2024

The Effectiveness of Concept Maps on Students' Achievement in Science: A Meta-Analysis

Dimitris Anastasiou
Southern Illinois University Carbondale

Clare Nangsin Wirngo
Illinois State University, cnwirng@ilstu.edu

Pantelis Bagos
University of Thessaly

META-ANALYSIS

The Effectiveness of Concept Maps on Students' Achievement in Science: A Meta-Analysis

Dimitris Anastasiou[1] [ORCID] **· Clare Nangsin Wirngo[2] · Pantelis Bagos[3]**

Abstract

This study aims to evaluate the effectiveness of concept maps on science achievement among elementary and secondary education students, including low-achieving students. A systematic search located 55 studies about concept mapping in science achievement published in peer-reviewed journals and dissertations between 1980 and 2020. We extracted 58 independent standardized mean difference effect sizes from 55 eligible studies involving 5,364 students from Grade 3 to Grade 12 who used concept maps for learning in physics/earth science, chemistry, and biology that met the specified design criteria. A random-effects model meta-analysis revealed that the mean effect size was moderate for overall science ($g = 0.776$). The mean effect sizes varied from moderate to large based on the subject area ($g = 0.671$ for biology; $g = 0.590$ for chemistry; $g = 1.040$ for physics and earth science); these differences between groups were not statistically significant ($p = 0.220$). Concept maps were generally associated with increased science learning across several learning and teaching conditions, and methodological features (low-achieving students, higher teaching guidance, intermediate grades, low- or middle-income countries, journal publications, and late year of publication). However, we found significant heterogeneity in most subsets. Implications for future research and practice recommendations are discussed.

Keywords Concept maps · Effectiveness · Science achievement · Elementary and secondary education · Meta-analysis

Introduction

Concepts are fundamental to advancing and understanding biology, physics, and chemistry (Mayr, 1997). Ernst Mayr (1997), a twentieth century's leading evolutionary biologist and philosopher of biology, pointed out the central role of concepts in life and physical sciences. He stated that "[t]he two major contributors to a

---

new theory in the life sciences are the discovery of new facts (observations) and the development of new concepts" (p. 62).

Concept mapping is an instructional strategy usually used to enhance science achievement for elementary and secondary education students, who are the focus of the present study. Science achievement refers to vocabulary development, understanding, and reading comprehension of scientific texts. Concept mapping, a versatile tool, extends its utility beyond the realm of science to encompass various disciplines such as mathematics, social science, humanities, and business management. However, our meta-analysis centers specifically on the utilization of concept mapping within the realms of natural science, specifically in the fields of physics, chemistry, and biology.

A concept map is a graphical tool, an alternative to written text, for representing verbal information and illustrating and organizing a set of relationships among concepts (Nesbit & Adesope, 2013; Novak & Cañas, 2008; Novak, 2010a). A fundamental characteristic of a concept map is its structural and relational nature, establishing connections between various concepts (Davies, 2011). A concept map embodies a "hierarchical tree" in levels featuring superior and subordinate components, encompassing primary, secondary, and tertiary ideas (Davies, 2011). The more general concepts are often placed at the top or center of a concept map (Davies, 2011; Novak, 1991, 2010a; Novak & Cañas, 2008). Concepts are usually enclosed in circles or boxes, and relationships between concepts are illustrated by connecting lines; these lines contain words such as "is," "are," and "includes," or short prepositional phrases such as "leads to," "results from," "is part of," etc. linking two concepts showing their relationships (Davies, 2011; Novak, 1991, 2010a; Novak & Cañas, 2008). *Examples* are sometimes included to clarify the meaning of a concept (Davies, 2011). In the 55 studies of our meta-analysis, we observed that concept maps were typically hierarchically structured, with the central concept typically positioned at the top or center, perhaps because they were addressed to elementary, intermediate, and secondary education students, and this structure is more straightforward. However, we note that non-hierarchical types of concept maps (i.e., network concept maps) are also in use (Amadieu et al., 2009; Davies, 2011).

Mind mapping is similar to a concept mapping tool, defined as visual and nonlinear representations of ideas and their interconnectedness (Davies, 2011). Mind mapping and concept mapping differ in terms of precision and formality. Whereas mind maps are less formal and structured, concept maps are formal and generally more tightly structured (Davies, 2011).

Not surprisingly, concept maps have been extensively used in school settings as an instructional strategy to organize information into a graphic form, creating a visual representation of the text structure (Kwon & Cifuentes, 2007; Sturm & Rankin-Erickson, 2002). Elementary and secondary education students are more likely to understand and remember relationships among concepts when involved in the mapping process (Karpicke & Blunt, 2011; Novak & Cañas, 2008; Rewey et al., 1989). Thus, students can actively identify critical ideas and details in the new content (Carnine & Carnine, 2004; Marzetta et al., 2018). Concept maps have been used in several ways: (a) to present information to students, (b) for students to represent their knowledge in learning activities, and (c) to assess what students know (Nesbit

& Adesope, 2013). Concept mapping may enhance organizational cues for retrieving concepts and information from memory (Kwon & Cifuentes, 2007). Concept maps can also be *scaffolding tools* to relate new learning material to students' *prior knowledge* (Marzetta et al., 2018; Novak, 2010a).

Theories Explaining Concept Mapping Effects on Learning

Several theories can explain the effectiveness of concept maps. We will concentrate on the three learning and instructional theories that may offer a closer explanation of the learning outcomes associated with concept maps.

First, concept mapping was initially linked to *meaningful learning*, a concept formulated by Ausubel (1962) within the framework of his subsumption (assimilation) learning theory (Novak & Cañas, 2008; Novak, 1991, 2010a, 2010b). In the subsumption learning theory, meaningful learning is contrasted with rote learning (Ausubel, 1962). *Rote learning* occurs when new information is acquired by verbatim memorization and arbitrarily incorporated into prior knowledge (Novak & Gowin, 1984; Novak, 2010a). Instead, *meaningful verbal learning* refers to anchoring new knowledge (or concepts) to relevant prior knowledge and existing cognitive structure (Ausubel, 1962, 1963; Novak & Gowin, 1984; Novak, 2010a).*"Cognitive structure"* refers to hierarchically organized concepts in any field of knowledge (e.g., biology, physics, chemistry) that represent the residue of information and ideas (i.e., schemas in their long-term memory) (Ausubel, 1963; Novak & Gowin, 1984). Ausubel's primary hypothesis was that learning is facilitated, and retention can be maintained over time when knowledge is meaningfully acquired and hierarchically organized (Ausubel, 1963; Kirschner & Hendrick, 2020; Novak, 2010a). The organized structure of concept maps aligns with Ausubel's idea of organizing knowledge in a meaningful way. Requiring students to reflect on relationships, connections, and hierarchies between concepts enhances meaningful learning (Novak, 2010a).

Second, since concept maps contain graphical elements and sometimes symbols and pictures, Allan Paivio's (1986) *dual coding theory* (DCT) is relevant to understanding concept mapping effectiveness. According to the DCT, information can be stored in two distinct but interconnected coding systems: a verbal system that deals with language and a non-verbal system that deals with information concerning nonverbal objects (i.e., images, symbols) (Clark & Paivio, 1991; Paivio, 1986). Concept maps involve the use of both verbal and visual elements. The verbal elements consist of labels, keywords, or short written text, while the visual elements include shapes, lines, and spatial organization. Thus, the two-modal representation (verbal and visual) of concept mapping aligns with the principles of DCT, which propose that information encoded using two modalities can be more efficiently integrated and retained than relying solely on a single, typically verbal system (Kirschner & Hendrick, 2020). DCT may provide a more precise *mechanism of action* (the potential of dual coding) for explaining how concept maps enhance learning outcomes compared to Ausubel's (1962, 1963) subsumption learning theory. In situations where experimental methods demonstrate the effectiveness of instructional

practices, it becomes essential to understand the underlying mechanisms of action (Bunge, 2013).

Third, *cognitive load theory* (CLT), a theory about instructional design developed by John Sweller and his colleagues (Sweller, 1994, 2020; Sweller et al., 2011), has been used to explain the functioning and usefulness of concept mapping. The starting point of the CLT is that there is limited working memory to code information. The limitations of working memory can be circumvented by organizing schemas (mental frameworks) of knowledge that allow students to work on complex material without exceeding their cognitive capacity (Kirschner & Hendrick, 2020). In other words, because working memory is limited, teachers should design learning experiences to reduce the *extraneous cognitive load* related to the additional cognitive load imposed by instructional design and presentation of information (Chen et al., 2023; Sweller et al., 1998, 2011). *Intrinsic cognitive load* refers to the inherent complexity or difficulty of the material being learned, making it intrinsic to the material being addressed (Sweller et al., 2011). Inherent load is defined by the number of novel elements, the element interactivity that has to be processed, and the expertise of the learners (Chen et al., 2023; Sweller et al., 1998, 2011). Thus, by visually simplifying unnecessary verbal information, the concept maps can reduce the extraneous cognitive load. Furthermore, the graphical and relational structure of concept maps can provide a higher degree of organization when compared to traditional text; this increased organization can assist students in structuring their learning materials into schemas or activating and reorganizing relevant schemas, thus reducing the extraneous cognitive load (Amadieu et al., 2009; Davies, 2011; Rivet & Krajcik, 2007). By providing a visual representation of otherwise verbal information and logically organizing the information, concept maps can be beneficial for low-achieving students and students experiencing learning difficulties (Kirschner, 2002; Sweller, 1994).

In summary, concerning concept maps, both CLT and DCT theories hold significance and offer potential *mechanisms of action* (Bunge, 2013). CLT's emphasis on minimizing extraneous cognitive load is relevant to concept mapping, as concept maps often provide a higher level of organization. DCT's emphasis on integrating verbal and visual elements to improve retention and comprehension aligns with concept mapping, as concept maps inherently involve this integration.

Previous Meta-analyses

Previous meta-analyses, such as those by Horton et al. (1993), Nesbit and Adesope (2006), and Schroeder et al. (2018), were more generic and did not specifically focus on science education at the elementary, intermediate, and secondary school levels. For example, in Horton et al.'s (1993) meta-analysis, six of the 19 studies included in the analysis were conducted at the college/post-secondary level. In Nesbit and Adesope's meta-analysis (2006, pp. 441–446), 38 out of the 55 studies that specified grade levels were conducted at the post-secondary level. The same meta-analysis also focused on a broader range of learning subjects, including statistics, humanities, law, and social studies (Nesbit & Adesope, 2006). Schroeder et al.

([2018](#)) categorized their studies into STEM (science, technology, engineering, and mathematics) or non-STEM fields. In addition, Nicoara et al. ([2017](#)) focused on the medical field.

Science instruction represents a specialized instructional domain distinguished by its unique curriculum (physics, chemistry, biology), content, and perhaps particular instructional challenges. A meta-analysis tailored to the science instruction context is essential to provide insights and guidance relevant to science educators. Furthermore, while broader meta-analyses provide valuable insights, they may not capture the specificities of science education at the elementary, intermediate, and secondary levels. Our search indicated no meta-analysis focusing exclusively on the effectiveness of concept maps in teaching science to elementary and secondary students (Grade 3 to Grade 12). Therefore, this study evaluates the impact of concept mapping on student achievement in science within elementary, intermediate, and secondary education settings. We aim to bridge existing research gaps that pertain to concept mapping effectiveness, particularly the emphasis on science education and the curriculum relevance of concept maps in schools. It is important to note that the complexity of the curriculum for elementary and secondary school students may differ from that of higher education or adult learning contexts.

Effects of Moderator Variables

The effectiveness of concept maps is examined through various moderators, demonstrating the diverse and dynamic nature of educational contexts, learning subjects, and learners. This wide range of moderator variables can shed light on concept maps's impact on learning outcomes and their specific aspects of usage. Among the crucial moderators we explore are the following:

Learning Subject Physics, chemistry, and biology are distinct scientific disciplines, each with its concepts, principles, processes, and facts. Students may encounter different learning challenges in physics, chemistry, and biology due to differences in the nature of experiments, equipment and settings, physical laws, chemical reactions and equations, biological processes, or the complexity of the mathematical concepts involved. Ernst Mayr ([1997](#), p. 63) suggested that concepts may be more critical to biology (e.g., natural selection, adaptation, evolution, and ecology) than physical sciences, where discovering new facts may be more critical. Nevertheless, instruction in physical sciences and biology varies from the process of discovering new knowledge in these fields. Furthermore, analyzing the effectiveness of concept mapping in each learning subject allows for a tailored examination of how an instructional tool aligns with the distinct content of these three instructional domains.

Level of Achievement Characteristics of learners may be relevant to the effectiveness of concept maps (Nesbit & Adesope, [2013](#)). But little is known about the effectiveness of concept mapping for low-achieving or high-achieving students. Few studies have suggested that concept maps could be an effective tool for science

instruction for low-achieving students (Guastello et al., 2000; Morfidi et al., 2017; Roberts & Joiner, 2007).

In our review, most studies included students with mixed abilities or did not define their participants' learning achievement well. Several studies have not stated whether low-achieving students or high-achieving were among their participants. For example, only three studies that fulfilled the selection criteria had their exclusive focus on "low achieving" students (Guastello et al., 2000), "poor readers" (Morfidi et al., 2017), and autism (Roberts & Joiner, 2007) and their control groups had similar low-achieving learning characteristics to the experimental groups. In addition, among the set of seven studies with experimental groups designated by ability, we identified two other studies (Haugwitz et al., 2010; Schmid & Telaro, 1990), which also included control "low achieving" groups as clearly defined subgroups, and had all the information to run a second more focused meta-analysis. Two other studies (Brown, 2000; Udupa, 1992) out of the aforementioned seven studies included experimental groups clearly defined as low achieving, but their control groups were of mixed ability; therefore, these two studies were excluded from the second meta-analysis, which specifically concentrated on exceptional learners, including either low-achieving or high-achieving students. Our literature search identified four studies (Gulati, 2005; Ling & Boo, 2007; Schmid & Telaro, 1990; Snead, 2000), including experimental and control "high-achieving" groups as clearly defined subgroups in their analysis. That is, their control groups had similar high-achieving learning characteristics to the experimental groups. Although one study (Brown, 2000) included an experimental group clearly defined as high achieving, the control group was of mixed ability; for this reason, this study was not included in the second meta-analysis.

Teaching Guidance Concept mapping is an adaptable instructional tool that can be implemented in various ways, from highly structured and directed by the teacher or researcher to minimally guided and constructed by students. We distinguish here three levels of teaching guidance: (a) concept maps constructed by learners through independent work, that is, lowly directed; (b) concept maps given or closely directed by teachers/researchers, that is, highly directed; and (c) concept maps scaffolded by teachers/researchers following a mixed or interactive approach (see Chang et al., 2002). Exploring the impact of different levels of teaching guidance allows us to understand how this teaching aspect of concept mapping can influence learning outcomes.

Grade Level and Level of Schooling The complexity of concept maps may vary depending on the grade level or level of schooling. Cognitive processing also varies with age and level of schooling (Demetriou et al., 2011). For younger students in elementary school, concept maps may emphasize more straightforward concepts and entail fewer relationships. For example, the complexity of the cognitive process is determined by the number of interacting concepts involved (Demetriou et al., 2011; Halford et al., 1998). At higher grade levels, the complexity and interactivity of the concept elements may increase, aligning with the advanced subject matter (Sweller

et al., 2011) and resulting in the construction of more intricate concept maps. Understanding how concept maps impact students at different grades and levels of schooling can help tailor teaching strategies to their cognitive needs.

Type of Technology Concept maps can be either hand-drawn or computer-generated. Hand-drawn concept maps may prompt students to add or develop ideas and require students to put in a significant amount of effort (Erdogan, 2009). On the other hand, computer-generated concept maps can be more easily changed and expanded. Computer-generated concept mapping software usually offers pre-made templates of various maps that can be easily altered and visually improved by adding colorful pictures, symbols, or clip art (Sturm & Rankin-Erickson, 2002). There is evidence from individual studies that both can benefit students by promoting the development of science vocabulary, reading comprehension, and writing (Flanagan & Bouck, 2015; Morfidi et al., 2017; Sturm & Rankin-Erickson, 2002).

Collaboration Level Concept maps can be individually or collaboratively constructed (Kwon & Cifuentes, 2007, 2009). The collaboration level among students (individual construction of concept maps vs. collaborative work during the intervention phase vs. mixed or not defined approach) may impact the effectiveness of concept mapping. For example, when students individually construct concept maps, they can engage deeply with the content and independently organize their understanding. However, some students may struggle to work with concept mapping independently, mainly if they have limited prior experience. Collaborative concept mapping encourages peer interaction, perhaps enabling some students to rectify misconceptions. However, successful collaborative work necessitates effective communication and coordination among students. In certain instances, dominant group members may exert disproportionate influence over the construction or understanding of the concept map.

Economic Level of Countries The economic level of countries, whether classified as low and middle-income or high-income, may influence the effectiveness of concept mapping in several ways. For example, high-income countries have better access to educational resources, including technology, teaching materials, and well-equipped classrooms. In contrast, low and middle-income countries may face resource constraints that can affect the implementation and effectiveness of concept mapping. Conversely, in low-income countries where teacher training resources may be limited compared to high-income countries, concept mapping may be considered a more significant innovation.

Publication Type The publication type (journals vs. dissertations and conference papers) may impact the pooled effectiveness of concept mapping. While journal articles typically undergo a rigorous peer-review process ensuring the quality and reliability of the research findings, the inclusion of dissertations and conference papers, that is, the inclusion of "grey literature," helps mitigate publication bias and the "file drawer problem" (Paez, 2017). Thus, we included grey literature to reflect the

existing evidential base fully, and we assessed the impact of this variation through a moderator analysis.

Considering moderator variables, the research questions for the current meta-analysis are structured as follows, aiming to delve into the multifaceted aspects of concept mapping's effectiveness:

(1) What effect does concept mapping have on elementary and secondary education students' science achievement (physics, chemistry, and biology)?
(2) What are the characteristics and conditions of concept mapping instruction that enhance science achievement?
(3) What effect does concept mapping have on the science achievement of low-achieving and high-achieving students?

Methods

We performed the meta-analysis following the PRISMA guidelines (Forero et al., 2019; Moher et al., 2009). We initially conducted a systematic review of the literature for studies published in English. Studies were selected from the Academic Search Complete, ERIC, PsycINFO, Web of Science, ProQuest Dissertation Database, and conference papers from the American Educational Research Association (AERA) and National Association for Research in Science Teaching (NARST) (see Fig. 1).

Inclusion Criteria

We used inclusion criteria similar to those of Nesbit and Adesope (2006) and Schroeder et al.'s (2018) meta-analyses. The criteria for including studies in this meta-analysis were the following:

The criteria for including studies in this meta-analysis were the following:

(1) Published studies in English;
(2) Articles published in peer-reviewed journals and other sources such as dissertations, theses, and conference papers (e.g., AERA, NAST);
(3) Using concept maps as a method to measure students' achievement in science (biology, physics/earth science, and chemistry);
(4) Participants in Grades 3–12;
(5) Inclusion of a control group;
(6) Reported sufficient data to allow an estimate of standardized mean difference effect size;
(7) Assigned participants to groups before differing treatments;
(8) Randomly assigned participants to groups or used a pretest or other prior variable correlated with the outcome to control for preexisting differences between the groups. Studies were excluded from this meta-analysis when reported a pretest
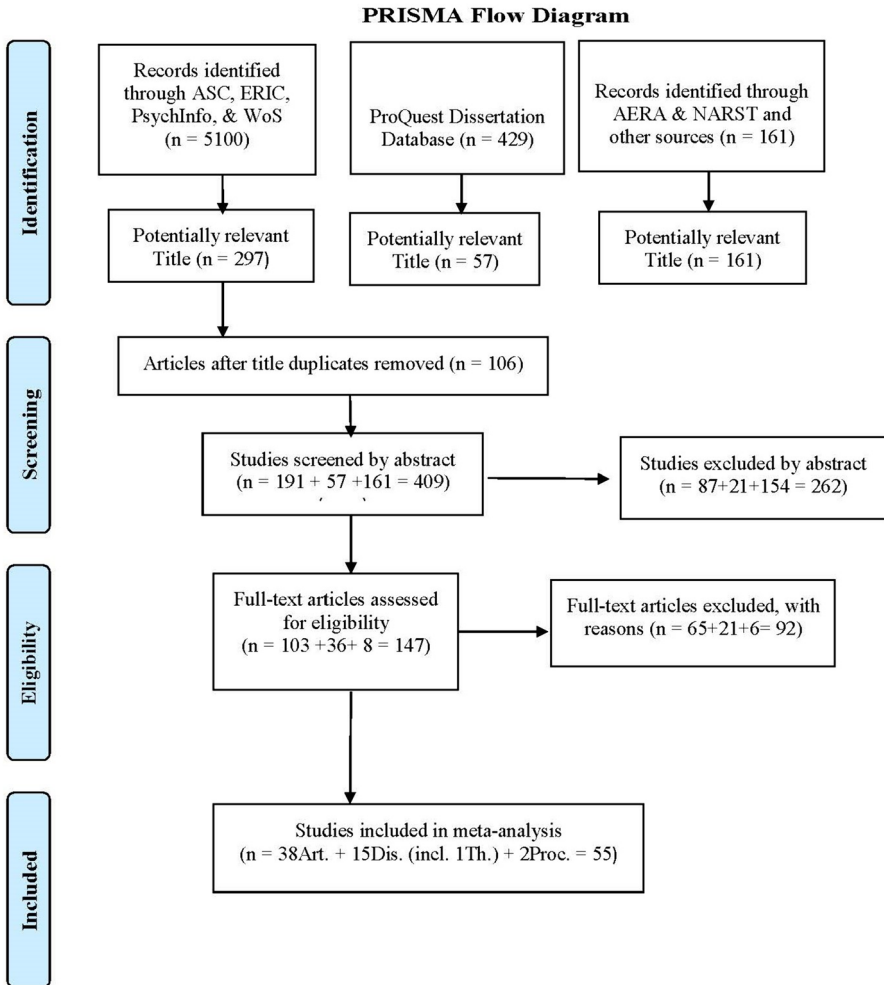
Fig. 1 A Prisma flow diagram of the search methods

effect size outside the range $-0.40 < d < 0.40$ (Nesbit & Adesope., 2006). Furthermore, when the same authors reported more than one source (e.g., dissertation and journal article), we considered the version published in a journal article.

Literature Search

The literature search focused on studies published between 1980 and May 2020. We used the following combination of broad search terms in a Boolean search:

("concept map* OR knowledge map* OR node-link map*) AND ("science") to identify potentially relevant studies.

Coding Procedure

After identifying studies that corresponded to the selection criteria according to our research questions, it was necessary to determine a perspective for a correct coding that concerned the present meta-analysis. Two researchers independently read each of the 55 studies retained in the inclusion phase (see Fig. 1) to (a) select group comparisons consistent with the research questions of the meta-analysis and (b) code each comparison according to a predefined coding form.

The coding form included the (b) *year of study*, (b) *subject area* (biology vs. physics and earth science vs. chemistry), (c) *publication type* (journal vs. dissertation and conference papers), (d) *country of research,* which was recoded according to the World Bank (2020) classification criteria for *country economic level* into *high-income country level vs. middle- and low-income country level* (Serajuddin & Hamadeh, 2020), (e) *grade level* of students and their relevant *level of schooling*, (f) *level of achievement* (low-achieving students vs. groups designated by ability vs. mixed levels of achievement), (g) *duration of intervention* (low duration: < 2 weeks vs. medium duration: 2–4 weeks vs. high duration > 4 weeks; for studies reported in minutes, one week calculated as equivalent to 120 min), (h) *the type of technology* used by students (pencil-and-paper vs. computer-generated maps or digitalized), (i) *degree of teaching guidance* (highly directed: maps given or closely directed by teacher/researcher vs. lowly directed: constructed by students vs. mixed-scaffolded or interactive approach), (j) *collaboration level among students* (individual construction of concept maps vs. collaborative work during the intervention phase vs. mixed or not defined approach) (see Table 1). The first author coded those pertinent features of studies and looked for comparison treatments, including random assignment of participants or control for pre-intervention differences in the experimental and control groups. The second author double-coded the aforementioned features of studies, resulting in interrater reliability of 0.96. Any disagreements were resolved by consensus.

Statistical Methods

We avoided introducing statistically dependent data, as Lipsey and Wilson (2001) have suggested. Thus, we avoided double or multiple data insertion from studies with one control group when experimental groups underwent different intervention approaches. In this case, we selected only one of the two or more experimental groups, the most relevant group to our meta-analysis, to calculate the standardized mean difference effect size. As effect size, we used the standardized mean difference (SMD):

Table 1 Overview of the studies included in the meta-analysis

First author	Year	Subject	Publication Type	Econ. Level of Country	Grade	Type of learners	Interv. Duration	Technology	Teaching Guidance	Collaboration
Abayomi	1988	Physics/Earth Sc	Dissertation	High	8	Mixed	Medium	pencil/paper	Low	Individually
Ade-Ogunnowo	2015	Biology	Conference	Middle/Low	12	Mixed	Low	pencil/paper	High	Individually
Akay	2012	Biology	Article	Middle/Low	11	Mixed	Low	pencil/paper	Low	Collaborative
Akpinar	2008	Biology	Article	Middle/Low	6	Mixed	Medium	Digitalized	Mixed	Collaborative
Al khawaldeh	2009	Biology	Article	Middle/Low	11	Mixed	Medium	pencil/paper	Mixed	Collaborative
Aldaon	2012	Biology	Thesis	High	10	Mixed	Medium	pencil/paper	Mixed	Individually
Asan	2007	Physics/Earth Sc	Article	Middle/Low	5	Mixed	Medium	Digitalized	Low	Collaborative
Balim	2013	Physics/Earth Sc	Article	Middle/Low	6	Mixed	Medium	Digitilized	Low	Collaborative
Berry	2013	Physics/Earth Sc	Article	High	3	Mixed	Medium	pencil/paper	Mixed	Collaborative
Bodolus	1986	Physics/Earth Sc	Dissertation	High	9	Mixed	Medium	pencil/paper	Mixed	Individually
BouJaoude	2003	Chemistry	Conference	Middle/Low	10	Mixed	High	pencil/paper	High	Individually
Bramwell-Lalor	2014	Biology	Article	Middle/Low	12	Mixed	High	pencil/paper	Mixed	Collaborative
Brown	2000	Biology	Dissertation	High	10	Achievement clarified	Medium	pencil/paper	Low	Mixed
Chang	1994	Biology	Dissertation	High	7	Mixed	High	pencil/paper	Low	Individually
Chang	2016	Physics/Earth Sc	Article	High	9	Mixed	High	Digitalized	Mixed	Mixed
Chen	2016	Biology	Article	High	5	Mixed	Low	Digitalized	Mixed	Individually
Conclin	2007	Biology	Dissertation	High	9	Mixed	High	Digitalized	Low	Collaborative
Damaso	2019	Chemistry	Article	Middle/Low	11	Mixed	High	pencil/paper	Low	Mixed
Fatokun	2014	Chemistry	Article	Middle/Low	11	Mixed	High	pencil/paper	Mixed	Individually
Ferreira	2013	Physics/Earth Science	Article	High	7	Mixed	Medium	pencil/paper	Mixed	Mixed
Gongden	2016	Chemistry	Article	Middle/Low	12	Mixed	Low	pencil/paper	High	Individually

Table 1 (continued)

First author	Year	Subject	Publication Type	Econ. Level of Country	Grade	Type of learners	Interv. Duration	Technology	Teaching Guidance	Collaboration
Guastello	2000	Physics/Earth Science	Article	High	7	Low achieving	Medium	pencil/paper	High	Individually
Gulati	2005	Biology	Dissertation	High	9	Achievement clarified	High	pencil/paper	Low	Collaborative
Hagemans a	2012	Physics/Earth Science	Article	High	9	Mixed	Low	Digitalized	Mixed	Individually
Hagemans b	2012	Physics/Earth Science	Article	High	8	Mixed	Low	Digitilized	Mixed	Individually
Haugwitz	2010	Biology	Article	High	8	Achievement clarified	Low	pencil/paper	Low	Collaborative
Hwang	2011	Biology	Article	High	5	Mixed	Medium	Digitalized	Low	Collaborative
Hwang	2020	Biology	Article	High	5	Mixed	Low	Digitalized	Low	Individually
Hwang	2013	Biology	Article	High	6	Mixed	Low	Digitalized	Mixed	Individually
Jegede	1990	Biology	Article	Middle/Low	10	Mixed	High	pencil/paper	Low	Individually
Jena	2019	Biology	Article	Middle/Low	7	Mixed	Medium	pencil/paper	Mixed	Individually
Keown	2008	Physics/Earth Sc	Dissertation	High	5	Mixed	High	Digitalized	Low	Individually
Kim	2008	Physics/Earth Sc	Article	High	5	Mixed	Medium	Digitalized	Mixed	Individually
Kwon	2009	Biology	Article	High	10	Mixed	Low	Digitalized	Low	Individually
Kwon	2007	Physics/Earth Sc	Article	High	6	Mixed	Low	Digitalized	Low	Mixed
Leary	1993	Chemistry	Dissertation	High	10	Mixed	Medium	pencil/paper	Low	Individually
Ledger	2003	Physics/Earth Sc	Dissertation	High	8	Mixed	High	pencil/paper	Low	Collaborative
Ling	2007	Biology	Article	High	4	Achievement clarified (Gifted)	High	pencil/paper	Low	Collaborative
McFall	1999	Biology	Dissertation	High	4	Mixed	High	pencil/paper	Low	Collaborative

Table 1 (continued)

First author	Year	Subject	Publication Type	Econ. Level of Country	Grade	Type of learners	Interv. Duration	Technology	Teaching Guidance	Collaboration
Morfidi	2018	Physics/Earth Sc	Article	High	5	Low achieving	Low	Digitalized	Mixed	Individually
Ogomaya	2016	Biology	Article	Middle/Low	7	Mixed	Medium	pencil/paper	High	Individually
Okebukola	1992	Biology	Article	Middle/Low	11	Mixed	Medium	pencil/paper	Low	Mixed
Olgun	2008	Physics/Earth Sc	Article	Middle/Low	5	Mixed	High	pencil/paper	Mixed	Individually
Pankratius	1990	Physics/Earth Sc	Article	High	12	Mixed	High	pencil/paper	Low	Individually
Redford a	2012	Biology	Article	High	7	Mixed	Low	pencil/paper	Mixed	Individually
Redford b	2012	Biology	Article	High	7	Mixed	Low	pencil/paper	Mixed	Individually
Roberts	2007	Biology	Article	High	7	Low achieving	High	Digitalized	Mixed	Individually
Schmid	1990	Biology	Article	High	11	Achievement clarified	Low	pencil/paper	Low	Individually
Snead	2000	Physics/Earth Sc	Dissertation	High	8	Achievement clarified	High	pencil/paper	Low	Individually
Spaulding a	1989	Chemistry	Dissertation	High	10	Mixed	Medium	pencil/paper	Low	Individually
Spaulding b	1989	Biology	Dissertation	High	10	Mixed	Medium	pencil/paper	Low	Individually
Stanisavljevic	2017	Biology	Article	Middle/Low	6	Mixed	Low	pencil/paper	Mixed	Individually
Stensvold	1989	Chemistry	Dissertation	High	9	Mixed	Low	pencil/paper	Low	Individually
Tastan	2008	Biology	Article	Middle/Low	11	Mixed	High	pencil/paper	Mixed	Individually
Tekkaya	2003	Biology	Article	Middle/Low	9	Mixed	Medium	pencil/paper	Mixed	Mixed
Udupa	1992	Biology	Dissertation	High	10	Achievement clarified (at risk vs. typically achieving)	High	pencil/paper	Low	Collaborative
Uygur	2019	Physics/Earth Sc	Article	Middle/Low	7	Mixed	High	Digitalized	Mixed	Mixed
Willerman	1991	Physics/Earth Sc	Article	High	8	Mixed	Medium	pencil/paper	High	Individually

$$d_i = \frac{X_{1i} - X_{2i}}{S_{pi}}$$

where X_{1i} and X_{2i} are the means of the two groups under comparison in the ith study, and S_{pi} is the pooled standard deviation:

$$S_{pi} = \sqrt{\frac{(n_{1i} - 1)S_{1i}^2 + (n_{2i} - 1)S_{2i}^2}{n_{1i} + n_{2i} - 2}}$$

Usually, the sample estimate of the SMD is termed Cohen's d, but it tends to overestimate the absolute value in small samples (Thompson & Sharp, 1999). This bias can be corrected using the Hedges' g suggested by Hedges and Olkin (1985), which generates an unbiased estimate using a correction factor called J. Therefore, the correction leads to $g_i = Jd_i = d_i - 3d_i/(4n_i - 9)$.

The estimated variance of d is given by:

$$\text{var}(d_i) = s_i^2 = \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}} \right) + \frac{d_i^2}{2(n_{1i} + n_{2i})}$$

and subsequently $\text{var}(g_i) = J^2\text{var}(d_i)$. The interpretation of the results was based on Cohen's suggestion: a magnitude of effect < 0.20 was considered trivial, between 0.20–0.50 small, 0.50–0.80 moderate, and an effect size > 0.80 was considered large (Cohen, 1988).

We used the random-effects model for the meta-analysis as proposed by Der-Simonian and Laird (1986). The method introduces an additive component of the between-studies variance (τ^2), also known as between-study heterogeneity, and we hypothesize that the true effect varies between studies,

$$g_i \sim N\left(g, s_i^2 + \tau^2\right).$$

Then, to obtain a pooled (weighted) estimate of g, we used:

$$\widehat{g} = \frac{\sum_{i=1}^{k} w_i g_i}{\sum_{i=1}^{k} w_i}$$

The weights are given by the inverse-variance estimate used in the meta-analysis with $w_i = \left(\tau^2 + s_i^2\right)^{-1}$ (Normand, 1999; Petiti, 1994). Inferences about the overall effect are based on the normal approximation since:

$$\text{var}\left(\widehat{g}\right) = \frac{1}{\sum_{i=1}^{k} w_i}$$

Heterogeneity is not easily evaluated using τ^2. Thus, we used the standard Cochran's Q and the inconsistency index (I^2), which ranges between 0 and 100% and provides easy quantification of the variability between studies. The potential sources of heterogeneity were investigated using subgroup analysis or

meta-regression analysis using study-level characteristics (Thompson & Sharp, 1999).

Influential studies were identified by removing each time a study from the meta-analysis and observing the change in the pooled estimate. Publication bias was estimated using the rank correlation method (Begg & Mazumdar, 1994), the regression method of Egger's, and its random effects analog (Egger et al., 1997). The non-parametric "trim and fill" method of Duval and Tweedie (2000) was also used to estimate the number of "missing studies." The potential time trend of the combined effect over the years was evaluated by applying the standard cumulative meta-analysis approach (Lau et al., 1995) using two different formal tests dealing with time trend: the "first vs. subsequent" comparison (Ioannidis & Trikalinos, 2005) and the Generalized Least Squares (GLS) Regression-based test (Bagos & Nikolopoulos, 2009). The statistical package used in the analysis was Stata 13 (Stata Corporation, 2013), while statistical significance was set at $p \leq 0.05$.

Results

On May 12, 2020, we searched for studies published from 1980 until the date the search was conducted in the following databases: Academic Search Complete (ASC; 1245), ERIC (835), PsycINFO (970), and Web of Science (WoS; 2050); the number of studies returned is indicated in parentheses. In the first search (identification phase), we found 5100 studies, many of which addressed topics other than the effectiveness of concept mapping in science. Thus, after reviewing the records, 4803 studies were eliminated as irrelevant. After removing duplicates (n = 113) across databases, 191 articles were maintained and further screened by their abstracts. On May 14, 2020, we searched for dissertation studies in the ProQuest Dissertations and Theses Database published from 1980 until the search date, and 429 dissertations were revealed. After reviewing the records, 372 studies were eliminated as irrelevant, and 57 dissertations were further screened by their abstract. On May 26, 2020, we searched (a) the titles of papers presented at the AERA conferences between 2010 and 2019 in their online paper repository and (b) the titles and abstracts of papers presented at the NARST program conferences between 2007 and 2019. We also conducted a Google Scholar search to find additional papers and gray literature using the same terms. Titles and abstracts were screened for eligibility and discarded if they did not meet the inclusion criteria. This search yielded 123 and 38 papers of potential relevance from online paper repositories and Google Scholar, respectively; thus, 161 papers were further screened by their abstracts. The search results from all data sources are presented in Fig. 1, employing the PRISMA template (Moher et al., 2009).

Of the 409 studies screened by abstract, 262 were irrelevant to concept mapping effectiveness in science (87 from databases, 21 from the ProQuest Dissertations and Theses database, and 154 from conferences, etc.). After meticulously reviewing the remaining 147 studies for their content and applying our inclusion criteria, 55 eligible articles and papers were included in the present meta-analysis. Among those 55 studies, 38 were journal articles, 15 were dissertations (including one Master's

thesis), and two were proceeding papers. Figure 1 illustrates the stages of the selection process of the studies included in the final analysis in a PRISMA Flow Diagram.

The 55 studies yielded 58 comparisons with non-dependent data (Lipsey & Wilson, 2001). Three of the 55 original studies (Hagemans et al., 2013; Redford et al., 2012; Spaulding, 1989) were larger research projects. Each of these three studies comprised two distinct studies with their own separate control groups. Based on the 58 comparisons in the meta-analysis, the total number of participants was 5364. Of them, 2,903 participants were part of the experimental groups, and 2,461 were part of the control groups. The overall ES based on 58 comparisons was estimated at 0.776 (95% CI: 0.597, 0.956; $p < 0.0001$), suggesting a moderate to large effect of the intervention (see Fig. 2).

The between-studies heterogeneity is high, with an I^2 equal to 89.1% (see Higgins et al., 2003). The methods for detecting publication bias suggest a substantial chance for this or other small-study-related bias ($p = 0.035$ for Begg's test, $p = 0.007$ for Egger's test, and $p < 0.0001$ for the random effects regression test; Fig. 3).

The trim and fill method detected an asymmetry to the funnel plot and estimated an additional 16 studies that, if appended to the dataset, would "correct" it. This addition would produce an attenuated estimate but still highly significant ($g = 0.403$, 95% CI: 0.198, 0.608). The influential analysis identified the study of Guastello et al. (2000) as the most influential one (i.e., the one with the largest
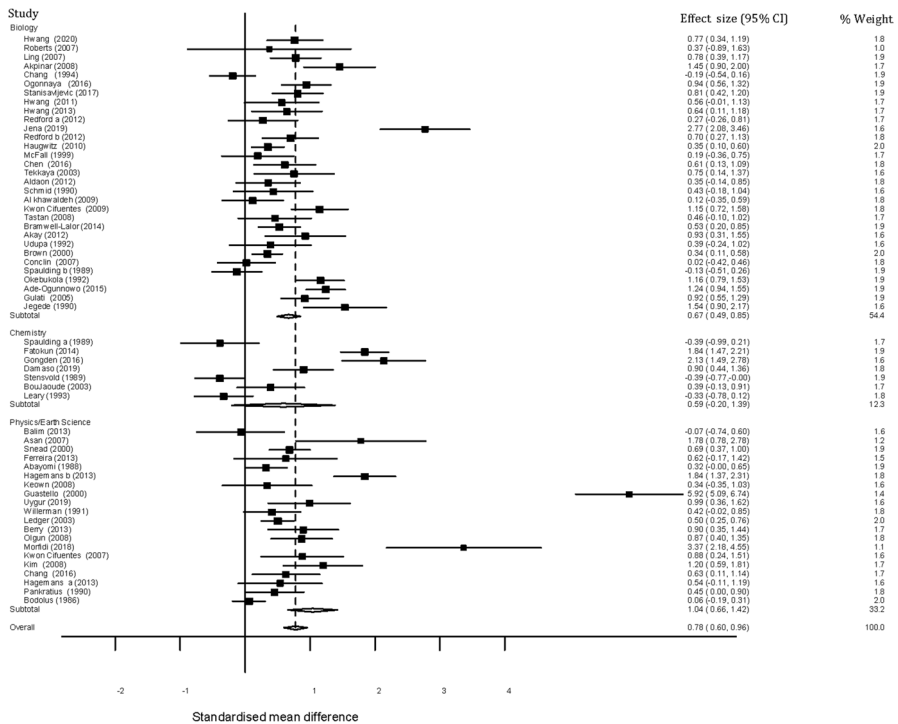

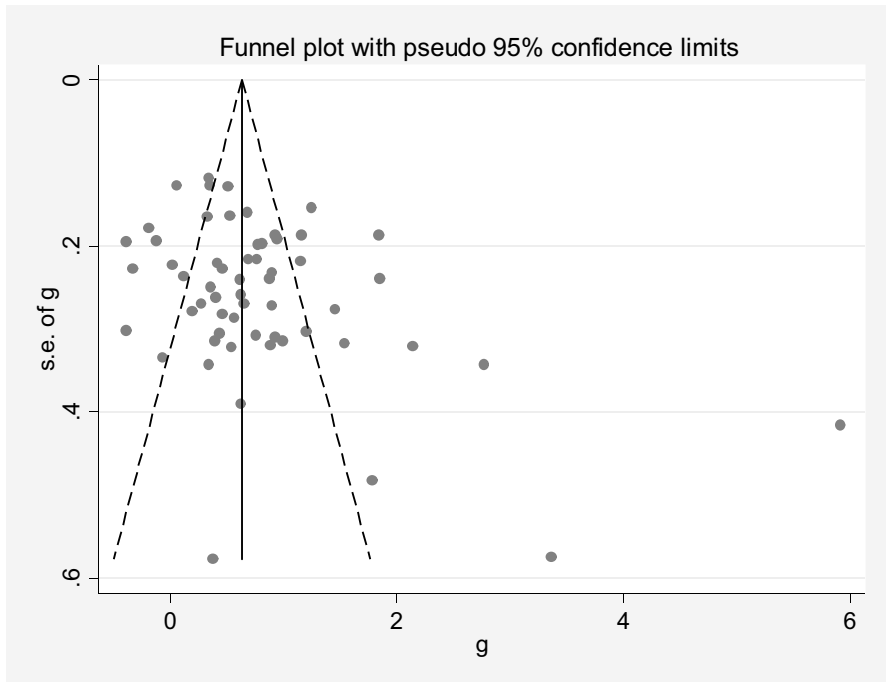**Fig. 2** Results from meta-analysis by subject and weighted mean effect sizes

Fig. 3 Funnel plot for detecting publication bias

effect sizes), followed by the studies of Morfidi et al. (2017) and Jena (2019). Interestingly, when the studies of Guastello et al. and Morfidi et al. were excluded from the analysis, the pooled estimate was reduced to 0.668 (95% CI: 0.519, 0.817), and the overall heterogeneity was reduced to 84%, and the evidence of publication bias was eliminated ($p = 0.182$ for Begg's test, $p = 0.073$ for Egger's test and $p = 0.082$ for the random effects regression test). The standard cumulative meta-analysis provides evidence for increased effect size over time (Fig. 4).

Mean effect sizes varied from moderate to large based on the subject area. For biology ($k = 31$ comparisons (from 30 studies)), the mean effect size was moderate ($g = 0.671$; 95% CI: 0.492, 0.851; $p < 0.0001$), but the between studies heterogeneity was high, with an $I^2 = 80.8\%$. For chemistry ($k = 7$ studies, the mean effect size was small ($g = 0.590$; 95% CI: -0.204,1.385; $p < 0.145$), and a high between studies heterogeneity with $I^2 = 95.0\%$. For physics ($k = 20$, comparisons from 19 studies), the mean effect size was large ($g = 1.040$; 95% CI: 0.656, 1.423; $p < 0.0001$), but the between studies heterogeneity was also high, with an $I^2 = 92.2\%$. In subgroup analyses, the effect sizes obtained from *biology*, *physics/earth science*, and *chemistry* did not yield significant differences among them ($p = 0.220$).

The *level of achievement* (low achieving students vs. groups designated by ability vs. mixed levels of achievements) also yielded significant differences (see Table 2), but this was related to the influential studies of Guastello et al. (2000) and Morfidi
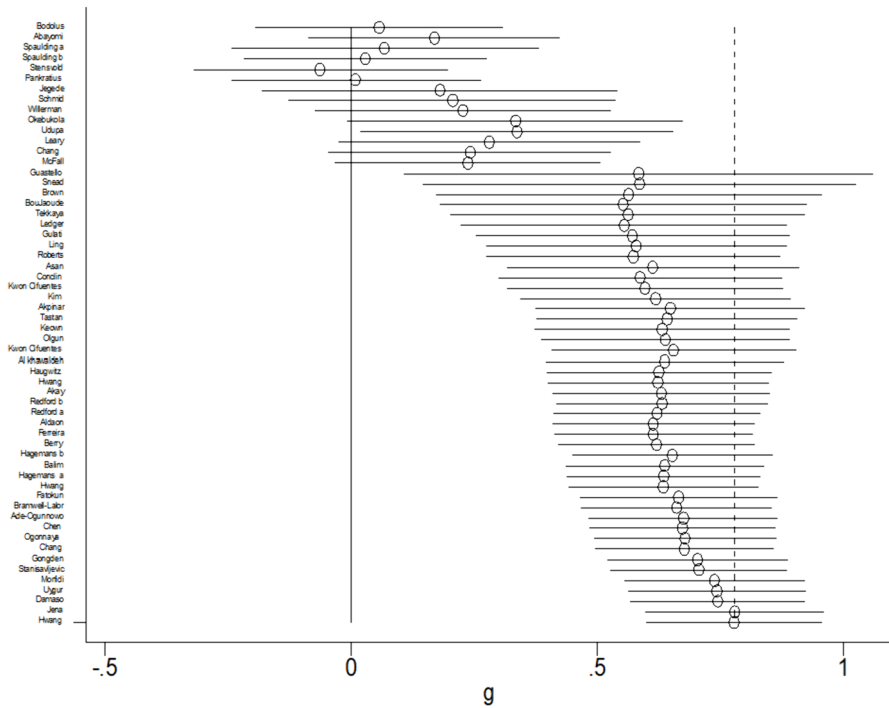
Fig. 4 The standard cumulative meta-analysis shows increased effect size over time

et al. (2017), which were performed on low-achieving learners; should these two studies be excluded, the effect disappears.

The degree of *teaching guidance* (high guidance vs. low guidance vs. mixed approach) seems to be a significant factor with a higher effect size for high teaching guidance (see Table 2). Still, after removing the two influential studies, this result was marginal ($p = 0.074$).

Regarding the *grade level*, we did not find significant differences among grades, but it is possible that the lack of significance in this result could be attributed to the limited number of studies within each grade. To explore this further, we made two distinct categorizations of the *schooling level*. First, similar to Horton et al. (1993), we assessed the effectiveness of concept maps at three schooling levels: elementary (Grades 3–6), middle (Grades 7–9), and high school (Grades 9–12). There were 16 studies containing 1,015 students from Grades 3–6 (elementary), 22 studies with 2,501 students from Grades 7–9 (middle level), and 20 studies with 2,059 students from Grades 10–12 (high level). In the subgroup analysis, the three categories did not yield significant differences ($g = 0.836, 0.821, 0.673$, respectively, $p = 0.76$).

Second, like Nesbit and Adesope (2006) and Schroeder et al. (2018), we made an alternative categorization of the *schooling level*, distinguishing between the intermediate level (Grades 3–8; encompassing 31 studies and 2783 students) and the secondary level of schooling (Grades 9–12; encompassing 27 studies and 2792 students). The results showed a significant difference ($g = 0.959$ vs. $0.587, p = 0.048$),

Table 2 Significant moderators

	Number of comparisons	Pooled SMD (g)	95% C.I	I^2	$Q(p)$
All Comparisons	58	0.776	0.597 0.956	89.1%	522.16 (<0.0001)
Comparisons excluding two influential studies	56	0.668	0.519 0.817	84.0%	343.56 (<0.0001)
Type of Learners-Level of Achievement					
Mixed levels or not defined	48	0.691	0.597 0.956	85.8%	329.93 (<0.0001)
With experimental groups designated by the ability	7	0.551	0.365 0.737	47.7%	11.48 (0.075)
Low-achieving	3	3.244	0.014 6.474	96.3%	53.70 (<0.0001)
Teaching Guidance					
High (given by teacher/researcher)	6	1.792	0.735 2.849	96.8%	156.34 (<0.0001)
Mixed-scaffolded or interactive	24	0.905	0.633 1.177	85.8%	161.47 (<0.0001)
Low (independent construction)	28	0.470	0.293 0.647	79.4%	131.12 (<0.0001)
Schooling Level					
Intermediate level	31	0.959	0.688 1.230	90.1%	301.60 (<0.0001)
Secondary level	27	0.587	0.351 0.823	87.8%	212.81 (<0.0001)
Country of Research by Economic Level					
Low- and middle-income countries	20	1.053	0.791 1.315	89.0%	110.81 (<0.0001)
High-income countries	38	0.626	0.412 0.841	82.9%	336.11 (<0.0001)
Publication Type					
Journals	40	1.039	0.808 1.270	88.0%	325.48 (<0.0001)
Dissertations and Conference papers	18	0.254	0.040 0.467	82.8%	98.95 (<0.0001)

suggesting that the effect is larger in intermediate grades (see Table 2). However, this significant difference may be attributable to the larger proportion of studies investigating chemistry in the secondary grades (7 compared to 0 studies) and the lower proportion of the studies investigating physics (4 compared to 16 studies). We note that in the overall analysis, chemistry had a smaller effect size compared to physics and biology (0.59 vs. 1.04 and 0.671), even though this learning subject difference was not significant.

Regarding the *economic level of countries*, the effect size from studies conducted in high-income countries is significantly smaller than those from low- and middle-income countries ($g=0.626$ vs. 1.053); this difference in effect sizes was not affected by the two aforementioned influential studies (see Table 2). Furthermore, journal publications had larger effect sizes than dissertations and conference papers, and this difference was statistically significant ($g=1.039$ vs. 0.254, see Table 2). The *year of publication* is also related to increased effect size over time, but the proportion of journal publications also increased over time (i.e., in Web of Science, Eric, and PsycINFO databases; see Nesbit & Adesope, 2013). Finally, various features of studies such as *technology* (pencil-and-paper vs. digitalized), *learning environment* (individual vs. collaborative vs. mixed or not clearly defined), and *duration of intervention* did not have significant differences (data not shown in Table 2).

Building a model with multiple variables, we ended up with publication type (journals vs. dissertations and conference papers), country of research by economic level (high income vs. middle- and low-income), and type of learners, resulting in a 46.63% reduction between studies variability. Regarding research question 3, we conducted a separate meta-analysis, including studies focusing on exceptional learners (low-achieving and high-achieving students). The mean effect size was large for low-achieving students ($g=2.019$; $k=5$) and moderate for high-achieving students ($g=0.610$; $k=4$), even though the between-groups differences did not reach statistical significance ($p=0.261$) due to a small number of studies in this area (Fig. 5).

Discussion

This meta-analysis examined the effectiveness of concept maps in improving students' science achievement in elementary and secondary education. Specifically, it focused on science school subjects and included 55 studies with 58 comparisons with experimental and quasi-experimental designs. Overall, the results of the present meta-analysis suggest that concept maps are an effective teaching tool to increase students' science achievement ($g=0.776$), even after two influential studies were excluded from the analysis ($g=0.668$). In a small subset of their meta-analysis, Nesbit and Adesope (2006) found a similar effect size ($g=0.742$, $k=10$) when learners constructed concept maps rather than attended a lecture or discussion. Nonetheless, their findings revealed more modest weighted mean effect sizes for concept map construction under two specific learning conditions: (a) in physical science ($g=0.283$, $k=6$), and (b) in the domains of general science, biology, and statistics ($g=0.522$, $k=9$). Our findings are also consistent with Schroeder et al. (2018) meta-analysis that indicated that concept maps were superior to other instructional
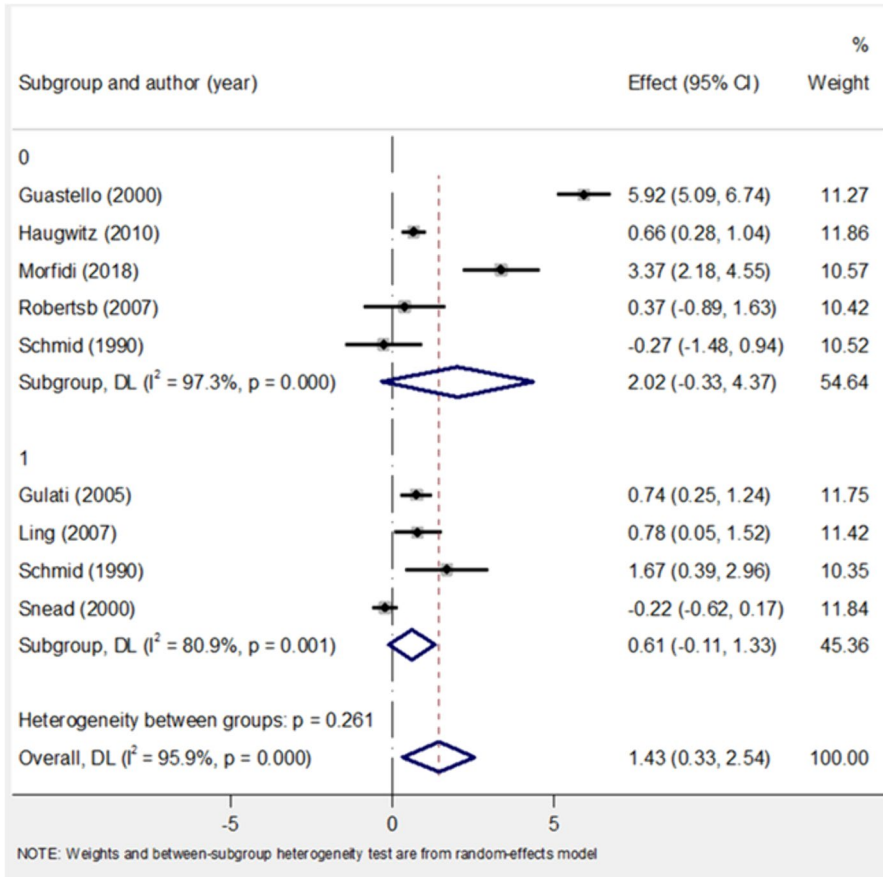
Fig. 5 Results from the second meta-analysis focusing exclusively on low-achieving and high-achieving learners

comparison conditions and effective across science, technology, engineering, and math (STEM) ($g = 0.60$, $k = 118$), although their meta-analysis was more generic and not specifically focused on science for students in the elementary and secondary education.

In short, evidence from the present meta-analysis suggests that concept maps, compared to traditional instructional methods such as attending lectures, reading text passages, and participating in class discussions, are more effective for facilitating science education. This effectiveness can be explained from three theoretical perspectives. First, concept mapping promotes meaningful engagement and meaningful learning (Novak & Cañas, 2008; Novak & Gowin, 1984; Novak, 2010a). Second, due to their incorporation of both verbal and visual components, concept maps align with the principles of dual coding theory, leveraging the memory system's capacity to store and organize information in a more interconnected and resilient manner (Kirschner & Hendrick, 2020). Third, concept

maps serve as a learning tool, simplifying complex verbal information, omitting unnecessary verbal clutter, and presenting key concepts concisely, thus *reducing extraneous cognitive load*. Furthermore, concept maps' graphical and organizational or relational format offers an additional layer of cognitive support. These visual representations provide a certain degree of organization, enabling students to structure their learning materials into coherent schemas. The organizational structure of concept maps can ultimately contribute to reducing the extraneous cognitive load (Amadieu et al., 2009; Kirschner & Hendrick, 2020; Schroeder et al., 2018). For example, biology instruction places a significant emphasis on understanding biological concepts and processes, such as photosynthesis, cellular respiration, homeostasis, and metabolism, and understanding relationships between different elements, such as species interactions within ecological webs. Concept maps can illustrate hierarchical, sequential, and cause-and-effect relationships. In short, concept maps, through their visual clarity and organizational capability, offer a *mechanism of action* to reduce *extraneous cognitive load* and enhance learning outcomes in science education.

Learning Subject Concept mapping seems to be an effective teaching strategy to enhance the science learning of students across science disciplines: biology ($g = 0.671$, $k = 31$), physics ($g = 1.040$, $k = 20$), and chemistry ($g = 0.590$, $k = 7$). Further research is needed for chemistry because of the small number of existing studies ($k = 7$) and very high heterogeneity.

Evidence-based practices (EBPs) fulfill sound criteria and rigorous indicators (Cook & Cook, 2013). While concept mapping has shown strong evidence-based potential in biology and physics, the case for its effectiveness in chemistry is still promising since the pooled effect size from these studies seemed smaller.

There may be some intrinsic reasons why concept maps typically do a better job in physics than chemistry. Instruction in physics typically places a strong emphasis on theoretical concepts and their application to real-world phenomena. Concept maps can help students bridge the gap between abstract theories and physical phenomena. In contrast, chemistry instruction may have a more experimental focus, where students engage in laboratory work, potentially reducing the reliance on concept maps to bridge theory and practice. Furthermore, in chemistry instruction, other visual aids, such as chemical diagrams, charts, and models, may assume a more prominent role when compared to physics and biology.

Type of Learners Our first meta-analysis focusing on the level of achievement (low-achieving students vs. groups designated by ability vs. mixed levels of achievements) yielded a statistically significant difference ($g = 3.244$ vs. 0.551 vs. 0.691), suggesting that the effect is larger in low-achieving students. However, this finding was predominantly influenced by the influential studies conducted by Guastello et al. (2000) and Morfidi et al. (2017), which focused on low-achieving learners; should these two studies be excluded, the effect disappears.

Our third research question focused on the effects that concept mapping has on the science achievement of low-achieving and high-achieving students. For this

purpose, we conducted a secondary meta-analysis, specifically incorporating studies that exclusively centered on exceptional learners: low-achieving and high-achieving students. The mean effect size was large for low-achieving students ($g=2.019$; $k=5$) and moderate for high-achieving students ($g=0.610$; $k=4$). However, the difference between these two groups did not attain statistical significance, owing to the limited number of studies in the comparison. Overall, our findings are encouraging regarding the effectiveness of concept mapping for low-achieving students.

It has also been hypothesized that by using concept mapping rather than plain text, low-achieving students can better understand the graphic representations of relationships among concepts (Miller, 2016). In addition, the typical hierarchical organization of ideas in concept mapping could improve information organization by low-achieving students (Boyle, 1996; Morfidi et al., 2017).

Teaching Guidance Our results indicate a statistically significant higher effect size for concept mapping interventions with high levels of teaching guidance ($g=1.792$) when compared to those with low teaching guidance ($g=0.470$) or scaffolded/interactive approaches ($g=0.905$). Nevertheless, it is worth noting that after excluding the two influential studies that specifically targeted low-achieving students, this finding became marginally significant ($p=0.074$). In practical terms, this marginal finding suggests that students tend to benefit more when teachers exert high levels of teaching guidance in concept mapping. Conversely, when students are expected to construct concept maps independently, the impact on their learning outcomes tends to be lower. However, this difference may be influenced by the type of learners involved in concept mapping instruction. Low-achieving students may derive greater benefits from a kind of explicit instruction in mapping concepts, which includes elements such as clear goals, logical sequencing, guided practice, immediate corrective feedback, comprehension checks, and mastery learning (Hughes et al., 2017; Kirschner & Hendrick, 2020; Kirschner et al., 2022; Pullen & Hallahan, 2015). Additionally, concept maps that are less directed by teachers may result in a lower degree of structure, which does not significantly facilitate learners' ability to organize the material effectively, thus potentially failing to reduce adequately the cognitive load (Amadieu et al., 2009).

Level of Schooling Our results indicated that the impact of concept mapping is more pronounced at the intermediate level of schooling (Grades 3–8) than at the secondary level. However, this statistically significant difference may be influenced by the learning subject, that is, the higher proportion of studies focusing on chemistry within the secondary grades group and the relatively lower proportion of studies addressing physics.

Consistent with the above finding, Horton et al. (1993) found moderate to high mean effect sizes in elementary school and college and small effect sizes in middle school and high school. Also, Nesbit and Adesope (2006) found high effect sizes at the intermediate school level (Grades 4 to 8) and a negligible effect size at high school (Grades 9–12). Further research is needed to analyze how concept maps affect information processing and cognitive load at different levels of schooling.

Economic Level Our results demonstrate a significantly higher effect size for concept mapping interventions when carried out in low- or middle-income countries compared to those conducted in high-income countries. Low- and middle-income countries often have limited educational resources and may face more significant challenges in delivering advanced instructional methods. In such contexts, concept mapping may be seen as an innovative approach that stands out as a more effective tool for enhancing learning outcomes.

Publication Type We found a significantly higher effect size for concept mapping studies published in journals than those published in grey literature (journals vs. dissertations and conference papers). In general, empirical evidence has shown consistent differences between studies published in journals and those in the grey literature; studies published in journals are more likely to contain statistically significant results and larger effect size estimates compared to unpublished studies (McAuley et al., 2000; Paez, 2017). Like in other areas of interventions, our meta-analysis shows that excluding grey literature may exhibit an overrepresentation of studies with statistically significant findings, primarily because studies with significant outcomes are more prone to publication than those with null or negative results (Paez, 2017). This can result in inflated and less precise effect size estimates for concept mapping interventions compared to a meta-analysis incorporating grey literature.

In addition, we found statistically significant differences regarding the *late year of publication*. Finally, technology (pencil-and-paper vs. computer-generated maps or digitalized), learning environment (individual vs. collaborative vs. mixed or not clearly defined), and duration (low vs. medium vs. high duration) did not yield statistically significant differences.

Implications for Research and Practice

The present meta-analysis indicates that concept mapping holds significant promise as an evidence-based instructional approach in biology and physics school instruction. However, while promising, the evidence supporting its effectiveness in chemistry instruction appears to be less robust, as indicated by the combined effect sizes from these studies, which tend to be smaller. Additional research is needed to establish its effectiveness in chemistry instruction due to the limited number of existing studies and the significant heterogeneity observed.

Moreover, further research is needed to explore the instructional aspects and learning conditions that influence the effectiveness of concept maps as a teaching tool in science school classrooms. Only five studies that referred to low-achieving students were included in the meta-analysis without providing information on whether students had learning disabilities (LD). Therefore, well-designed and rigorous empirical studies are required to examine whether concept mapping tools are effective for students across different performance levels. Specifically, research on the effectiveness of concept maps is needed to advance our knowledge about evidence-based practices for low-achieving students and students with LD. Similarly,

the perception that concept maps are a student-friendly instructional method for students with LD needs further investigation.

Noteworthy, Oliver (2009) pointed out the shortcomings of concept maps as a singular learning strategy without using other techniques. Concept maps may not necessarily operate similarly to improve the learning outcome across different achievement levels. However, they may be effective if tailored to students' cognitive needs and learning abilities. Teachers should consider aligning concept maps with instruction and other evidence-based practices.

Limitations

Although the current meta-analysis followed all available guidelines and best practices, it cannot avoid limitations. High heterogeneity is a significant concern of the present meta-analysis, denoting that the studies in the meta-analysis produced estimates that differ significantly from one another. Publication bias was also detected in this meta-analysis. The funnel plot showed evidence of publication bias, as there were influential studies with low-achieving students that reported highly positive effects of concept mapping instruction. When the studies of Guastello et al. and Morfidi et al. were excluded from the analysis, although the pooled estimate was reduced from 0.776 to 0.668, the evidence of publication bias was eliminated.

References

References marked with an asterisk indicate studies included in the meta-analysis.

* Abayomi, B. I. (1988). *The effects of concept mapping and cognitive style on science achievement* (Publication No. 8814330) [Doctoral dissertation, Georgia State University Atlanta]. ProQuest Dissertations and Theses.
* Ade-Ogunnowo, F. E., Ogundele, O. E., & Kalejaiye, B. A. (2015). *Mind mapping and concept mapping as effective learning tools in biology education*. Proceedings of the 1st Interdisciplinary Conference of UCC-TASUED 2015, University of Cape Coast Ghana.
* Adlaon, R. B. (2012). *Assessing effectiveness of concept map as instructional tool in high school biology* [Unpublished master's thesis]. Louisiana State University.

* Akay, S. Ö., Baştürk, K. A. Y. A., & Kilic, S. (2012). The effects of concept maps on the academic success and attitudes of 11th graders while teaching urinary system. *International Online Journal of Primary Education*, *1*(1), 23-30.

* Akpinar, E., & Ergin, O. (2008). Fostering primary school students' understanding of cells and other related concepts with interactive computer animation instruction accompanied by teacher and student-prepared concept maps. *Asia-Pacific Forum on Science Learning and Teaching, 9*(1), 1–13.

* Al khawaldeh, S. A., & Al Olaimat, A. M. (2009). The contribution of conceptual change texts accompanied by concept mapping to eleventh-grade students understanding of cellular respiration concepts. *Journal of Science Education and Technology*, *19*(2), 115–125. https://doi.org/10.1007/s10956-009-9185-z

Amadieu, F., Van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction, 19*(5), 376–386. https://doi.org/10.1016/j.learninstruc.2009.02.005

* Asan, A. (2007). Concept mapping in science class: A case study of fifth grade students. *Educational Technology & Society*, *10*(1), 186–195.

Ausubel, D. (1962). A subsumption theory of meaningful verbal learning and retention. *The Journal of General Psychology, 66*, 213–222. https://doi.org/10.1080/00221309.1962.9711837

Ausubel, D. (1963). *The psychology of meaningful verbal learning*. Grune & Stratton.

Bagos, P. G., & Nikolopoulos, G. K. (2009). Generalized least squares for assessing trends in cumulative meta-analysis with applications in genetic epidemiology. *Journal of Clinical Epidemiology, 62*(10), 1037–1044. https://doi.org/10.1016/j.jclinepi.2008.12.008

* Balım, A. G. (2013). Use of technology-assisted techniques of mind mapping and concept mapping in science education: A constructivist study. *Irish Educational Studies*, *32*(4), 437–456. https://doi.org/10.1080/03323315.2013.862907

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*(4), 1088–1101. https://doi.org/10.2307/2533446

* Berry, J., Potter, J., & Hollas, V. (2013). Concept Maps and Informational Read-Alouds: Strengthening Both Science and Reading for Elementary Students. *Texas Journal of Literacy Education*, *1*(1), 29–44.

* Bodolus, J. E. (1986). *The use of a concept mapping strategy to facilitate meaningful learning for ninth grade students in science* [Unpublished doctoral dissertation]. Temple University.

* BouJaoude, S., & Attieh, M (2003 March 23–26). *The effect of using concept maps as study tools on achievement in chemistry* [Paper presentation]. Annual Meeting of the National Association for Research in Science Teaching, Philadelphia, PA, United States.

Boyle, J. R. (1996). The effects of a cognitive mapping strategy on the literal and inferential comprehension of students with mild disabilities. *Learning Disability Quarterly, 19*(2), 86–98. https://doi.org/10.2307/1511250

* Bramwell-Lalor, S., & Rainford, M. (2014). The effects of using concept mapping for improving advanced level biology students' lower- and higher-order cognitive skills. *International Journal of Science Education*, *36*(5), 839-864. https://doi.org/10.1080/09500693.2013.829255

* Brandt, L., Elen, J., Hellemans, J., Heerman, L., Couwenberg, I., Volckaert, L., & Morisse, H. (2001). The impact of concept mapping and visualization on the learning of secondary school chemistry students. *International Journal of Science Education*, *23*(12), 1303-1313. https://doi.org/10.1080/09500690110049088

* Brown, D. S. (2000). *The effect of individual and group concept mapping on students' conceptual understanding of photosynthesis and cellular respiration in three different academic levels of biology classes* (Publication No. 9970734) [Doctoral dissertation, University of Missouri-Kansas City]. ProQuest Dissertation and Theses.

Bunge, M. (2013). *Medical philosophy: Conceptual issues in medicine*. World Scientific.

Carnine, L., & Carnine, D. (2004). The interaction of reading skills and science content knowledge when teaching struggling secondary students. *Reading & Writing Quarterly, 20*(2), 203–218. https://doi.org/10.1080/10573560490264134

Chang, K.-E., Sung, Y.-T., & Chen, I.-D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education, 71*(1), 5–23. https://doi.org/10.1080/00220970209602054

* Chang, W. (1994). *The effects of using concept mapping to supplement class notes on the biology test scores of seventh grade students in Taiwan, R.O.C.* (Publication No. 9513557) [Doctoral dissertation, University of Iowa]. ProQuest Dissertation and Theses.

* Chang, C., Yeh, T., & Shih, C. (2016). The effects of integrating computer-based concept mapping for physics learning in junior high school. *Eurasia Journal of Mathematics, Science and Technology Education*, *12*(9). https://doi.org/10.12973/eurasia.2016.1284a

* Chen, C., Chou, Y., & Huang, C. (2016). An augmented-reality-Based concept map to support mobile learning for science. *The Asia-Pacific Education Researcher*, *25*(4), 567–578.https://doi.org/10.1007/s40299-016-0284-3

Chen, O., Paas, F., & Sweller, J. A. (2023). Cognitive load theory approach to defining and measuring task complexity through element interactivity. *Educational Psychology Review, 35*(63), 1–18. https://doi.org/10.1007/s10648-023-09782-w

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum

* Conklin, E. (2007). *Concept mapping: Impact on content and organization of technical writing in science* (Publication No. 3254433) [Doctoral dissertation, Walden University]. ProQuest Dissertation and Theses.

Cook, B. G., & Cook, S. C. (2013). Unraveling evidence-based practices in special education. *Journal of Special Education, 47*, 71–82. https://doi.org/10.1177/0022466911420877

* Damaso, K., & Banda, A. (2019). Effect of concept mapping on learner achievement on mole concept. *International Journal of Advanced Research and Publications*, *3*(11), 18–26.

Davies, M. (2011). Concept mapping, mind mapping and argument mapping: What are the differences and do they matter? *Higher Education, 62*, 279–301. https://doi.org/10.1007/s10734-010-9387-6

Demetriou, A., Spanoudis, G., & Mouyi, A. (2011). Educating the developing mind: Towards an overarching paradigm. *Educational Psychology Review, 23*(4), 601–663. https://doi.org/10.4324/9781003187486

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89–98. https://doi.org/10.2307/2669529

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629

Erdogan, Y. (2009). Paper-based and computer-based concept mappings: The effects on computer achievement, computer anxiety and computer attitude. *British Journal of Educational Technology, 40*(5), 821–836. https://doi.org/10.1111/j.1467-8535.2008.00856.x

* Fatokun, K. V. F., & Eniayeju, P. A. (2014). The effect of concept mapping-guided discovery integrated teaching approach on Chemistry students achievement and retention. *Educational Research and Reviews*, *9*(22), 1218-1223.https://doi.org/10.5897/ERR2014.1848

* Ferreira, P., & Zygouris-Coe, V. (2013). Concept mapping and middle school science students' vocabulary and comprehension: A quasi-experimental study. *Journal of Reading Education*, *39*(1), 23–31.

Flanagan, S. M., & Bouck, E. C. (2015). Mapping out the details: Supporting struggling writers' written expression with concept mapping. *Preventing School Failure, 59*(4), 244–252. https://doi.org/10.1080/1045988X.2014.933400

Forero, D. A., Lopez-Leon, S., González-Giraldo, Y., & Bagos, P. G. (2019). Ten simple rules for carrying out and writing meta-analyses. *PLOS Computational Biology, 15*(5), 1–7. https://doi.org/10.1371/journal.pcbi.1006922

* Gongden, E. J., & Delmang, T. K. (2016). The efficacy of concept mapping instructional strategy in remedying students' problem-solving difficulties in stoichiometry. *African Research Review*, *10*(2), 144.https://doi.org/10.4314/afrrev.v10i2.9

* Greene, P. (2011). *Concept mapping and the science achievement of third grade students* (Publication No. 3450514) [Doctoral dissertation, Walden University]. ProQuest Dissertation and Theses.

* Guastello, E. F., Beasley, T. M., & Sinatra, R. C. (2000). Concept mapping effects on science content comprehension of low-achieving inner-city seventh graders. *Remedial and Special Education*, *21*(6), 356–364.https://doi.org/10.1177/074193250002100605

* Gulati, S. (2005). *a comparison of inquiry-based teaching through concept maps and traditional teaching in biology* (Publication No. 3188183) [Doctoral dissertation, University of South Dakota]. ProQuest Dissertation and Theses.

* Hagemans, M. G., Van der Meij, H., & De Jong, T. (2013). The effects of a concept map-based support tool on simulation-based inquiry learning. *Journal of Educational Psychology*, *105*(1), 1-24. https://doi.org/10.1037/a0029433

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *The Behavioral and Brain Sciences, 21*, 803–864. https://doi.org/10.1017/S0140525X98001769

* Haugwitz, M., Nesbit, J. C., & Sandmann, A. (2010). Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learning and Individual Differences*, *20*(5), 536–543. https://doi.org/10.1016/j.lindif.2010.04.004

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*, 557–560. https://doi.org/10.1136/bmj.327.7414.557

Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education, 77*(1), 95–111. https://doi.org/10.1002/sce.3730770107

Hughes, C. A., Morris, J. R., Therrien, W. J., & Benson, S. K. (2017). Explicit instruction: Historical and contemporary contexts. *Learning Disabilities Research & Practice, 32*(3), 140–148. https://doi.org/10.1111/ldrp.12142

* Hwang, G., Shi, Y., & Chu, H. (2011). A concept map approach to developing collaborative Mindtools for context-aware ubiquitous learning. *British Journal of Educational Technology*, *42*(5), 778–789. https://doi.org/10.1111/j.1467-8535.2010.01102.x

* Hwang, G., Yang, L., & Wang, S. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, *69*, 121–130. https://doi.org/10.1016/j.compedu.2013.07.008

* Hwang, G., Zou, D., & Lin, J. (2020). Effects of a multi-level concept mapping-based question-posing approach on students' ubiquitous learning performance and perceptions. *Computers & Education*, *149*. https://doi.org/10.1016/j.compedu.2020.103815

Ioannidis, J. P., & Trikalinos, T. A. (2005). Early extreme contradictory estimates may appear in published research: The Proteus phenomenon in molecular genetics research and randomized trials. *Journal of Clinical Epidemiology, 58*, 543–549. https://doi.org/10.1016/j.jclinepi.2004.10.019

* Jegede, O. J., Alaiyemola, F. F., & Okebukola, P. A. (1990). The effect of concept mapping on students' anxiety and achievement in biology. *Journal of Research in Science Teaching*, *27*(10), 951–960. https://doi.org/10.1002/tea.3660271004

* Jena, A. K. (2019). Concept map in photosynthesis and its effects on achievement and self-classroom management. *International Journal of Research in Teacher Education*, *10*(4), 1–18.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science, 331*(6018), 772–775. https://doi.org/10.1126/science.1199327

* Keown, S. L. (2008). *Effects of the use of thematic organizers in conjunction with concept mapping on learning, misconceptions, and retention in middle school science classes* (Publication No. 3341861) [Doctoral dissertation, Tennessee State University]. ProQuest Dissertation and Theses.

* Kim, P., & Olaciregui, C. (2008). The effects of a concept map-based information display in an electronic portfolio system on information processing and retention in a fifth-grade science class covering the earth's atmosphere. *British Journal of Educational Technology*, *39*(4), 700–714. https://doi.org/10.1111/j.1467-8535.2007.00763.x

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction, 12*(1), 1–10. https://doi.org/10.1016/S0959-4752(01)00014-7

Kirschner, P. A. & Hendrick, C. (2020). *How learning happens: Seminal works in educational psychology and what they mean in practice*. Routledge.

Kirschner, P. A. & Hendrick, C., & Heal, J. (2022). *How teaching happens: Seminal works in teaching and teacher effectiveness and what they mean in practice*. Routledge.

* Kwon, S. Y., & Cifuentes, L. (2007). Using computers to individually-generate vs. collaboratively-generate concept maps. *Journal of Educational Technology & Society*, *10*(4), 269–280.

* Kwon, S. Y., & Cifuentes, L. (2009). The comparative effect of individually-constructed vs. collaboratively-constructed computer-based concept maps. *Computers & Education*, *52*(2), 365–375. https://doi.org/10.1016/j.compedu.2008.09.012

Lau, J., Schmid, C. H., & Chalmers, T. C. (1995). Cumulative meta-analysis of clinical trials builds evidence for exemplary medical care. *Journal of Clinical Epidemiology, 48*, 45–57. https://doi.org/10.1016/0895-4356(94)00106-Z

* Leary, R. (1993). *Effect of concept maps on concept learning and problem-solving achievement in high school chemistry* (Publication No. 3341861) [Doctoral dissertation, Arizona State University]. ProQuest Dissertation and Theses.

* Ledger, A. F. (2003). *The effects of collaborative concept mapping on the achievement, science self-efficacy and attitude toward science of female eighth grade students* (Publication No. 3090329) [Doctoral dissertation, University of Massachusetts Lowell]. ProQuest Dissertation and Theses.

* Ling, Y., & Boo, H. K. (2007). Concept mapping and pupils' learning in primary science in Singapore. *Asia-Pacific Forum on Science Learning and Teaching, 8*(2), 1–28.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage.

Marzetta, K., Mason, H., & Wee, B. (2018). 'Sometimes they are fun and sometimes they are not': Concept mapping with English Language acquisition (ELA) and gifted/talented (GT) elementary students learning science and sustainability. *Education Sciences, 8*(1), 1–13. https://doi.org/10.3390/educsci8010013

Mayr, E. (1997). *This is Biology: The science of the living world*. The Belknap Press of Harvard University Press.

McAuley, L., Pham, B., Tugwell, P., & Moher, D. (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *The Lancet, 356*(9237), 1228–1231. https://doi.org/10.1016/S0140-6736(00)02786-0

* Mcfall, R. E. (1999). *Effects of concept maps and dialogue journals on bilingual students' conceptual understanding in science* (Publication No. 9928205) [Doctoral dissertation, Boston University]. ProQuest Dissertation and Theses.

Merchie, E., & Van Keer, H. (2016). Mind mapping as a meta-learning strategy: Stimulating preadolescents' text-learning strategies and performance? *Contemporary Educational Psychology, 46*, 128–147. https://doi.org/10.1016/j.cedpsych.2016.05.005

Miller, R. D. (2016). Contextualizing instruction for English Language learners with learning disabilities. *Teaching Exceptional Children, 49*(1), 58–65. https://doi.org/10.1177/0040059916662248

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(6), e1000097. https://doi.org/10.1371/journal.pmed1000097

* Morfidi, E., Mikropoulos, A., & Rogdaki, A. (2017). Using concept mapping to improve poor readers' understanding of expository text. *Education and Information Technologies*, *23*(1), 271–286. https://doi.org/10.1007/s10639-017-9600-7

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413–448. https://doi.org/10.3102/00346543076003413

Nesbit, J. C., & Adesope, O. O. (2013). Concept maps for learning, theory, research, and design. In G. Schraw, M. T. McCrudden, & D. Robinson (Eds.), *Learning through visual displays* (pp. 303–328). Information Age Publishing.

Nicoara, S. M., Szabo, B.A., Micu, C. M, & Badea, A. F. (2017). Meta-analysis on the study with concept maps on the medical field. *Transylvanian Journal of Psychology, 18*, 133–170. https://doi.org/10.24193/epsz.2017.2.4

Normand, S. L. (1999). Meta-analysis: Formulating, evaluating, combining, and reporting. *Statistics in Medicine, 18*(3), 321–359. https://doi.org/10.1002/(SICI)1097-0258(19990215)18:3%3C321::AID-SIM28%3E3.0.CO;2-P

Novak, J. D., & Cañas A. J. (2008). *The theory underlying concept maps and how to construct and use them* (Technical Report IHMC Cmap Tools 2006–01 Rev 01–2008). Florida Institute for Human and Machine Cognition.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173469

Novak, J. D. (1991). Clarify with concept maps: A tool for students and teachers alike. *The Science Teacher, 58*, 45–49.

Novak, J. D. (2010a). *Learning, creating, and using knowledge: Concept maps as facilitative tools for schools and corporations* (2nd ed.). New York: Routledge. https://doi.org/10.4324/9780203862001

Novak, J. D. (2010b). Learning, creating and using knowledge: Concept maps as facilitating tools in schools and corporations. *Journal of e-Learning and Knowledge Society, 6*(3), 21–30. https://doi.org/10.4324/9780203862001

Novak, J. D., & Cañas, A. J. (2006). The origins of the concept mapping tool and the continuing evolution of the tool. *Information Visualization, 5*(3), 175–184. https://doi.org/10.1057/palgrave.ivs.9500126

* Ogonnaya, U. P., Okafor, G., Abonyi, O. S., & Ugama, J. O. (2016). Effects of concept mapping instruction approach on students' achievement in basic science. *Journal of Education and Practice*, *7*(8), 79–84.

* Okebukola, P. A. (1992). Concept mapping with a cooperative learning flavor. *The American Biology Teacher*, *54*(4), 218–221. https://doi.org/10.2307/4449458

* Olgun, Ö. S. Ç. (2008). Examining the fifth graders' understanding of heat and temperature concepts via concept mapping. *H. U. Journal of Education*, *34*(34), 54–62.

Oliver, K. (2009). An investigation of concept mapping to improve the reading comprehension of science texts. *Journal of Science Education and Technology, 18*(5), 402–414. https://doi.org/10.1007/s10956-009-9157-3

Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine, 10*(3), 233–240. https://doi.org/10.1111/jebm.12266

Paivio, A. (1986). *Mental representations: A Dual-coding approach*. New York, NY.

*Pankratius, W. J. (1990). Building an organized knowledge base: Concept mapping and achievement in secondary school physics. *Journal of Research in Science Teaching*, *27*(4), 315–333. https://doi.org/10.1002/tea.3660270404

Pullen, P. C., & Hallahan, D. P. (2015). What is special education instruction? In B. Bateman, J. W. Lloyd, & M. Tankersley (Eds.), *Enduring issues in special education: Personal perspectives* (pp. 37–50). Routledge.

Petiti, D. B. (1994). *Meta-Analysis, decision analysis, and cost-effectiveness analysis: Methods for quantitative synthesis in medicine (Monographs in epidemiology and biostatistics)*. Oxford University Press.

* Redford, J. S., Thiede, K. W., Wiley, J., & Griffin, T. D. (2012). Concept mapping improves metacomprehension accuracy among 7th graders. *Learning and Instruction*, *22*(4), 262-270. https://doi.org/10.1016/j.learninstruc.2011.10.007

Reiska, P., Soika, K., Möllits, A., Rannikmäe, M., & Soobard, R. (2015). Using concept mapping method for assessing students' scientific literacy. *Procedia - Social and Behavioral Sciences, 177*, 352–357. https://doi.org/10.1016/j.sbspro.2015.02.357

Rewey, K. L., Dansereau, D. F., Skaggs, L. P., Hall, R. H., & Pitre, U. (1989). Effects of scripted cooperation and knowledge maps on the processing of technical material. *Journal of Educational Psychology, 81*(4), 604–609. https://doi.org/10.1037/0022-0663.81.4.604

Rivet, A. E., & Krajcik, J. S. (2007). Contextualizing instruction: Leveraging students' prior knowledge and experiences to foster understanding of middle school science. *Journal of Research in Science Teaching, 45*(1), 79–100. https://doi.org/10.1002/tea.20203

* Roberts, V., & Joiner, R. (2007). Investigating the efficacy of concept mapping with pupils with autistic spectrum disorder. *British Journal of Special Education*, *34*(3), 127–135. https://doi.org/10.1111/j.1467-8578.2007.00468.x

* Schmid, R. F., & Telaro, G. (1990). Concept mapping as an instructional strategy for high school biology. *The Journal of Educational Research*, *84*(2), 78–85. https://doi.org/10.1080/00220671.1990.10885996

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review, 30*(2), 431–455. https://doi.org/10.1007/s10648-017-9403-9

Serajuddin, U., & Hamadeh, N. (2020). *New World Bank country classifications by income level: 2020–2021*. The World Bank. https://blogs.worldbank.org/opendata/new-world-bank-country-classifications-income-level-2020-2021

* Snead, D. (2000). *Concept mapping and science achievement of middle grade students* (Publication No. 9968097) [Doctoral dissertation, University of Kentucky]. ProQuest Dissertations and Theses.

* Spaulding, D. T. (1989) *Concept mapping and achievement in high school biology and chemistry* (Publication No. 8920449) [Doctoral dissertation, Florida Institute of technology]. ProQuest Dissertation.

* Stanisavljević, J. D., Bunijevac, M. M., & Stanisavljević, L. Ž. (2017). The application of concept maps in the teaching of pollination and pollinators in elementary school. *Journal of Baltic Science Education*, *16*(5), 746–760. https://doi.org/10.33225/jbse/17.16.746

Stata Corporation. (2013). *Stata Statistical Software: Release 13*. StataCorp LP.

* Stensvold, M. S. (1989). *An assessment of concept mapping as a learning tool for the science instructional laboratory* (Publication No. 9004949) [Doctoral dissertation, University of Iowa] ProQuest Dissertations and Theses.

Sturm, J. M., & Rankin-Erickson, J. L. (2002). Effects of hand-drawn and computer-generated concept mapping on the expository writing of middle school students with learning disabilities. *Learning Disabilities Research and Practice, 17*(2), 124–139. https://doi.org/10.1111/1540-5826.00039

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development, 68*, 1–16. https://doi.org/10.1007/s11423-019-09701-3

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer-Verlag.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296. https://doi.org/10.1023/A:1022193728205

* Tastan, I., Dikmenli, M., & Cardak, O. (2008). Effectiveness of the conceptual change texts accompanied by concept maps about students' understanding of the molecules carrying genetical information. In *Asia-Pacific Forum on Science Learning and Teaching,* 9(*1*), 1–13.

* Tekkaya, C. (2003). Remediating high school students' misconceptions concerning diffusion and osmosis through concept mapping and conceptual change text. *Research in Science & Technological Education*, *21*(1), 5-16.https://doi.org/10.1080/02635140308340

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine, 18*(20), 2693–2708. https://doi.org/10.1002/(SICI)1097-0258(19991030)18:20%3C2693::AID-SIM235%3E3.0.CO;2-V

* Udupa, P. S. (1992) *Concept mapping/cooperative learning as a technique to improve the learning of "at-risk" and nondisabled students* (Publication No. 9239159) [Doctoral dissertation, University of Minnesota]. ProQuest Dissertation.

* Uygur, M. (2019). The effects of using digitally supported concept maps method in science classes in primary education on the academic success and students' opinions. *Science Education International*, *30*(3), 209–216. https://doi.org/10.33828/sei.v30.i3.7

* Willerman, M., & Mac Harg, R. A. (1991). The concept map as an advance organizer. *Journal of Research in Science Teaching*, *28*(8), 705–711.https://doi.org/10.1002/tea.36602808

Authors and Affiliations

Dimitris Anastasiou[1] (ORCID) **· Clare Nangsin Wirngo[2] · Pantelis Bagos[3]**

✉ Dimitris Anastasiou
anastasiou@siu.edu

Clare Nangsin Wirngo
cnwirng@ilstu.edu

Pantelis Bagos
pbagos@compgen.org

[1] School of Education, Southern Illinois University Carbondale, Mail Code 4624, 625 Wham Drive, Carbondale, IL 62901, USA

[2] Department of Special Education, Illinois State University Normal, DeGarmo 533, Normal, IL 61790-591, USA

[3] Department of Computer Science and Biomedical Informatics, University of Thessaly, 2-4 Papasiopoulou, 35131 Lamia, Greece