)	Brazil	22	Computer science	Quasi-experimental	Programming knowledge	Expository	University	Programming test	1 day	0.08
DiCerbo <i>et al.</i> (2010)	USA	97	Computer science	Quasi-experimental	Networking knowledge	Expository	University and high school	Networking test	1 day	0.3
Costa (2019)	Authors	22	Computer science	Quasi-experimental	Programming knowledge	Expository	High school	Programming test	1.5 months	-0.43
Pfahl, Koval, and Ruhe (2001)	Germany	9	Computer science	Quasi-experimental	Software project knowledge	Expository	University	Software project	3 days	0.45
Schneider <i>et al.</i> (2013)	USA	28	Neurosciences	AB/BA cross over design	Neurosciences knowledge	Cognitive	University	Neurosciences test	1 day	1.51
Lehrer, Lee, and Jeong (1999)	USA	50	Computer science	Quasi-experimental	Programming knowledge	Cognitive	K-12	Logo test	1 month	1.24
Moons and Backer (2013)	Belgium	39	Computer science	Quasi-experimental	Programming knowledge	Cognitive	University	Recursion test	Five lessons of 4 h each	0.89
Kluge (2008)	Germany	183	Chemistry	3×3 factorial control-group design	Problem solving	Cognitive	University	Problem-solving test	1 day	0.66
Zohar and Peled (2007)	Israel	41	Biology (seed germination)	2×2 design	Strategic knowledge	Cognitive	K-12	Strategic knowledge interview	Nine sessions, with 30 min each	1.2