Learning Environments Research
https://doi.org/10.1007/s10984-021-09373-y

ORIGINAL PAPER

Four-component instructional design (4C/ID) model: a meta-analysis on use and effect

Joana Martinho Costa[1] · **Guilhermina Lobato Miranda**[2] · **Mário Melo**[3]

Received: 10 July 2020 / Accepted: 7 June 2021

Abstract
The four-component instructional design model (4C/ID) has been increasingly used in face-to-face and online learning environments. We present a meta-analysis on the use and effect on performance of educational programs developed with the 4C/ID model after more than 20 years of its application and research in different academic areas and technical training. We performed the meta-analysis through the combination of effect sizes from studies using Cohen's d. The combination of the studies suggests that the use of educational programs developed with 4C/ID has a high impact on performance ($d = 0.79$ standard deviations), regardless of the academic area, the design of the study and the outcome (knowledge and complex skills). The grade under study was a significant moderator on the effect, showing that the higher-education level is more suitable for application of the 4C/ID model. Our results suggest that the use of the 4C/ID model should be prioritized as an instructional model in college and university learning environments.

Keywords 4C/ID model · Complex skills · Instructional design · Meta-analysis · Performance

Introduction

Instructional Design is the domain of human activity that concerns learning and performance improvement (Merrill, 2002), it aims to optimize learning rather than describe or explain it, and therefore it has a prescriptive and normative orientation (Bruner, 1966). Reiser (2001) highlights that instructional design is more than the use of certain resources and technologies; its purpose is to improve performance. Recently, Reigeluth et al. (2017) gave a vision of changes taking place in instruction theory, arising from the technological society from the Information Age, that contribute to the shift from teacher-centered to learner-centered models. These changes have had repercussions in several domains associated with the instructional

✉ Joana Martinho Costa
joana.martinho.costa@iscte-iul.pt

1 Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal

2 Instituto de Educação, Universidade de Lisboa, Lisbon, Portugal

3 Instituto de Educação, UIDEF, Universidade de Lisboa, Lisbon, Portugal

design: instructional management, assessment and even curriculum. According to these authors, the development of the theory of instruction arose during the 1980s when systemic or process models predominated, such as the ADDIE model (Analysis, Design, Development, Implementation, Evaluation) which still is used nowadays. During the 1990s, based on the urgency of a new paradigm and its different theories, research centered on the learner rather than on the teacher (Reigeluth et al., 2017). From 2009, academics tried to find a common language and knowledge in this field without replacing the different models and instructional theories that emerged mainly during the 1990s.

Each instructional theory is associated with a scientific theory of cognitive development or human learning that explains how we acquire and transfer knowledge (Bruner, 1966). Although instructional design as a disciplinary field emerged with Skinner's (1954) Programmed Instruction in the early fifties, today there are three major approaches to instructional design: instructivism, rooted in the behaviorist theories of learning; constructivism, rooted in cognitive developmental psychology; and cognitivism, rooted in information process theories or cognitive psychology.

Cognitive models are based on experimental research that is developed within information processing psychology or cognitive psychology, with memory being the main mechanism that allows human beings to learn (Anderson, 1983, 1993; Baddeley, 1997; Nisser 1967). They value the processes of knowledge construction that follow the way memory works: a working memory with limited capacity (Baddeley, 1997) which processes different stimuli through the sensorial systems and transform them into mental representations. This process makes the connection between this new knowledge and previously-acquired knowledge organized in longterm memory in mental schemas. In human cognition framework, learning means acquiring new mental schemas, integrating others already acquired into broader schemas and automating other schemas (e.g. Anderson, 1983, 1993; Blessing & Anderson, 1996).

One of the cognitive instructional models that is currently used in many scientific areas as a dimension of powerful learning environments is van Merriënboer's Four-Component Instructional Design Model (4C/ID) (van Merriënboer & Paas, 2003). According to its authors, the 4C/ID model is an instructional-design model mainly suitable for teaching complex learning (Frerejean et al., 2019). It is based on the cognitive psychology or information-processing approach to cognition in which memory plays a central role in the process of learning. The 4C/ID model also involves a task-centered instructional design that the traditional cognitive models, such as the conditions-based model of Gagné, do not share. According to Merrill (2002), this model is "perhaps, the most comprehensive recent model of instructional design that is problem-centered" (p. 56) as it involves all of the first principles of instruction.

The 4C/ID model allows the development of powerful learning environments that integrate some of the characteristics of categories of learning environments proposed by Bransford et al. (2000): student-centered, knowledge-centered, assessment-centered and community-centered. For Bransford et al. (2000), learning environments that manage to bring all these perspectives together are the most effective for learning, both inside and outside school.

4C/ID model

The 4C/ID model was developed by van Merriënboer and his colleagues in the early 1990s. Its main objective is to support the teaching of complex learning that requires the integration of knowledge, complex skills and attitudes to solve real-world problems (van Merriënboer & de Croock, 1992; van Merriënboer & Kirschner, 2001). This model covers the current knowledge about human cognitive architecture, the limitations of Working

Memory, and a set of principles proposed by the cognitive theory of multimedia learning (e.g. Mayer, 2014) and the cognitive load theory (e.g. Sweller et al., 2011, 2019), and it aims to promote the acquisition and transfer of knowledge (Kester, Kirschner & van Merriënboer 2004).

Van Merriënboer et al. (1992) made the first presentation of the 4C/ID model to the scientific community responsible for training in complex skills. They described the different stages of implementing the instructional model for training complex skills (design, presentation of information to support learning and training). Their work also presented a practical application of the model to the development of professional skills in distinct areas, such as engineering, programming and statistics. The authors predicted that the application of the 4C/ID model would produce more-reflective knowledge and better performance in learning transfer (van Merriënboer et al., 1992). Following this work, van Merriënboer et al. (2002) described a framework for the application of the 4C/ID model for complex competence training, suggesting the structure of the training plan and the associated instructional methods.

The 4C/ID model is composed of four interconnected components (Fig. 1): (a) *learning tasks*, which are the central feature of the model, are preferentially based on real-world problems; (b) *supportive information,* that helps the learner to solve more efficiently the problems defined in the learning tasks, connecting previously-acquired knowledge to the newly-presented knowledge, (c) *procedural information,* which provides information about how the routine aspects of the tasks should be performed; this information is organized into small segments of information presented at the exact moment they are needed – just-in-time information; and (d) *part-task practice* that allows the learner to practise routine skills that are part of the *learning tasks* and should become fully automated.

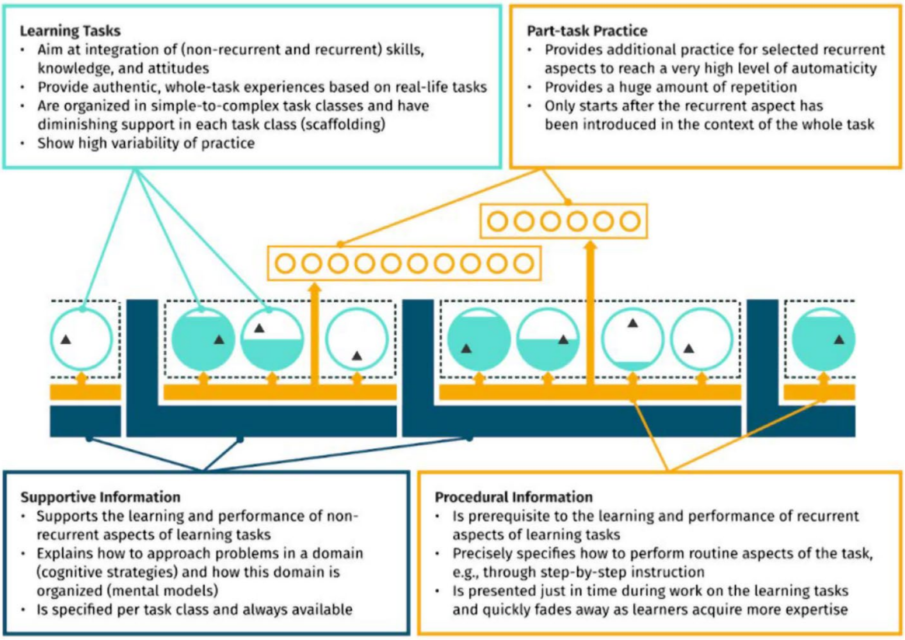

Fig. 1 4C/ID model (reprinted from 4CID.org, 2020)

In the 4C/ID model, it is critical to manage learners' cognitive load carefully as it focuses on whole-task learning tasks. Because of the complexity of this kind of learning tasks, it is crucial to accurately manage the cognitive load imposed on the learners.

In order to manage these types of cognitive load, the 4C/ID model suggests the following specific strategies in terms of task sequencing and information presentation:

(a) *Sequence learning tasks from simple to complex*. Task complexity could compromise the learning process because of the excessive cognitive overload for learners. Thus, in order to fix this problem, the first task class should be the simplest version of the whole tasks that the learners encounter in the work context. The final task class should be the most complex, including the real-world tasks.
(b) *Sequence learning tasks with decreasing learner support (fading from high to no support)*. Another way to prevent learners' cognitive overload is to use various types of learning tasks ranging from high built-in-support, via tasks with an intermediate level of support, to conventional tasks without support (e.g., worked examples and completion tasks).
(c) *Sequence learning tasks in a variable order (high contextual interference)*. Research indicates that high variability of practice affects the development of schemata and promotes subsequent learning transfer.
(d) *Present supportive information before learners start working on the learning tasks, and make it accessible to learners during the practice*, because this information involves mental models (e.g., what is this?, how is this organized, how does this work?) and cognitive strategies, with high complexity. So, performing practice and studying supportive information simultaneously can cause cognitive overload (Kester et al., 2001).
(e) *Present just-in-time information when learners need it*, to reduce ineffective cognitive load caused by temporal split attention (Mayer & Sims, 1994).

Research on the 4C/ID model

The research carried out under the 4C/ID model has been developed in two main areas. First, one area is focused on improving the applicability conditions of the learning environment that integrate the model, such as the format and timing of the presentation of supportive information, how to organize the learning tasks or how to sequence the task classes (e.g. Nadolskyi et al. 2004, 2006; Paas, 1992). Second, the other area is related to the efficiency of the model for the acquisition of knowledge and its transfer to other tasks that are more or less distant from the previous ones (e.g. Martínez-Mediano & Losada, 2017; van Es & Jeuring, 2017) or even for combining performance (acquisition and transfer) with the perceived mental effort (e.g. Melo & Miranda, 2015; Sarfo & Elen, 2007). We carried out a meta-analysis of the above studies in an attempt to understand the effects of the whole model on the acquisition and transfer of knowledge.

Method

To understand the effects of the 4C/ID model on learning, we developed a systematic review following the PRISMA statement (Moher et al., 2009).

Information sources and search

We searched for the keywords ("4C" and "instructional design") to include the abbreviations 4C/ID, 4CID and 4C-ID presented in the literature, or the full name of the model as the search term ("four-component instructional design") in the topic, abstract or title. We only searched for articles from 1992 because this was the year of the first publication of the model (van Merriënboer et al., 1992). We performed all database searches in the first week of May 2020.

The selected databases were the major ones that covered high-quality studies in the field, namely, Web of Science (WoS) of Clarivate Analytics, including all databases in the WoS Core Collection, Scopus, ERIC, DOAJ and IEEEXplore. We also looked for PhD theses in DigiNole using the same search keywords.

Eligibility criteria

After the search phase, the studies that satisfied the following criteria were included:

i. The study included the use of instructional materials based on the whole 4C/ID model. Studies that used only a segment of the 4C/ID model were excluded.
ii. The study presented quantitative empirical results.
iii. Results provided or allowed the calculation of effect size.
iv. Results were not provided from self-assessment.
v. The study measured knowledge or complex skills.

Figure 2 shows the flow of information through this systematic review based on Moher's et al. (2009) PRISMA diagram.

From the 127 records identified in searching databases (Scopus 50, DOAJ 4, IEEEXplore 7, WoS 54, DigiNole 3), 44 were repeated through databases and seven (mainly book chapters) were unavailable. We also searched the127 records' references identified in search databases. From the records' references, we identified two additional records that satisfied the inclusion criteria.

Statistical analysis

We used Cohen's *d* to calculate the effect size for the selected studies with the support of Lenhard and Lenhard's application (2016). The combination of the studies with effect sizes allows interpretation of studies through a single measure (Borenstein et al. 2009) and also explanation of the characteristics that are responsible for differences in the results (Coe, 2002).

Cohen's *d* represents an estimate of the number of subjects in the experimental group who are expected to exceed the medium value of the control group when represented by a proportion (Conboy, 2003). This measure has the advantage of using an unbiased estimative of true population variability, whether the null hypothesis is true or not. Also, it is less biased than Hedges' *g* (Conboy, 2003).

Additionally, we used a *p*-curve method to investigate whether the significant results present in the literature can be explained by phenomena such as *p*-hacking and publication

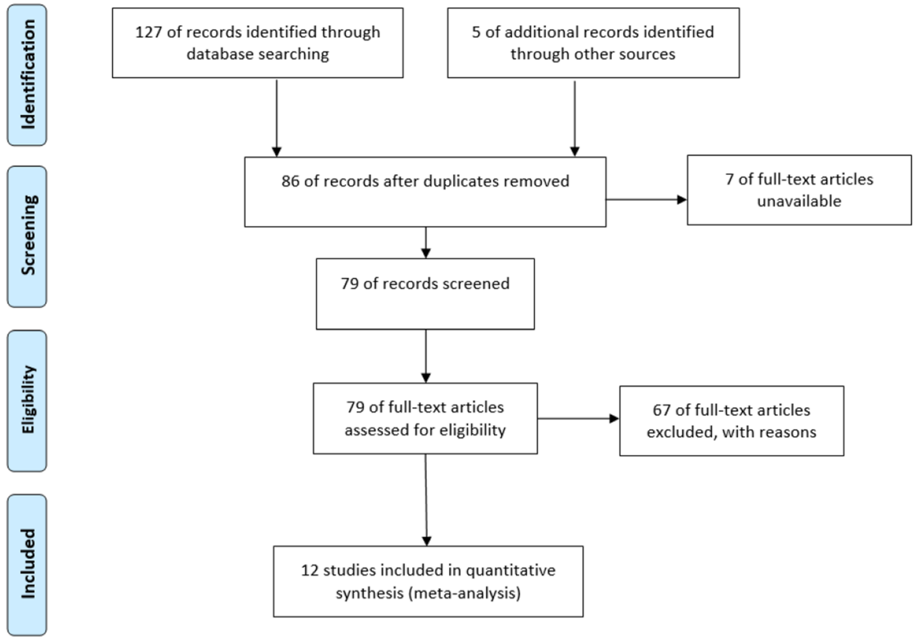

Fig. 2 PRISMA flow diagram through the 4C/ID model meta-analysis

bias, or if they signify real evidentiary value (Simonsohn et al. 2014). We used an improved version of *p*-curve analysis that is more robust to errors and certain kinds of *p*-hacking (Simonsohn et al. 2015).

Moderators

Considering the characteristics of the studies, we selected five possible moderators that could influence the results because they are all related to methodological aspects or participants' characteristics: duration of the intervention; academic area of the study context; design of the study; grade of the participants; and the outcomes under study. Outcomes were divided into knowledge, which involves the measurement of the acquired knowledge level and complex skills which can never be fully automated and rely on/encompass knowledge available in long-term memory. Table 1 represents the moderators characteristics: name, description, type and values.

Results

Table 2 presents a synthesis of the 12 selected studies. Columns EG and CG represent the number of subjects in the experimental and control groups, respectively. For the studies that had one single group, the N column provides the total number of participants. The Group column synthesizes the treatment given to experimental groups versus control groups and the procedures for the non-experimental groups.

Table 1 Moderators characteristics

Name	Description	Type	Values
Duration	Duration of intervention	Dichotomous variable	Short (less than one month)
			Long (more than one month)
Area	Academic area in which study was conducted	Categorical variable	The studies were categorized based on Fan and Chen's (2001) characterization for academic areas: Math/quantitative
			Reading/language arts, Science, Social studies, Other (e.g. music)
			Unspecified
Design	Design of study	Categorical variable	Experimental
			Quasi-experimental
			Non-experimental
Grade	Grade level of participants	Dichotomous variable	Higher education
			High school
Outcome	Dependent variable	Dichotomous variable	Knowledge
			Complex skills

From Table 2 we verify that all studies lasted between a day and a year. Therefore, the effect of the 4C/ID model in most cases has been measured with short experiences, but longer experience might have a greater impact on student learning. However, it is not clear whether the duration of previous studies refers to the experience as a whole or just the time devoted to the task. Because these data are relevant for analysis, it would be important for studies to distinguish both types.

The studies were conducted at three distinct grade levels: elementary school (van Es & Jeuring, 2017), high school (Flores, 2011; Melo & Miranda, 2014, 2015; Sarfo & Elen, 2007) and higher education (Kolcu et al., 2020; Lim & Park, 2012; Lim et al., 2009; Martínez-Mediano & Losada, 2017; Peng et al., 2017; Rosenberg-Kima, 2012). There were also some differences in the number of subjects in each study. Zhao et al. (2017) had a large sample of 160 students. On the other hand, Lim and Park (2012) had a sample of only 22 students. Because the process of meta-analysis considers the number of participants in calculating effect sizes, those differences were considered in the results.

From Table 2, we also verify that the studies were conducted in distinct domains, mostly in Computer Science (Lim & Park, 2012; Lim et al., 2009; Peng et al., 2017; Rosenberg-Kima, 2012; van Es & Jeuring, 2017), but also in Physics (Melo & Miranda, 2014, 2015), Mathematics (Flores, 2011), Engineering (Martínez-Mediano & Losada, 2017; Sarfo & Elen, 2007) and Health (Kolcu et al., 2020, Zhao et al. 2017). In the next section (Sensitivity Analysis), we consider whether these differences influence the effect size.

More than half of the studies were conducted with young adults in a higher education environment, with the remaining involving youth and children. All studies used tests to quantify the performance of the subjects, but the dependent variable or outcome was skill acquisition (both complex and procedural) or knowledge. Because most 4C/ID instruction was created for learning complex skills, the Sensitivity Analysis section below considers whether these differences in the outcome under study bias the effect size.

We also consider some differences in the design of each study which are analyzed by a sensitivity analysis in the next section. The studies were conducted under a quasi-experimental design (Flores, 2011; Lim & Park, 2012; Lim et al., 2009; Martínez-Mediano & Losada, 2017; Melo & Miranda, 2014, 2015; Sarfo & Elen, 2007; van Es & Jeuring, 2017; Zhao et al. 2017) or a non-experimental design (Kolcu et al., 2020; Peng et al., 2017; Rosenberg-Kima, 2012).

Table 3 provides the effect size for the 12 selected studies, their confidence interval (95%) and their weight in the meta-analysis. Table 3 and Fig. 3 present the effect size of the selected studies. These results suggest a large effect for the use of 4C/ID model on knowledge acquisition ($d = 0.79$ [0.50; 1.08]. According to Coe (2002), between 76 and 79% of the subjects in the experimental group can be expected to perform better than the average of the control group.

However, a heterogeneity test revealed a high degree of heterogeneity ($I^2 = 78.6\%$ [63.1%; 87.6%]; $Q = 51.32$; $p < 0.0001$) which means that the moderators could have had some effect (Borenstein et al. 2009).

We tested each moderator to check whether they were related to effect size differences (Borenstein et al. 2009). To achieve this goal, we checked the statistical significance of each moderator by performing a *t*-test on the β-weight of each moderator (Harrer et al., 2019).

To ensure that our meta-regression was robust, we checked intercorrelations between the moderators to identify any high correlations (Table 4). Figure 4 shows that Duration and Grade moderators are highly correlated. A correlation being significant does not establish the presence of a latent variable. Therefore, we decided to exclude the duration of the

Table 2 Characteristics of the selected studies

Study	EG	CG	N	Age	Duration	Type	Area	Country	Group	Design	Grade	Outcome	d
Flores (2011)	16	19	35	16	1 day	PhD thesis	Math	USA	EG: adaptative tutorial designed with 4C/ID CG: Non-adaptative tutorial	Quasi-experimental	High school	Performance (problem solving skills)	$F = 1.16$ $d = 0.376$
Kolcu et al. (2020)	–	–	26	25	21 days	Paper	Dentistry	Turkey	Participants learned endodontics procedures in a distance-education program based on 4C/ID	Non-experimental (single group with pre-test and posttest)	Higher education	Performance (psychomotor skills)	t-test = 5.97 $d = 1.26$
Lim and Park (2012)	12	10	22	30	1 semester	Conference paper	Computer Science	USA	EG: Whole task (instruction designed on the 4C/ID model) CG: Part task	Quasi-experimental	Higher education	Performance (knowledge test: preparing a grade test using data provided)	$t\,(20) = 2.28$ $d = 0.976$

Table 2 (continued)

Study	EG	CG	N	Age	Duration	Type	Area	Country	Group	Design	Grade	Outcome	d
Lim et al. (2009)	25	26	51	21	1 day (2 × 60 min)	Paper	Computer Science	USA	EG: Whole task—(instruction was designed based on the 4C/ID model) CG: Part task (traditional instructions)	Quasi-experimental	Higher education	Performance (complex skills)	EG: 30.9 (4.36) CG: 24.6 (6.67) $F(1, 47) = 15.87$ $d = 1.114$
Martínez-Mediano and Losada (2017)	20	20	40	18–20	1 semester	Paper	Engineering (Physics)	Bulgaria	EG: Used platform with system based on 4C/ID CG: Traditional classroom environment	Quasi-experimental	Higher education	Performance knowledge	EG: 4.59 (0.96) CG: 3.58 (1.1) $d = 0.983$
Melo and Miranda (2015)	78	51	129	14	2 weeks (90 min/week)	Paper	Physics	Portugal	EG: Classes based on 4C/ID CG: Classes based on expository teaching	Quasi-experimental	High school	Performance (transfer test)	EG: 11.2 (1.9) CG: 8.8 (2.6) $d = 1.09$

Table 2 (continued)

Study	EG	CG	N	Age	Duration	Type	Area	Country	Group	Design	Grade	Outcome	d
Melo and Miranda (2014)	81	50	131	14	2 weeks (90 min/week)	Conference Paper	Physics	Portugal	EG: Classes based on 4C/ID CG: Classes based on expository teaching	Quasi-experimental	High school	Performance (transfer test)	EG: 0,55 (0.87) CG: −0.42 (1.01) $d = 1.48$
Peng et al. (2017)	–	–	29	3-year college students	5 weeks	Conference Paper	Computer Science	China	Participants learned computer programming through system based on 4C/ID	Non-experimental (single group with pre-test and posttest)	Higher education	Performance (knowledge test)	Pretest: 46.34 (17.29) Post-test: 53.31 (15.38) $d = 0.43$
Rosenberg-Kima (2012)	31	33	64	21	1 day (105 min)	PhD thesis	Computer Science	USA	EG: Learned with task-centered instruction (4C/ID) CG: Learned with topic-centered instruction (traditional)	Non-experimental (single group with pre-test and posttest)	Higher education	Performance (skill development)	EG: 8.96 (2) CG: 7.51 (2.87) $F(1, 63) = 5.442$ $d = 0.58$

Table 2 (continued)

Study	EG	CG	N	Age	Duration	Type	Area	Country	Group	Design	Grade	Outcome	*d*
Sarfo and Elen (2007)	41	41	82	18	6×90 min	Paper	Engineering	Ghana	EG: 4C/ID learning environment without ICT CG: traditional method	Quasi-experimental	High school	Performance	EG: 8.84 (3.12) CG: 5.44 (3.46) $d = 1.03$
van Es and Jeuring (2017)	75	55	129	9–12	3–5 weeks (5 lessons)	Conference Paper	Computer Science (programming)	Netherlands	EG: Learned through 4C/ID instruction CG: Learned through constructionism instruction	Quasi-experimental	Elementary school	Performance (programming skills)	$t\,(128) = 1.123$ $d = 0.019$

Table 2 (continued)

Study	EG	CG	N	Age	Duration	Type	Area	Country	Group	Design	Grade	Outcome	*d*
Zhao et al. (2017)	80	80	160	N/A	1 year	Conference paper	Health (Nursing)	China	EG: Operating room nurse training with 4C/ID CG: Operating Room Nurse training with conventional training	Quasi-experimental	Higher education	Performance (knowledge test)	Chi-square = 5.0 $d = 0.34$

Table 3 Effect sizes for the 12 studies

Study	*d*	95% CI	%W (random)
Flores (2011)	0.4100	− 0.2368; 1.0568	6.9
Kolcu et al. (2020)	1.2600	0.6720; 1.8480	7.5
Lim and Park (2012)	0.9800	0.1176; 1.8424	5.2
Lim et al. (2009)	1.1100	0.5220; 1.6980	7.5
Martinez-Mediado and Losada (2017)	0.9900	0.3432; 1.6368	6.9
Melo and Miranda (2015)	1.0900	0.7176; 1.4624	9.6
Melo and Miranda (2014)	1.4800	1.0880; 1.8720	9.4
Peng et al. (2017)	0.4300	− 0.0796; 0.9396	8.2
Rosenberg-Kima (2012)	0.5800	0.2272; 0.9328	9.8
Sarfo and Elen (2007)	1.0300	0.5792; 1.4808	8.8
van Es and Jeuring (2017)	0.0200	− 0.3132; 0.3532	10.0
Zhao et al. (2017)	0.3400	0.0264; 0.6536	10.2

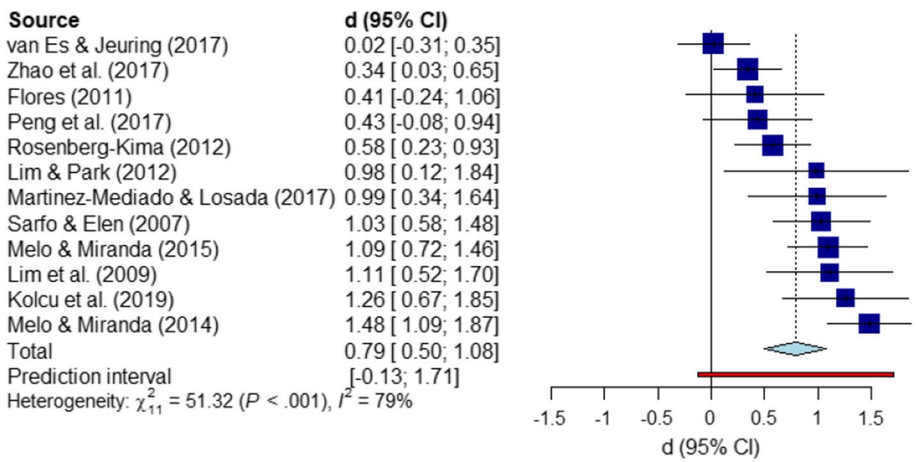

Fig. 3 Forest plot with the selected studies

intervention (Harrer et al., 2019) because, in most of the 12 studies, not being able to distinguish the time and frequency could lead to erroneous conclusions.

The moderator Area (where studies were conducted) did not have a significant influence on the effect size ($p > 0.05$). The moderator Outcome was also not statistically significant ($p > 0.05$), suggesting that 4C/ID is suitable for the development of both knowledge and skills (both complex and procedural). Also, the Design of the study did not demonstrate any evidence of significant influence on the effect size ($p > 0.05$).

On the other hand, the moderator Grade was statistically significant ($p < 0.05$) explaining 76.22% of the heterogeneity. To assess the fit of our model, we performed a permutation test for the significant moderator to ascertain whether we captured a real pattern in our analysis by re-calculating the *p*-values of the meta-regression (Harrer et al., 2019). The permutation test was significant ($p < 0.05$), suggesting that the grade of the participants indeed influences the effect size of the studies.

Table 4 Intercorrelation matrix for moderators

Moderator	Correlation				
	Duration	Area	Design	Grade	Outcome
Duration	1.00000000	0.1807016	0.09759001	0.6390644	0.2390457
Area	0.1807016	1.00000000	0.00000000	0.2245444	0.4724556
Design	0.09759001	0.00000000	1.00000000	−0.2425356	0.4082483
Grade level	0.63906444	0.2245444	−0.2425356	1.00000000	−0.2970443
Outcome	0.23904572	0.4724556	0.4082483	−0.2970443	1.00000000

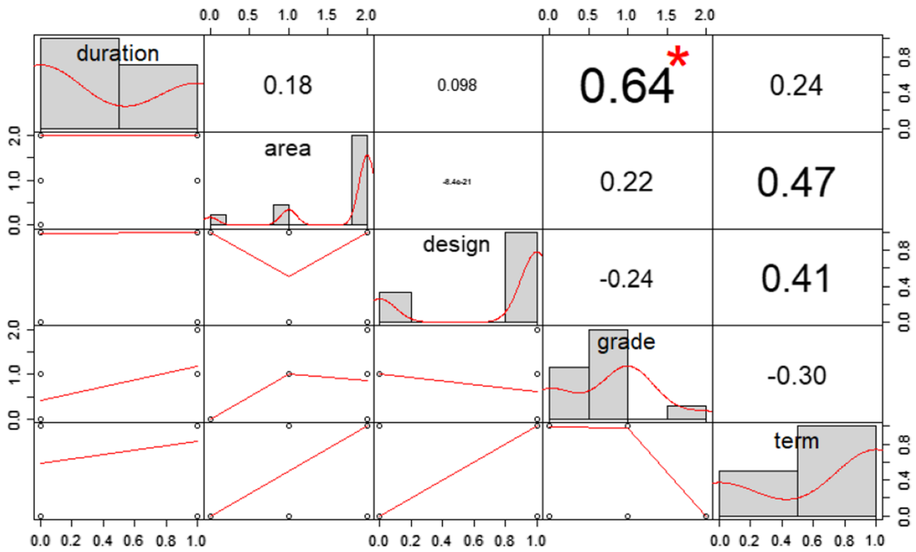

Fig. 4 Intercorrelation between moderators

Discussion

Our study aimed to understand the effects of educational programs developed with the 4C/ID model on student performance. Figure 3 and Table 3 reveal that there is a set of studies with effect sizes that are higher than average (Kolcu et al., 2020; Lim & Park, 2012; Lim et al., 2009; Martinez-Mediano and Losada 2017; Melo & Miranda, 2014, 2015; Sarfo & Elen, 2007). In this set of studies, the effect of using educational programs developed with the 4C/ID model produced more pronounced effects on the participants' performance. How can we explain this difference?

Sensitivity analysis suggested that differences between studies were due to the academic area in which the study was conducted, with the design and the outcome under study explaining only a low percentage of the heterogeneity. This means that they have a non-significant influence on the effect size, indicating that the 4C/ID model might be helpful for developing educational programs that have positive effects on both knowledge and complex skills in a vast set of academic areas. The difference between knowledge and skills and their integration in a competent action, proposed by the 4C/ID model, has its origins in the history of psychology. The cognitive psychology of information processing established the difference between declarative knowledge and procedural knowledge (cf Anderson, 1983), the former associated with semantic memory and the latter with procedural memory. Hence it can be concluded that a model like 4C/ID, which aims to answer the problem of integrating knowledge and complex skills in a given domain, could not fail to consider these two dimensions. Furthermore, it is this integration that facilitates the transfer of learning.

We also verified that the selected studies reported a correlation between their duration and the grade level of students. However, because the studies were not detailed enough about the duration of the intervention and the whole duration of the study, we could not be more detailed in the analysis. This limitation of our meta-analysis leads us to consider that studies in this area should be more accurate about the duration variable. In future work, an in-depth analysis of the effect of the duration of the intervention with 4C/ID should answer this issue.

In all studies except Martínez-Mediano and Losada (2017), the researchers used validated and reliable instruments to evaluate students' performances. In the study reported by Martínez-Mediano and Losada (2017), there was no reference to procedures for determining the reliability of the knowledge assessment test. However, because several experimental and control groups were set up over time during two semesters, it can be inferred that the researchers were careful and undertook some analysis of the reliability of the questionnaire used for performance evaluation.

The meta-analysis provides evidence that the grade level of students has a significant influence on the effect size, with a more-powerful effect in higher education than in lower grades. This reason explains most of the heterogeneity in this selection. This was the case in the study by van Es and Jeuring (2017) that had reported a lower effect size. This study was conducted in two different elementary schools with different curricular organizations, which might explain the magnitude of effects being close to zero. These results are in line with van Merriënboer's work (2019) in which he argues that the 4C/ID model is more suitable for designing instruction for professional learning. But these results also lead us to raise two questions. Are the lessons learned at these education levels less complex than those of certain professional activities if we take into account the children's level of development? Or is it that the model is not suitable for these learnings and should be adapted? Learning to read, write, calculate and think logically are complex and demand the integration of knowledge, attitudes and complex skills. They take time and a lot of training. Therefore, perhaps the lack of significant results in the aforementioned study are attributable to the non-adaptation of the model to these levels of education. But these questions can only be fully answered with further research.

Conclusion

The use of the 4C/ID model has increased in several areas and contributed to the enhancement of online (e.g. Frerejean et al., 2019; Sluijsmans et al., 2006) and face-to-face learning environments (e.g. Costa, 2019; Deep et al., 2020). Despite not having met the inclusion criteria mostly because of not having published data on performance effects related to the use of the whole model, there has been high-quality research on this model in several domains such as health and medical education (e.g. Maggio et al., 2015; Tjiam et al., 2012), information-problem solving (e.g. Wopereis et al., 2015) and higher education (e.g. Frerejean et al., 2019), as well as research on other possible related variables such as socioeconomic status (e.g. Costa & Miranda, 2019) or socio-demographic factors (e.g. Postma and White 2017).

For the future, longer-term studies should involve the complete model in learning environments in upper-secondary, vocational and higher education. This meta-analysis suggests that educational programs developed with the 4C/ID model have a significant impact on student performance in several academic domains, and therefore should be more widely used in universities for learning complex skills.

More-robust methodologies also will be required, especially quasi-experimental designs with pretest–posttest measures and experimental/control groups. The presence of measures that allow the calculation of effect sizes would make these meta-analyses more robust. Another important point is that researchers should use valid and reliable measuring instruments. By incorporating these aspects, it will become possible to gather more-detailed evidence on how specific design guidelines of the 4C/ID model contribute to the improvement of complex learning.

Acknowledgements This work was partially supported by Fundação para a Ciência e a Tecnologia, I.P. (FCT) [ISTAR Projects: UIDB/04466/2020 and UIDP/04466/2020].

References

4CID.org. (2020). Retrieved from https://www.4cid.org/home

Anderson, J. R. (1983). *The architecture of cognition*. Harvard University Press.

Anderson, J. R. (1993). *Rules of the mind*. Erlbaum.

Baddeley, A. (1997). *Human memory. Theory and practice* (revised). Psychology Press.

Blessing, S. B., & Anderson, J. R. (1996). How people learn to skip steps. *Journal Experimental Psychology: Learning, Memory, Cognition, 22*, 576–598.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Chichester, United Kingdom: Wiley.

Bransford, J., Brown, A., & Cocking, R. (2000). *How people learn: Brain, mind, experience and school* (expanded ed.). Washington, DC.: National Academy Press.

Bruner, J. S. (1966). *Toward a theory of instruction*. The Belnap Press of Harvard University Press.

Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important. In: *Paper presented at the British Educational Research Association Annual Conference, Exeter*.

Conboy, J. (2003). Algumas medidas típicas univariadas da magnitude do efeito. *Análise Psicológica, 2*(21), 145–158.

Costa, J. (2019). Microworlds with different pedagogical approaches in introductory programming learning: Effects in programming knowledge and logical reasoning. *Informatica, 43*, 145–174. https://doi.org/10.31449/inf.v43i1.2657

Costa, J., & Miranda, G. L. (2019). Using Alice software with 4C/ID model: Effects in programming knowledge and logical reasoning. *Informatics in Education, 18*(1), 1–15. https://doi.org/10.15388/infedu.2019.01

Deep, A., Murthy, S., & Bhat, J. (2020). Geneticus Investigatio: A technology-enhanced learning environment for scaffolding complex learning in genetics. *Research and Practice in Technology Enhanced Learning*. https://doi.org/10.1186/s41039-020-00145-5

Flores, R. (2011). *Examining the design and impact of adaptively faded worked examples on high school students´ mathematics problem solving skills*. Doctoral dissertation, Texas Tech University. Retrieved from: https://ttu-ir.tdl.org/handle/2346/ETD-TTU-2011-08-1755.

Frerejean, J., van Merriënboer, J. J. G., Kirschner, P. A., Roex, A., Aertgeerts, B., & Marcellis, M. (2019). Designing instruction for complex learning: 4C/ID in higher education. *European Journal of Education, 54*(4), 513–524.

Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). Doing meta-analysis in R: A hands-on guide. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/. Accessed 01 June 2020. https://doi.org/10.1111/ejed.12363

Kester, L., Kirschner, P. A., van Merriënboer, J. J. G., & Baumer, A. (2001). Just-in-time information presentation and the acquisition of complex cognitive skills. *Computers in Human Behaviour, 17*, 373–391.

Kester, L., Kirschner, P., & van Merriënboer, J. (2004). Information presentation and troubleshooting in electrical circuits. *International Journal of Science Education*, *26*(6), 239–256. https://doi.org/10.1080/69032000072809.

Kolcu, M. İB., Öztürkçü, Ö. S. K., & Kaki, G. D. (2020). Evaluation of a distance education course using the 4C-ID model for continuing endodontics education. *Journal of Dental Education, 84*(1), 62–71.

Lenhard, W., & Lenhard, A. (2016). *Calculation of effect sizes*. Retrieved from: https://www.psychometrica.de/effect_size.html. Dettelbach (Germany): Psychometrica. https://doi.org/10.13140/RG.2.2.17823.92329

Lim, J., & Park, S. (2012). An instructional method for competency-based e-learning: A whole-task approach. In P. Resta (Ed.), *Proceedings of Society for Information Technology and Teacher Education International Conference* (pp. 580–585). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

Lim, J., Reiser, R., & Olina, Z. (2009). The effects of part-task and whole-task instructional approaches on acquisition and transfer of a complex cognitive skill. *Educational Technology Research and Development, 57*, 61–77. https://doi.org/10.1007/s11423-007-9085-y

Maggio, L. A., Cate, O. T., Irby, D. M., & O'Brien, B. C. (2015). Designing evidence-based medicine training to optimize the transfer of skills from the classroom to clinical practice: Applying the four component instructional design model. *Academic Medicine, 90*, 1457–1461. https://doi.org/10.1097/ACM.0000000000000769

Martínez-Mediano, C., & Losada, N. R. (2017). Internet-based performance support systems in engineering education. *Revista Iberoamericana De Tecnologias Del Aprendizaje, 12*(2), 86–93. https://doi.org/10.1109/RITA.2017.2697778

Mayer, R. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*, 389–401.

Melo, M., & Miranda, G. L. (2014). Applying the 4C-ID model to the design of a digital educational resource for teaching electric circuits: Effects on student achievement. In H. M. Fardou, and J. A. Gallud (Eds). *Proceedings of the 2014 Workshop on Interaction Design in Educational Environments (IDEE'14)* (pp. 12–18). New York: ACM. https://doi.org/10.1145/2643605

Melo, M., & Miranda, G. L. (2015). Learning electrical circuits: The effects of 4C/ID instructional approach in the acquisition and transfer of knowledge. *Journal of Information Technology Education: Research, 14*, 313–337. https://doi.org/10.28945/2281

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development, 50*(3), 43–59.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*, 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Nadolskyi, R., Kirschner, P., & van Merriënboer, J. (2004). Optimizing the number of steps in learning tasks for complex skills. *British Journal of Educational Psychology, 75*, 223–237. https://doi.org/10.1348/000709904X22403

Nadolskyi, R., Kirschner, P., & van Merriënboer, J. (2006). Process support in learning tasks for acquiring complex skills in the domain of law. *Learning and Instruction, 16*, 266–278. https://doi.org/10.1016/j.learninstruc.2006.03.004

Nisser, U. (2014/1967). *Cognitive psychology*. New York: Psychology Press.

Paas, F. (1992). Training strategies for attaining transfer of problem solving skill in statistic: A cognitive load approach. *Journal of Educational Psychology, 84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429

Peng, J., Wang, M., & Sampson, D. (2017). Visualizing the complex process for deep learning with an authentic programming project. *Educational Technology and Society, 20*, 275–287.

Postma, T. C., & White, J. G. (2017). Socio-demographic and academic correlates of clinical reasoning in a dental school in South Africa. *European Journal of Dental Education, 21*(1), 58–65.

Reigeluth, C.M., Beatty, B.J., & Myers, R.D. (Eds.). (2017). *Instructional-design theories and models, Volume IV: The learner-centered paradigm of education.* New York: Routledge

Reiser, R. (2001). A history of instructional design and technology. Part 1: A history of instructional media. *Educational Technology Research and Development, 49*(1), 53–64.

Rosenberg-Kima, R. (2012). *Effects of task-centered vs. topic-centered instructional strategy approaches on problem solving: Learning to program in flash*. Doctoral dissertation, Florida State University). Available from http://fsu.digital.flvc.org/islandora/object/fsu%3A183092.
Sarfo, F., & Elen, J. (2007). Developing technical expertise in secondary technical schools: The effect of 4C/ID learning environments. *Learning Environments Research, 10*, 207–221. https://doi.org/10.1007/s10984-007-9031-2
Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: a key to the file-drawer. *Journal of experimental psychology: General*, *143*(2), 534. https://doi.org/10.1037/a0033242.
Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better P-curves: Making P-curve analysis more robust to errors, fraud, and ambitious P-hacking, a Reply to Ulrich and Miller (2015). *Journal of Experimental Psychology: General*, *144*(6), 1146–1152. https://doi.org/10.1037/xge0000104.
Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review, 24*(2), 86–97.
Sluijsmans, D., Prins, F., & Martens, R. (2006). The design of competency-based performance assessment in e-learning. *Learning Environments Research, 9*, 45–66. https://doi.org/10.1007/s10984-005-9003-3
Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.
Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292.
Tjiam, I. M., Schout, B. M., Hendrikx, A. J., Scherpbier, A. J., Witjes, J. A., & van Merriënboer, J. J. (2012). Designing simulator-based training: An approach integrating cognitive task analysis and four-component instructional design. *Medical Teacher, 34*(10), e698–e707.
van Es, N., & Jeuring, J. (2017). Designing and comparing two Scratch-based teaching approaches for students aged 10–12 years. *Proceedings of Koli Calling 2017*. New York: ACM. https://doi.org/10.1145/3141880.3141883.
van Merriënboer, J. J. (2019). *The four-component instructional design model: An overview of its main design principles*. Maastricht University.
van Merriëenboer, J. J., Clark, R. E., & De Croock, M. B. M. (2002). Blueprints for complex learning: The 4C/ID-model. *Educational Technology Research and Development, 50*(2), 39–64. https://doi.org/10.1007/BF02504993
van Merriënboer, J., & de Croock, M. (1992). Strategies for computer-based programming instruction: Program completion versus program generation. *Journal of Educational Computing Research, 8*, 365–394.
van Merriënboer, J., Jelsma, O., & Paas, F. (1992). Training for reflective expertise: A four-component instructional design model for complex cognitive skills. *Educational Technology Research and Development, 40*(2), 23–43. https://doi.org/10.1007/BF02297047
van Merriënboer, J. J. G., & Paas, F. (2003). Powerful learning and the many faces of instructional design: Towards a framework for the design of powerful learning environments. In E. De Corte, L. Verschaffel, N. Entwistle, & J. J. G. van Merriënboer, (Eds.), *Powerful learning environments: Unravelling basic components and dimensions* (pp. 3–20). Oxford, UK: Elsevier Science.
van Merriënboer, J., & Kirschner, P. (2001). Three worlds of instructional design: State of the art and future directions. *Instructional Science*, *29*, 429–441.
Wopereis, I., Frerejean, J., & Brand-Gruwel, S. (2015). Information problem solving instruction in higher education: A case study on instructional design. In S. Kurbanoglu, J. Boustany, S. Špiranec, E. Grassian, D. Mizrachi, & L. Roy (Eds.), *Information literacy: Moving toward sustainability* (Vol. 552, pp. 293–302). Springer.
Zhao, F., Wang, C., & Yin, R. (2017). The Application of 4C Training Modelin the Operating Room Nurse Training. *BIO Web of Conferences*, *8*. https://doi.org/10.1051/bioconf/201708010.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.