



Co-operative Learning in Undergraduate Mathematics and Science Education: A Scoping Review

Anja Møgelvang¹ · Jorun Nyléhn²

Received: 14 June 2022 / Accepted: 24 October 2022 / Published online: 22 November 2022
© The Author(s) 2022

Abstract

To cope with an unpredictable future, higher education in mathematics and science (MS) needs to educate a knowledgeable and skilled workforce. Co-operative learning (CL) is a teaching method associated with increased academic achievement and development of generic skills. Thus, the purposes of this scoping review are to assess the evidence base of CL in undergraduate MS education to inform teaching practices and to identify potential knowledge gaps to inform future research. The review covers 24 empirical studies conducted from 2010 to 2020 on the prevalence, uses, and outcomes of CL elements in undergraduate MS education. The results show that there are few such studies, and these are rarely conducted outside the US or in disciplines other than chemistry. The most frequently implemented CL elements in the included studies are heterogeneous group formation, the use of roles, and different CL structures. The most prevalent student outcome of implemented CL elements in the reviewed studies is enhanced academic success, followed by student attitudes, generic skills, and psychological health. The results have implications for future implementation of and research on CL in international MS higher education.

Keywords Co-operative learning · Generic skills · Review · Student outcomes · Undergraduate mathematics and science

✉ Anja Møgelvang
anja.mogelvang@gmail.com

Jorun Nyléhn
jorun.nylehn@uib.no

¹ Centre for Excellence in Biology Education, BioCEED, University of Bergen, Thormøhlens Gate 53A, N-5008 Bergen, Norway

² Department of Biological Sciences, University of Bergen, Thormøhlens Gate 53A, N-5008 Bergen, Norway

Introduction

Disciplines in mathematics and science (MS) with their focus on sustainability, innovation, and technology are often viewed as a key to the future (Taylor, 2016). Thus, MS higher educations are expected to prepare a skilled and knowledgeable workforce (Donaldson et al., 2020, p. 722). Since content knowledge within many scientific disciplines is at risk of rapidly becoming outdated (Soler & Dadlani, 2020), MS education communities have long underlined the value of developing generic skills (Johnson & Tisdall, 2002; Leggett et al., 2004), also known as twenty-first century skills (Organization for Economic Co-operation and Development [OECD], 2018; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2016). Generic skills such as collaboration, creativity, and critical thinking (Keane et al., 2016, p. 769) operate across a wide range of contexts (Taber, 2016, p. 226) and may equip the students with tools to update their content knowledge and to navigate in and adapt to an unpredictable future.

Co-operation is recognised as a useful generic skill because co-operation features an essential way of working (Binkley et al., 2012, p. 18). An approach for developing co-operation is *co-operative learning* (CL). CL may be defined as ‘...a highly structured form of group work’ (Millis, 2010, p. 5) and ‘...the instructional use of small groups in which students work together to maximize their own and each other’s learning’ (Johnson et al., 1998b, p. 14). Traditionally, CL has been most common in primary and secondary schools all over the world (Millis & Cottell, 1998), and here CL in MS education seems to have been primarily related to enhanced MS performance (Acar & Tarhan, 2007; Ebrahim, 2012; Eymur & Geban, 2017). As CL found its way into higher education, it continued to be chiefly linked to enhanced student achievement, at least in higher education in general (Loh & Ang, 2020). However, due to its inherent group and task structures, CL may also stimulate the development of the generic skills (Millis & Cottell, 1998; Slavin, 1996), called for by MS higher education communities (Johnson & Tisdall, 2002; Leggett et al., 2004). Nonetheless, we do not know to what degree and how MS discipline higher education puts CL into use. It is also unclear how CL relates to student outcomes in MS discipline higher education.

The aims of this scoping review are twofold. The first is to review the evidence base regarding the uses and outcomes of CL in undergraduate MS education to inform teaching practices. The second is to identify knowledge gaps within the field to inform future research. Thus, we pose four review questions:

1. Which disciplines, countries, and research methods are prevalent in studies of co-operative learning elements in undergraduate MS education?
2. What are the characteristics of the co-operative learning elements used in and principles guiding undergraduate MS education?

3. *What are the student outcomes of co-operative learning elements in undergraduate MS education?*
4. *How are the various co-operative learning elements associated with student outcomes?*

Background

Co-operative learning (CL) stems from social interdependence theory (Deutsch, 2012) and has primarily been developed by educational psychologists and brothers David and Roger Johnson (Johnson & Johnson, 1989). According to Millis and Cottell (1998, p. 11), it is important that the implementation of CL in higher education adheres to two principles in particular: positive interdependence and individual accountability.

CL Principles: Positive Interdependence and Individual Accountability

The purpose of the first principle—positive interdependence—is to create an authentic sense of mutual gain and a shared goal (Millis & Cottell, 1998, p. 11). Positive interdependence is achieved by structuring a range of elements in ways that makes group members dependent on each other and causes them to work together to successfully complete the task (Ballantine & Larres, 2007, p. 128). The purpose of the second principle—individual accountability—is to promote responsibility and prevent social loafing (Millis & Cottell, 1998, p. 12). Individual accountability is achieved when the teacher includes a mechanism, for example, individual tests, for holding group members accountable for learning the material and completing the group task (Ballantine & Larres, 2007, p. 128). When introducing CL into one's teaching, these two principles should guide every element of the CL process from group features, goals, tasks, resources, roles, structures, to rewards. A focus on and a conscious approach to all these elements makes CL a much more highly structured teaching and learning strategy than other forms of small-group learning, e.g. collaborative learning. Opposed to CL, collaborative learning is characterised by looser structures and rely mainly on very few elements, except for task and goal, to guide the collaborative process (Davidson, 2021; Millis & Cottell, 1998, pp. 7–10). Thus, collaborative learning teachers only rarely or never consider elements such as group features, role assignments, team-building activities, co-operative structures, equal participation, or activist interventions in their teaching (Davidson, 2021). According to CL literature, the success of CL—and hence the reasons faculty should consider implementing CL in their teaching—lies in the structured and conscious approach to elements such as group features and CL structures. These are elaborated upon in the following and exemplified in Box 1.

Box 1 Frequent and well-known co-operative learning elements in the reviewed studies

Co-operative Instructional Modelling: a teaching method based on the group features of CL and the Modelling Theory of Science (Hestenes, 1987). This method is characterised by active engagement of students in co-operative groups and emphasis on conceptual development by modelling scientific activities.

Co-operative Peer Review Structures: covers a wide concept, comprising CL principles and peer review or peer feedback (Ladyshevsky, 2013). Group members give and receive peer review on product and process, installing both positive interdependence and individual accountability in the group.

Group Contract: provides guidelines for group work and group tasks. The purpose of the contract is to establish common expectations and provide the group members with tools to develop constructive communication and manage potential conflicts (Oakley et al., 2004)

Jigsaw: each group member takes responsibility for learning a specific part of a complex whole and teaching it to the rest of the group. This way the group, by working together, put all the pieces of the jigsaw together (Millis & Cottell, 1998, p. 127).

POGIL: Process Oriented Guided Inquiry Learning is an instructional group-learning strategy comprising a set of rules and structures based on Kolb's learning cycle and CL principles such as small, fixed groups and rotating roles (Process Oriented Guided Inquiry Learning [POGIL], 2019). It was developed for chemistry education but is currently used in a wide range of subjects and disciplines.

Rotating Roles: complementary tasks and responsibilities are prescribed to ensure both the principle of positive interdependence and individual accountability. Popular roles are *Facilitator/leader*, *Recorder/evaluator*, *Elaborator*, *Summariser*, and *Monitor* and an important feature is that the roles rotate between the group members on a regular basis (Cohen, 2010).

STAD: Student Team Achievement Division is a co-operative learning strategy where small groups of students with different levels of ability are working together to accomplish a shared goal (Slavin, 1991).

Think-Pair-Share/Square: This CL technique is suitable for many different teaching scenarios, ranging from lectures, seminars, and laboratory exercises. The teacher poses a question that needs reflection and gives each student time to reflect individually. Next, the students are asked to pair up and discuss their thoughts or responses to the question before they share their joint answer with the entire class (share) or in their groups (square) (Millis & Cottell, 1998, p. 73).

CL Group Features: Group Size, Formation, and Duration

Most literature on CL in higher education agrees that group size should be between three and five students and most seem to prefer groups of four (Kagan, 2021; Millis & Cottell, 1998, p. 13). When students work in small groups of four, social loafing might be avoided, less forthright students can express their opinion, pairing up is easy, and even if a person is missing, the group is still technically a group (Johnson et al., 1998b; Millis & Cottell, 1998). Compared to students working individually, students' performance, knowledge, and achievement seem to be higher when students work in such smaller groups (Bertucci et al., 2010; Lou et al., 1996, 2001).

Diversity of opinion and experiences may create a cognitive disequilibrium (Piaget, 1985) and force the students to take different perspectives and argue their case. Thus, CL literature (Johnson et al., 1998b; Kagan, 2021; Millis & Cottell, 1998) recommends that groups should be formed by the teacher based on heterogeneous principles, i.e. different academic ability, background, age, and gender. Lou et al.

(1996) found that low-ability students learn more in heterogeneous groups and Jacobs et al. (2006) argue that higher-ability students may also benefit from CL, building their sense of autonomy and an opportunity to care for others. In male-dominated groups, the level and nature of knowledge transfers within groups is significantly lower (Hansen et al., 2015) than in female-dominated groups, and the proportion of women in groups positively predicts discussion quality that in turn predicts group (academic) performance (Curşeu et al., 2018).

Depending on purpose, CL groups may last a short or long period of time. Formal CL groups typically last from one class to several weeks or months and are suited to teach specific content. Informal CL groups are ad hoc groups which last from few minutes to one class, and they are used to ensure that students actively process information during a lecture. CL base groups typically last at least one year and are meant to provide long-term support in order to make academic progress and build committed relationships (Johnson et al., 1992, 1998b).

CL Structures

CL structures are content-free strategies (Kagan, 2021) which organise the interaction of students by prescribing student behaviour step-by-step to complete the assignment (Johnson et al., 1998b; Kagan, 2021). The benefits of these structures are that they may be employed in any subject and on any educational level, including higher education while being designed to ensure that positive interdependence and individual accountability occur. Highly structured groups and group tasks help students understand how they are to work together, contribute, take responsibility, and help each other learn (Johnson & Johnson, 1999). Gillies (2003, 2008) discovered that students in structured groups compared to peers in unstructured groups exhibited more co-operative behaviour and demonstrated more complex thinking and problem-solving skills. In a systematic review of secondary and post-secondary courses, Romero (2009) found that the effect on student achievement was greater for structured than unstructured CL interventions. Thus, the evidence base seems to suggest that the highly structured and conscious approach characteristic of CL may benefit both the group process and individual outcomes.

Outcomes of CL

Academic Success In a meta-analysis by Johnson et al. (1998a), the effect of CL on academic achievement was found to be significantly higher compared to competitive learning environments and individualistic learning environments. Another meta-analysis in undergraduate STEM education by Springer et al. (1999) supported these findings. More recent meta-analyses (Apugliese & Lewis, 2017; Kyndt et al., 2013) and systematic reviews (Romero, 2009) show similar results concerning the association between CL and academic achievement in higher education generally.

Student Attitudes Johnson et al. (2014) claim that it is through discussions in co-operative groups that students learn and model the norms and values of university life and that CL thus makes up an effective tool for improving student attitudes.

A meta-analysis comprising CL studies conducted in universities internationally (Johnson et al., 1998a) show that CL seems to improve student attitudes compared to competitive university learning environments and individualistic learning.

Generic Skills Millis and Cottell (1998) and Slavin (1996) suggest that CL may lead to improved generic skills. Although failing to satisfy the inclusion criteria of this review due to insufficient information on the CL elements used (Rattanatumma & Puncreobutr, 2016; Sandi-Urena et al., 2012) or wrong study focus (Winschel et al., 2015), these three studies may cast light on the relationship between CL and generic skills in undergraduate STEM education. Two of these were conducted in undergraduate chemistry (Sandi-Urena et al., 2012; Winschel et al., 2015), and both found that different co-operative lab instructions relate to an increase in the students' problem-solving skills. A study in undergraduate mathematics (Rattanatumma & Puncreobutr, 2016) supported these findings. These studies show the potential for CL to strengthen the problem-solving skills of MS students in higher education.

Psychological Health Due to the structured group work, peer relationships, and negotiation of social skills, CL elements may also promote socialisation and psychological health (Gillies, 2016; Johnson et al., 2014). One of the health benefits hypothesised to be affected positively by CL is sense of belonging which is regarded a basic human need (Deci & Ryan, 2000). In their meta-analysis, Johnson et al. (2014) showed that co-operation fostered both greater interpersonal attraction and perceived social support among students than did competing with others or working alone.

Although not included in this review due to wrong population, a few college science studies have examined self-efficacy (Bandura, 1997) and specific types of academic self-efficacy in relation to CL (Espinosa et al., 2019; Rivera, 2013). In a study in introductory algebra, the implementation of CL elements improved the students' mathematics self-efficacy significantly (Rivera, 2013). Similarly, Espinosa et al. (2019) found that physics self-efficacy increased significantly for women and reduced the gender gap in physics self-efficacy following teaching approaches in an introductory physics class based on CL principles.

Methods

Research Design

To inform an ongoing research-based redesign process targeting student generic skills such as co-operation in undergraduate education at the Faculty of Mathematics and Science in a Norwegian University, we conducted a scoping review. A scoping review seeks to provide thorough reviews of available literature and identify possible knowledge gaps through analyses of the answers to the review questions (Arksey & O'Malley, 2005). Thus, we reviewed recent empirical studies to examine the prevalence, use of CL elements, and student outcomes in undergraduate MS

discipline education, using a systematic approach in five steps: (1) identifying the review questions, (2) identifying the relevant studies, (3) selecting the studies, (4) charting the data, and (5) collating, summarising, and reporting the results (Arksey & O'Malley, 2005).

Step 1: Identifying the Review Questions

We began by identifying key concepts such as study population, intervention, and outcome (Arksey & O'Malley, 2005). The PICO (Population, Intervention, Comparison, and Outcome) (Oliver et al., 2017, p. 76) model was useful in this process. By applying the components of the PICO model on the proposed review questions, we determined the following key concepts: **Population**=Students in undergraduate mathematics, physics, chemistry, biology, and geology, **Intervention**=Exposure to CL learning elements, **Comparison**=Not relevant in this review, and **Outcome**=All types of student CL outcomes.

The PICO model served two purposes in this review. The first purpose was to ensure validity through conceptual framework: only review questions containing key concepts such as population, CL elements intervention, and CL outcome were subject to examination and analysis. The second purpose was to ensure validity through methodology: the identified key concepts helped guide the review process, from search strategies via screening procedures to data extraction.

Step 2: Identifying the Relevant Studies

In the development of a search strategy to identify relevant studies, it is important to consider both sensitivity and specificity. Sensitivity ensures a high proportion of relevant studies and specificity ensures a low proportion of irrelevant studies (Brunton et al., 2017). Relevance is one of several means to ensure validity preventing both selection bias and publication bias (Booth et al., 2016). Thus, we searched databases containing studies in specific subjects or disciplines and databases containing studies of all disciplines. The selected databases searched were ERIC, Proquest Education, PsycINFO, Web of Science, and Google Scholar. To avoid publication bias (Krumsvik & Røkenes, 2016) and ensure further validity, this review also searched

Table 1 Database search field strings

Database search field	Search string
1st line	<i>(cooperati* NEAR/2 learn* OR collaborati* NEAR/2 learn* OR team* NEAR/2 learn*)</i>
2nd line	<i>("higher education" OR college* OR universit* OR postsecondary OR "post secondary" OR graduate* OR undergraduate* OR tertiary OR bachelor*)</i>
3rd line	<i>(stem OR math* OR physic* OR chemi* OR biolog* OR geolog*)</i>
4th line	<i>(Outcome* OR result* OR effect OR effects OR skill* OR competenc* OR knowledge OR achievement* OR performanc*OR benefit* OR impact*)</i>
5th line	1 AND 2 AND 3 AND 4

the grey literature (Booth et al., 2016, p. 120) database OpenGrey and the online source Higher Education Academy.

Database search strings were developed in collaboration with a university librarian and based on the key concepts in the review question. The key concepts were first linked by the Boolean operator OR and second by the Boolean operator AND. Truncation and proximity operators were additional tools used to balance sensitivity and specificity in all database search fields (Table 1). The search strategy varied according to database, and a full overview of all search strategies in each database is provided in Supplemental Material 1. Full overviews provide transparency, an important aspect of auditability and reliability (Booth et al., 2016; Brunton et al., 2017; Koffel & Rethlefsen, 2016).

Step 3: Selecting the Studies

To select only relevant studies, we developed a set of inclusion and exclusion criteria. These criteria were carefully selected to inform the aforementioned redesign process. Thus, the population identified for inclusion was students in global undergraduate (Bachelor) MS *discipline* courses comprised of the following subjects or disciplines: mathematics, physics, chemistry, biology, and geology. Studies needed to employ *highly structured* in-class CL elements (s) in groups of 3–6 people based on one or more of the guiding principles of CL. To assess the relationship between CL elements and student outcomes, only primary studies with this particular focus were included, and precise information about the amount and type of CL elements and outcomes used were required. Due to scarce information in abstracts and conference papers, these types of publication were excluded. To obtain only recent studies, the time limit was set to 2010–2020, and the language restrictions were based on the language skills of the reviewers and included studies published in English, Danish, Norwegian, and Swedish. For a full overview of the criteria, see Supplemental Material 2.

The selection of the studies for inclusion was conducted by means of the review tool Rayaan (Rayaan, 2022) and further followed the four-step PRISMA process as recommended by Moher et al. (2009), i.e. identification, screening, eligibility, and inclusion. Both titles and abstracts ($n = 1847$) and full-text articles ($n = 105$) were screened independently and ultimately; 24 studies were included (Fig. 1).

Discrepancies concerning the suitability of studies, during both screening stages, were solved through two processes: (1) discussion and clarification of the inclusion and exclusion criteria to ensure a common understanding and (2) thorough common review of the texts in question based on the inclusion and exclusion criteria anew. Both the clear-cut inclusion and exclusion criteria guiding the screening process and the systematic and independent approach guided by the PRISMA protocol prevent selection and publication bias and thus strengthen clarity, reliability, and validity of the review (Booth et al., 2016; Brunton et al., 2017).

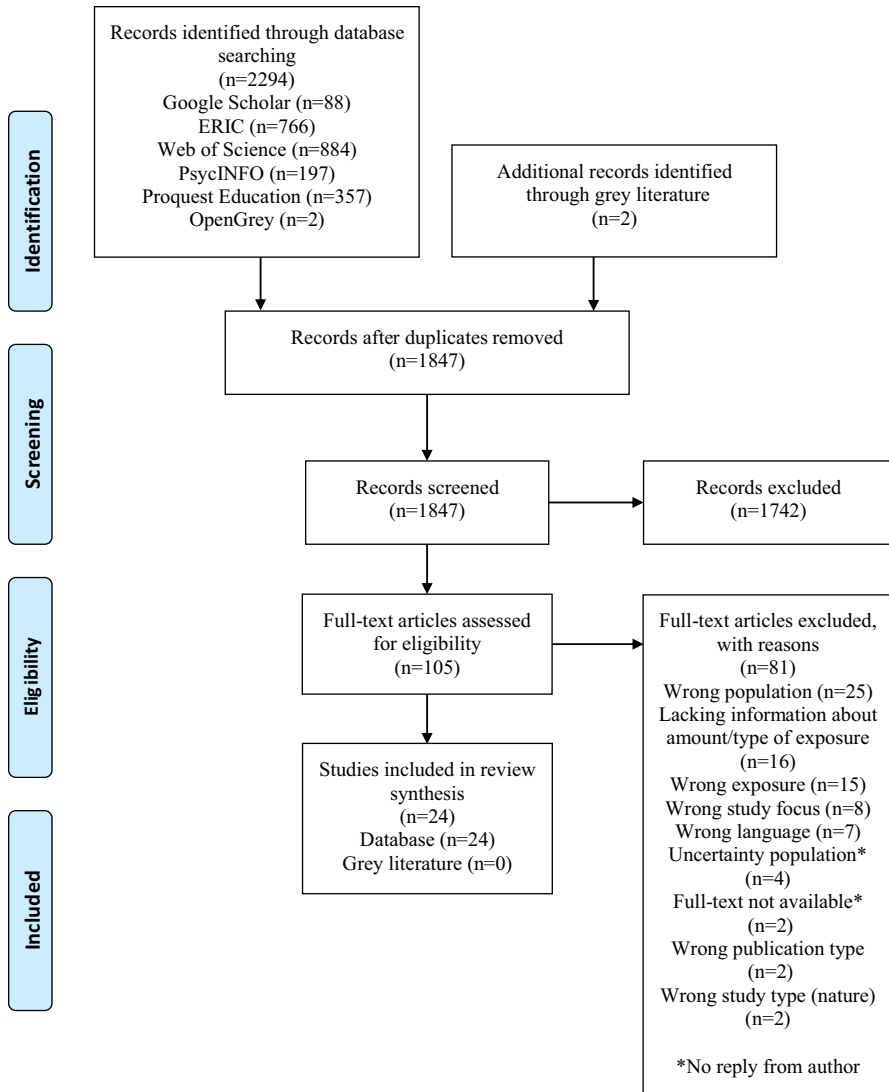


Fig. 1 PRISMA flow diagram (Moher et al., 2009)

Step 4: Charting the Data

To chart or extract data from the included studies means to report the key items of information obtained from the included studies (Arksey & O'Malley, 2005). The purposes of this step are to gain an overview of the included studies and to identify the information needed to answer the review questions. Thus, the review questions guided this process, and the first author extracted the following information: authors, year of publication, study locations, subject/discipline, research methods,

intervention duration, CL elements and CL principles, outcome measures, and results.

Step 5: Collating, Summarising, and Reporting the Results

In this scoping review, we mapped and organised the data chartered for each of the four review questions.

Results

The chartered data of the included studies, guided by the four research questions, are mapped, organised, and reported in Table 2. Thus, Table 2 covers the answers to all four research questions and contains the following chartered data from the 24 reviewed studies: author, discipline, country, research methods, data collection, group size, group formation, group duration, CL structure, CL principle, type of outcome, and result of outcome.

Disciplines, Countries, and Research Methods

The 24 studies included in this review were predominantly conducted in five disciplines, and chemistry ($n=11$) was by far the most represented. Following chemistry, we identified studies in biology ($n=7$), physics ($n=5$), and mathematics ($n=4$), among others. Further, we found great differences in geographical distribution. The USA ($n=16$) constituted an overwhelming majority and counting the neighbouring countries of Canada ($n=2$) and the commonwealth of Puerto Rico ($n=1$), North America in total made up 79% ($n=19$) of the reviewed studies. We also found studies from the African countries South Africa and Ethiopia ($n=2$) and from Turkey ($n=2$) and Indonesia ($n=1$) in Asia. None of the reviewed studies were from Europe. Most of the included studies made use of quantitative methods ($n=14$), followed by mixed methods studies ($n=7$), and qualitative studies ($n=3$).

CL Elements

In terms of group size, a clear majority of the studies employed groups of four members ($n=21$) and a minority, groups of three members ($n=3$). Most groups were formed by the teacher ($n=10$) and were heterogeneous ($n=8$). Some groups were also student-selected ($n=6$) and homogeneous ($n=2$). Most studies employed groups lasting from several hours and weeks to one semester ($n=17$), while a few lasted one class, test or the like ($n=4$), and the rest ($n=3$) did not report on duration.

Table 2 Author, disciplines, countries, research methods, co-operative learning (CL) elements, co-operative (CL) principles, and types and results of student outcomes in the reviewed studies

Author	(1) Bierema et al. (2017)	(2) Brewé et al. (2010)	(3) Byrne (2015)	(4) Canelas et al. (2017)	(5) Carson and Glaser (2010)	(6) Chase et al. (2013)	(7) Cheruvelil et al. (2020)	(8) Daniel (2016)
Discipline	Biology	Physics	Math	Chemistry	Chemistry	Chemistry	Biology	Biology
Country	USA	USA	USA	USA	USA	USA	USA	USA
Method	Qualitative	Quantitative	Mixed	Quantitative	Quantitative	Quantitative	Quantitative	Mixed
Data collection	Recording interview	Survey	Survey Assessment Interview	Exam scores Survey	Survey	Survey	Survey	Grades Attendance records Survey Minute-papers Quizzes Exams
Group size	3	3	3-4	4-6	3-5	3-4	3-4	4
Group formation	Student-selected	No info	Teacher-selected Heterogeneous	Random Teacher-selected Heterogeneous	Student-selected	No info	No info	Student-selected
Group duration	One task One semester	No info	One semester	One test	One semester	Three discussion sessions	14 weeks	One semester
CL structure	Co-operative instructional modelling	Co-operative instructional modelling	Roles	POGIL	Intra-group peer review and grading	POGIL Roles	Teamwork syllabus objective, exercises, reflection, assessment	Roles Jigsaw
CL principle(s)	IA	No info	PI	No info	IA & PI	No info	IA & PI	IA
Type outcome	Content knowledge	Academic achievement Content knowledge	Content knowledge	Academic achievement Content knowledge Generic skills	Academic achievement Generic skills Attitudes	Academic achievement Retention Attitudes Psychological health	Generic skills Attitudes	Academic achievement Attendance
Positive change in outcome(s)	Yes	Yes	Yes	Yes (only in generic skills)	Yes	No	Yes	Yes

Table 2 (continued)

Author	(9) Diaz-Vázquez et al. (2012)	(10) Furuto (2017)	(11) Gok (2018)	(12) Harlow et al. (2016)	(13) Hein (2012)	(14) Leung et al. (2017)	(15) Ott et al. (2018a)	(16) Ott et al. (2018b)
Discipline	Chemistry	Math	Physics	Physics	Chemistry	Physics	Chemistry	Biology Chemistry Environmental science Mechanical engineering Physics
Country	Puerto Rico	USA	Turkey	Canada	USA	Canada	USA	USA
Method	Quantitative	Mixed	Mixed	Quantitative	Quantitative	Quantitative	Quantitative	Mixed
Data collection	Survey Tests Exams	Observation Video recordings Interviews Journals Survey	Survey Test Written essay	Test Exam Evaluation	Exam	Test Survey	Exam grades Retention records	Survey Focus groups
Group size	4	3–4	2–3	3 vs. 4	4	4	4	4
Group formation	Teacher-selected Student-selected	Teacher-selected Heterogeneous	Random	Teacher-selected Heterogeneous Homogeneous	Student-selected	Random	Random	Teacher-selected Heterogeneous
Group duration	One semester	One semester	One class	One semester	Three hours	One laboratory session	Four weeks	One semester
CL structure	Roles Group investigation Intra-group peer review	Roles STAD	Think-pair-share	Group composition	POGIL	Assessment selection	POGIL Roles	Roles Team evaluation Group contract
CL principle	IA and PI	IA	IA and PI	No info	No info	IA and PI	IA and PI	IA and PI

Table 2 (continued)

Author	(9) Diaz-Vázquez et al. (2012)	(10) Furuto (2017)	(11) Gok (2018)	(12) Harlow et al. (2016)	(13) Hein (2012)	(14) Leung et al. (2017)	(15) Ott et al. (2018a)	(16) Ott et al. (2018b)
Type outcome	Academic achievement Content knowl- edge Generic skills	Participation Psychological health	Content knowl- edge Attitudes	Academic achievement Content knowl- edge	Academic achievement Content knowl- edge Attitudes Grading duty	Academic achievement Content knowl- edge Attitudes Grading duty	Academic achievement Retention	Generic skills Attitudes
Positive change in outcome(s)	Yes	Yes	Yes	No	Yes	Yes (only attitudes, and grading duties)	Yes	Yes (only generic skills)
Author	(17) Pilcher et al. (2015)	(18) Srougi et al. (2013)	(19) Stanford et al. (2016)	(20) Timungki (2015)	(21) Warfa et al. (2018)	(22) Wilton et al. (2019)	(23) Yapici (2016)	(24) Yimer and Feza (2020)
Discipline	Chemistry	Biological sci- ences Chemistry	Chemistry	Math	Chemistry	Biology	Biology	Math
Country	South Africa	USA	USA	Indonesia	USA	USA	Turkey	Ethiopia
Method	Mixed	Quantitative	Qualitative	Mixed	Qualitative	Quantitative	Quantitative	Quantitative
Data collection	Survey Observation Interviews	Course grades Survey	Video record- ings Group discus- sions Discourse analysis	Test Observation Interview	Audio record- ings of group dialogue Discourse analysis	Course grades Survey Retention records	Test Survey	Achievement test Survey
Group size	3-4	4-5	3-4	4-5	3-4	4-6	4-5	4-5
Group formation	Teacher-selected Homogeneous	Teacher-selected Heterogeneous	No info	Teacher-selected Heterogeneous	Student-selected	No info	Teacher-selected Heterogeneous	No info
Group duration	Six weeks	Two weeks	Five weeks	No info	One semester	Ten weeks	One semester	No info

Table 2 (continued)

Author	(17) Pilcher et al. (2015)	(18) Srougi et al. (2013)	(19) Stanford et al. (2016)	(20) Timungki (2015)	(21) Warfa et al. (2018)	(22) Wilton et al. (2019)	(23) Yapici (2016)	(24) Yimer and Feza (2020)
CL structure	Roles Jigsaw	Co-operative take-home exams Individual tests	POGIL Roles	TAI	POGIL Roles	Think-Pair-Share	STAD Group rules Group name	Jigsaw
CL principle	IA	IA	No info	IA and PI	PI	No info	IA and PI	No info
Type outcome	Content knowl- edge Generic skills Attitudes	Academic achievement Generic skills Attitudes	Content knowl- edge	Content knowl- edge	Content knowl- edge	Academic achievement Retention Psychological health	Academic achievement Psychological health Attitudes	Content knowl- edge Attitudes
Positive change in outcome(s)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: CL co-operative learning, POGIL process oriented guided inquiry learning, STAD student team achievement division, TAI team-assisted individualisation, IA individual accountability, PI positive interdependence

The most common CL structures featured in the reviewed studies were the use of roles ($n=10$) and POGIL ($n=6$). Other employed CL structures in the reviewed studies included jigsaw ($n=3$), STAD ($n=2$), Think-Pair-Share ($n=2$), co-operative instructional modelling ($n=2$), peer review structures ($n=2$), group contract ($n=2$), and several more. See Box 1 for a description of these CL structures.

The CL principles of positive interdependence and individual accountability underpin CL elements. However, less than half of the studies reported having included both CL principles ($n=9$), a scarce minority reported having included one of the two CL principles ($n=7$), and the rest of the studies ($n=8$) did not report having included either of these two CL principles.

CL Outcomes

In the reviewed studies, student outcomes of CL elements were largely related to academic success, in the form of content knowledge ($n=8$), academic achievement or performance, in this review combined and called “academic achievement” ($n=7$), or both ($n=6$). Other frequent student outcomes measured in the included studies were attitudes towards the discipline, the learning process or group work ($n=10$), different generic skills ($n=7$), and different types of psychological outcomes ($n=4$). The majority ($n=19$) of the included studies found only positive results of the implemented CL elements. A few studies identified both positive results and some negative results ($n=3$) or no positive results at all ($n=2$).

Associations Between CL Elements and Outcomes

Studies employing teacher-selected heterogenous groups ($n=8$) were first and foremost associated with academic success, i.e. content knowledge ($n=4$), academic achievement ($n=4$), or both ($n=2$). However, other types of outcomes such as generic skills ($n=3$), attitudes ($n=3$), psychological health ($n=2$), and participation ($n=1$) were also represented. Studies employing student-selected groups ($n=6$) were also mostly associated with content knowledge ($n=4$), academic achievement ($n=4$), or both ($n=2$). Other outcome types counted generic skills ($n=2$), attitudes ($n=1$), and attendance ($n=1$). Two of the studies employing teacher-selected heterogenous groups reported a positive change in generic skills only, not in other outcomes (4, 16), while the student-selected groups studies reported a positive change in all outcomes.

Studies employing longer lasting groups, i.e. formal groups ($n=17$) were associated with all chartered outcomes, and a vast majority ($n=14$) reported positive changes in all outcomes. Three studies did not report positive changes in all outcomes (6, 12, 16). Studies employing groups of short duration, i.e. informal groups ($n=4$) were mainly associated with academic success, i.e. content knowledge ($n=4$), academic achievement ($n=2$) or both ($n=2$), followed by attitudes ($n=2$), generic skills ($n=1$), and grading duties, i.e. workload ($n=1$). Half of these four studies failed to identify a positive change in academic success (4, 14).

The most used CL structure, roles ($n=10$), was primarily associated with either content knowledge ($n=6$), academic achievement ($n=4$), or both ($n=1$). Further, use of roles was associated with attitudes ($n=3$), generic skills ($n=3$), psychological health ($n=3$), retention ($n=2$), attendance ($n=1$), and participation ($n=1$). Eight of the ten studies applying roles led to a positive change in all outcomes, while one led to a positive change in one of the outcomes only (16), and one did not lead to any positive change in any of the outcomes (6). These two studies both failed to measure a positive change in attitudes following the use of roles.

The second most used CL structure, POGIL ($n=6$), was primarily associated with content knowledge ($n=4$), academic achievement ($n=4$), or both ($n=2$). Other outcomes associated with POGIL were retention ($n=2$), generic skills ($n=1$), attitudes ($n=1$), and psychological health ($n=1$). Four of these studies reported positive changes in outcome, but two of these studies did not find a positive change in academic achievement (4, 6).

The third most used CL structure, jigsaw ($n=3$), was associated with content knowledge ($n=2$), attitudes ($n=2$), academic achievement ($n=1$), generic skills ($n=1$), and attendance ($n=1$). All the jigsaw studies reported a positive change in all outcomes. In addition to jigsaw, studies employing other CL structures ($n=13$), e.g. Think-Pair-Share and STAD to name a few, were associated with all chartered outcomes, and all of these, except for two (12, 14), identified a positive change in outcomes.

Discussion

Analysis of Disciplines, Countries, and Research Methods

First, few studies exist in almost all undergraduate MS disciplines except chemistry, and further research on CL elements and outcomes in other MS disciplines is needed. The reason why chemistry stands out is unknown, but it may be connected to the popularity of the POGIL method (Walker & Warfa, 2017). Second, we identified few studies outside of North America, and no studies at all from Europe, which may be said to represent a knowledge gap. Research results are not necessarily transferable to other continents, countries, or cultures, and therefore further research and knowledge on CL elements and outcomes in undergraduate MS education in different parts of the world are needed. The reason that so many of the studies were conducted in the USA may be due to the American origin of CL (Deutsch, 2012; Johnson & Johnson, 1989). Third, most of the studies were quantitative. Although quantitative data are valuable, they may not give us a full in-depth understanding of students' perceptions nor explain why CL leads to certain student outcomes in undergraduate MS education. The present lack of qualitative studies represents yet another knowledge gap within the field of CL in MS higher education. For such knowledge, faculty planning studies within the field might consider employing qualitative methods.

Analysis of CL Elements

A vast majority of the reviewed studies met the recommendations of group size but not of heterogeneity from previous CL research (Johnson & Johnson, 1999; Millis, 2010; Millis & Cottell, 1998). Some of the studies reported on student-selected and homogeneous groups, while others made use of random group formation without specifying who formed the groups. A quarter of the studies did not report anything about group formation. Of the eight reviewed studies where groups were formed heterogeneously by the teacher, all seemed to take ability and/or gender into consideration when forming the groups (3, 4, 10, 12, 16, 18, 20, 23). Most of the groups in the studies lasted from several hours, classes, or weeks to one semester and may thus be characterised as formal CL groups, while a minority of the groups were informal CL groups, lasting only one class, one test, or the like. It may be of some concern that studies in international undergraduate MS education examine outcomes of CL while not necessarily following the recommendations given in the CL literature. This lack of coherence and alignment with theory may lead to invalid results as it may become unclear what these studies are actually studying.

The most common CL structures in the reviewed studies were roles and POGIL. Role is usually a fixed feature of POGIL, and to some degree that might explain the number of the reviewed studies employing roles ($n=10$). Six of the ten reviewed studies using roles mentioned that the students took rotating roles (3, 8, 10, 15, 16, 19), and five of these identified a positive change in outcomes. As shown in Box 1, rotating roles is a CL structure underpinned by both the principles of positive interdependence and individual accountability. Taken together, these findings may indicate that implementing CL structures which underpin positive interdependence and individual accountability seems to be of significance in undergraduate MS education. These indications support previous research on CL structures in other subject disciplines in both higher education and elsewhere (Gillies, 2003, 2008; Johnson & Johnson, 1999; Johnson et al., 1998a; Romero, 2009). Considering that the principles of positive interdependence and individual accountability underpin CL teaching (Gillies, 2016), it may be of some concern that many of the included studies did not mention them. Voicing the principles might create a more conscious approach, ensuring that future implementation of CL and research on CL are in accordance with the underlying theory.

Analysis of CL Outcomes

Of the 21 reviewed studies which included content knowledge and/or academic achievement as the outcome measure, 17 reported an improvement. Similarly, in eight of the ten studies examining student attitudes, improvement was found. Thus, the findings of this review add to the extensive evidence research base regarding the positive relationships between CL and academic success (e.g. Apugliese & Lewis, 2017; Kyndt et al., 2013; Romero, 2009) and CL and student attitudes (Johnson et al., 1998a)—albeit in undergraduate MS education. These relationships may according to Deutsch (2012) and CL literature (Johnson & Johnson, 1989, 1999;

Johnson et al., 2014) be explained by the common goal, and interaction takes place in CL groups. When students work together to achieve a common goal, i.e. when they are positively interdependent, academic success enhances—and it is in discussions in CL groups that students learn and model the norms and values of university, making CL an effective tool for improving student attitudes.

Seven of the reviewed studies examined the hypothesis that CL elements may lead to the development of student generic skills (Millis & Cottell, 1998; Slavin, 1996). All these studies found support for this hypothesis. In the reviewed studies, generic skills related to CL elements were teamwork skills ($n=4$), problem-solving skills ($n=1$), critical thinking/higher thinking skills ($n=2$), communication skills ($n=1$), and metacognitive skills ($n=1$). Prior studies have mainly concentrated on problem-solving skills in relation to CL elements in higher MS education (Rattana-tumma & Puncreobutr, 2016; Sandi-Urena et al., 2012; Winschel et al., 2015), but this review identifies several additional generic skills.

In four of the reviewed studies, CL elements were related to sense of belonging ($n=2$), academic self-efficacy ($n=1$), or both ($n=1$). Three of the four reviewed studies reported positive findings regarding sense of belonging (10, 22, 23) and academic self-efficacy (10), and that may be considered important. Research indicates that students with a strong sense of belonging create a positive student identity (Sanders & Munford, 2016), and high self-efficacy (Bandura, 1997) is a strong predictor for performance and persistence in MS education (Espinosa et al., 2019).

That the vast majority of the included studies found only positive results of the implemented CL elements and very few studies found partly or no positive results at all may be due to publication bias (Ekholm & Chow, 2018; Francis, 2012). Although this review searched grey literature (Booth et al., 2016; Krumsvik & Røkenes, 2016) in attempt to avoid publication bias, we cannot exclude that it has played a role. Further, it should be noted that several of these studies employed more than one CL structure. Thus, it is not possible to know if any positive increase in outcome would have been due to one certain CL structure over another, the combination of CL structures, or other reasons. Taken together, the results should be approached with some caution, and more research, which may cast light on such issues, is needed to strengthen the evidence base.

Analysis of the Association Between CL Elements and Outcomes

Group Formation Many of the reviewed studies did not meet the recommendations of most CL literature regarding group formation (Johnson & Johnson, 1999; Kagan, 2021; Millis & Cottell, 1998). Yet, when group formation was held up against outcomes, there was no evidence that teacher-selected heterogenous groups led to more positive outcomes than did the student-selected groups. This apparent gap is worth mentioning—but it is hard to identify a reason. It may be that the population, i.e. international undergraduate MS education, differs from other undergraduate populations or students in higher education differ from students in schools. It may also be that the effect of group composition lessened in combination with other CL

structures. Or it may also be due to more random reasons altogether. If causality is to be determined here, more research is needed.

Group Duration Two of the four studies employing informal CL groups (4, 14) found no improvement in academic success. This may indicate that duration could be important to obtain enhanced academic success from CL elements in MS undergraduate education and perhaps that formal CL groups could be more suited. Duration may also play a role in the development of psychological health. Three of the four studies (10, 22, 23) examining sense of belonging and/or academic self-efficacy were all characterised by groups lasting for a minimum of 10 weeks. By lasting a certain length of time and allowing the students to partake in several social and personal experiences, the CL intervention may have enhanced the students' sense of belonging and academic self-efficacy in the process.

Roles Not only was academic success the most measured outcome in the studies employing roles, but all of these, except for one (6), reported a positive change. Thus, roles may be a suited CL element to enhance academic success. On the other hand, two of three studies associating roles with attitudes found no improvement in students' attitudes (6, 16). This does not necessarily mean that roles are not suited to improve student attitudes as hypothesised by Johnson et al. (2014). However, if they are to do that in undergraduate MS education, it may according to (16) themselves be important that roles are perceived to have a purpose and contribute to team productivity. Also, it may be that roles when applied in POGIL are dependent on the study activities containing all required elements as prescribed by the POGIL method (6). Roles were the most used CL structure in studies with generic skills as outcome, and all of these reported a positive change. Thus, roles may also be appropriate to develop undergraduate MS students' generic skills. However, it should be noted that all studies, independent of CL element, identified a positive change in generic skills. This may both indicate that (i) CL elements may lead to the development of generic skills as hypothesised by CL literature (e.g. Millis & Cottell, 1998; Slavin, 1996) and (ii) that many different CL elements may be appropriate in doing so.

POGIL Four of the six studies featuring POGIL found that that POGIL increased academic success in undergraduate chemistry education. The two studies not identifying increased academic success assigned this lack to several causes: (i) the use of study activities which did not contain all required elements as prescribed by the POGIL method and implementation of POGIL in a small proportion of the courses (6) and (ii) lectures incorporating some student-centred activities and thus possibly reducing differences between control and experimental groups (4). Taken together, POGIL may be a suitable CL element to increase academic success in undergraduate chemistry education given the prescription by POGIL is followed and confounding variables controlled.

Jigsaw and Other Structures The studies employing jigsaw all found positive changes in outcomes such as academic achievement, content knowledge, attendance, generic

skills, and attitudes (8, 17, 24). Two of the 13 studies employing other CL structures such as Think-Pair-Share and STAD did not find enhanced academic success (12, 14). A reason for the lack of enhanced academic success may according to (12) themselves be that the studied classes already used student-centred teaching strategies which may have lessened the sensitivity to the implemented CL changes. In (14) the reasons for the lack of enhanced academic success are less clear, but their study underlines other benefits from the CL elements such as improved attitudes to cooperation and reduced assessment workload. Taken together, other CL structures may lead to a range of different positive outcomes, but if these impacts are to be measured, they may need to be isolated from other student-centred approaches.

Conclusions and Implications

The goal of this scoping review was to assess the evidence base of CL in undergraduate education in MS to inform teaching practices and to identify important knowledge gaps. We identified 24 studies and found that studies examining CL elements in undergraduate education in MS are relatively few, primarily quantitative in nature, almost non-existent outside the North American continent, and mainly conducted in chemistry. The reviewed studies employed many different CL elements of which some were not in accordance with CL theory and research. Further, relatively few of the included studies report on both of the guiding CL principles positive interdependence and individual accountability. Studies of CL elements in MS higher education are associated primarily and positively with enhanced academic success, but also generic skills and psychological outcomes seem to be linked positively to CL elements.

In sum, there is a need to design studies which explore CL using qualitative methods, in other countries than the USA and perhaps especially in Europe and in more undergraduate MS disciplines. If the impact of CL elements is to be measured in additional quantitative studies, it seems of importance to isolate CL elements from other student-centred approaches and control even more for confounding variables. Also, in filling the existing knowledge gaps, future research should concentrate on student outcomes other than enhanced content knowledge and academic achievement. Both research and teaching practices may benefit from addressing CL group features, structures, and principles corresponding to CL theory. Further, faculty contemplating CL elements in their undergraduate MS education may need to be aware that student outcomes seem to be somewhat dependent on the underpinning of the CL principles and duration of the groups. POGIL and roles are the most used CL elements in the reviewed studies, and both may be suited in undergraduate MS education given the POGIL prescriptions are followed, and roles are perceived as purposeful and contributing to team processes and outcomes.

Limitations

Our review focused solely on undergraduate MS *discipline* education, and this may have resulted in a limited number of relevant studies. During the screening

stage, it became clear that many studies of CL elements took place in undergraduate MS *professional* studies, particularly in study programs for pre-service teachers which might be transferable to MS discipline studies. Another limitation is that of inferring meaning from omission. Not voicing the principles for instance may be due to many reasons. Perhaps such omissions may simply be indicative of the nature of the journals in which they were published.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10763-022-10331-0>.

Acknowledgements The authors thank Anne Grete Danielsen and Kathrine Cohen from the University of Bergen in Norway for valuable help throughout the study process.

Funding Open access funding provided by University of Bergen (incl Haukeland University Hospital).

Declarations

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acar, B., & Tarhan, L. (2007). Effect of cooperative learning strategies on students' understanding of concepts in electrochemistry. *International Journal of Science and Mathematics Education*, 5(2), 349–373. <https://doi.org/10.1007/s10763-006-9046-7>
- Apugliese, A., & Lewis, S. E. (2017). Impact of instructional decisions on the effectiveness of cooperative learning in chemistry through meta-analysis. *Chemistry Education Research and Practice*, 18(1), 271–278. <https://doi.org/10.1039/c6rp00195e>
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32.
- Ballantine, J., & Larres, P. M. C. (2007). Cooperative learning: A pedagogy to improve students' generic skills? *Education and Training*, 49(2), 126–137. <https://doi.org/10.1108/00400910710739487>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Bertucci, A., Conte, S., Johnson, D., & Johnson, R. (2010). The impact of size of cooperative group on achievement, social support, and self-esteem. *The Journal of General Psychology*, 137, 256–272.
- Bierema, A.M.-K., Schwarz, C. V., & Stoltzfus, J. R. (2017). Engaging undergraduate biology students in scientific modelling: Analysis of group interactions, sense-making, and justification. *CBE Life Sciences Education*, 16(4), ar68. <https://doi.org/10.1187/cbe.17-01-0023>
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 17–66). Springer.
- Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review*. Sage.
- Brewe, E., Sawtelle, V., Kramer, L. H., O'Brien, G. E., Rodriguez, I., & Pamelá, P. (2010). Toward equity through participation in modeling instruction in introductory university physics. *Physical Review*

- Special Topics. Physics Education Research*, 6(1), 010106. <https://doi.org/10.1103/PhysRevSTPER.6.010106>
- Brunton, G., Stansfield, C., Caird, J., & Thomas, J. (2017). Finding relevant studies. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 93–122). Sage.
- Byrne, M. (2015). *Changes in student proving skills and attitudes following a cooperative learning seminar*. Proquest Dissertations Publishing.
- Canelas, D. A., Hill, J. L., & Novicki, A. (2017). Cooperative learning in organic chemistry increases student assessment of learning gains in key transferable skills. *Chemistry Education Research and Practice*, 18(3), 441–456. <https://doi.org/10.1039/C7RP00014F>
- Carson, K. M., & Glaser, R. E. (2010). Chemistry is in the news: Assessing intra-group peer review. *Assessment and Evaluation in Higher Education*, 35(4), 381–402. <https://doi.org/10.1080/02602930902862826>
- Chase, A., Pakhira, D., & Stains, M. (2013). Implementing process-oriented, guided-inquiry learning for the first time: Adaptations and short-term impacts on students' attitude and performance. *Journal of Chemical Education*, 90(4), 409–416. <https://doi.org/10.1021/ed300181t>
- Cheruvellil, K. S., De Palma-Dow, A., & Smith, K. A. (2020). Strategies to promote effective student research teams in undergraduate biology labs. *The American Biology Teacher*, 82(1), 18–27. <https://doi.org/10.1525/abt.2020.82.1.18>
- Cohen, M. W. (2010). Cooperative learning in educational psychology. Modeling success for future teachers. In B. J. Millis (Ed.), *Cooperative learning in higher education* (pp. 69–89). Stylus.
- Curşeu, P. L., Chappin, M. M. H., & Jansen, R. J. G. (2018). Gender diversity and motivation in collaborative learning groups: The mediating role of group discussion quality. *Social Psychology of Education*, 21(2), 289–302. <https://doi.org/10.1007/s11218-017-9419-5>
- Daniel, K. L. (2016). Impacts of active learning on student outcomes in large-lecture biology courses. *The American Biology Teacher*, 78(8), 651–655. <https://doi.org/10.1525/abt.2016.78.8.651>
- Davidson, N. (2021). Introduction to pioneering perspectives in cooperative learning. In N. Davidson (Ed.), *Pioneering perspectives in cooperative learning. Theory, research, and classroom practice for diverse approaches to CL* (pp. 1–16). Routledge.
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behaviour. *Psychological Inquiry*, 11(4), 227–268.
- Deutsch, M. (2012). A theory of cooperation – competition and beyond. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology: Volume 2* (pp. 275–294). SAGE Publications Ltd. <https://doi.org/10.4135/9781446249222.n40>
- Díaz-Vázquez, L. M., Montes, B., Echevarría Vargas, I., Hernandez-Cancel, G., Gonzalez, F., Molina, A., Morales-Cruz, M., Torres-Díaz, C., & Griebenow, K. (2012). An investigative, cooperative learning approach for general chemistry laboratories. *International Journal for the Scholarship of Teaching and Learning*, 6(2), Article 20. <https://doi.org/10.20429/ijstol.2012.060220>
- Donaldson, T., Fore, G. A., Filippelli, G. M., & Hess, J. L. (2020). A systematic review of the literature on situated learning in the geosciences: Beyond the classroom. *International Journal of Science Education*, 42(5), 722–743. <https://doi.org/10.1080/09500693.2020.1727060>
- Ebrahim, A. (2012). The effect of cooperative learning strategies on elementary students' science achievement and social skills in Kuwait. *International Journal of Science and Mathematics Education*, 10(2), 293–314. <https://doi.org/10.1007/s10763-011-9293-0>
- Ekholm, E., & Chow, J. (2018). Addressing publication bias in educational psychology. *Translational Issues in Psychological Science*, 4(4), 425–439. <https://doi.org/10.1037/tps0000181>
- Espinosa, T., Miller, K., Araujo, I., & Mazur, E. (2019). Reducing the gender gap in students' physics self-efficacy in a team- and project-based introductory physics class. *Physical Review Physics Education Research*, 15(1), 010132.
- Eymur, G., & Geban, Ö. (2017). The collaboration of cooperative learning and conceptual change: Enhancing the students' understanding of chemical bonding concepts. *International Journal of Science and Mathematics Education*, 15(5), 853–871. <https://doi.org/10.1007/s10763-016-9716-z>
- Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. *Psychonomic Bulletin & Review*, 19(6), 975–991. <https://doi.org/10.3758/s13423-012-0322-y>
- Furuto, M. (2017). Mathematics pedagogical strategies to create a positive college classroom community. *Malaysian Journal of Mathematical Sciences*, 11(1), 9–21.
- Gillies, R. M. (2003). Structuring cooperative group work in classrooms. *International Journal of Educational Research*, 39(1), 35–49. [https://doi.org/10.1016/S0883-0355\(03\)00072-7](https://doi.org/10.1016/S0883-0355(03)00072-7)

- Gillies, R. M. (2008). The effects of cooperative learning on junior high school students' behaviours, discourse and learning during a science-based learning activity. *School Psychology International*, 29, 328–347.
- Gillies, R. M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education*, 41(3), 39–54.
- Gok, T. (2018). The evaluation of conceptual learning and epistemological beliefs on physics learning by Think-Pair-Share. *Journal of Education in Science, Environment and Health*, 4(1), 69–80.
- Hansen, Z., Owan, H., & Pan, J. (2015). The impact of group diversity on class performance: Evidence from college classrooms. *Education Economics*, 23(2), 238–258. <https://doi.org/10.1080/09645292.2013.813908>
- Harlow, J. J. B., Harrison, D. M., & Meyertholen, A. (2016). Effective student teams for collaborative learning in an introductory university physics course. *Physical Review. Physics Education Research*, 12(1), 010138. <https://doi.org/10.1103/PhysRevPhysEducRes.12.010138>
- Hein, S. M. (2012). Positive impacts using POGIL in organic chemistry. *Journal of Chemical Education*, 89(7), 860–864. <https://doi.org/10.1021/ed100217v>
- Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics*, 55(5), 440–454.
- Jacobs, G. M., McCafferty, S. G., & Iddings, C. (2006). Roots of cooperative learning in general education. In S. G. McCafferty, G. M. Jacobs, & C. Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 9–17). Cambridge University Press.
- Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction Book Company.
- Johnson, D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory into Practice*, 38(2), 67–73. <https://doi.org/10.1080/00405849909543834>
- Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1992). *Advanced cooperative learning* (2nd ed.). Interaction Book Company.
- Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). *Active learning: Cooperation in the university classroom* (2nd ed.). Interaction Book Company.
- Johnson, D. W., Johnson, R. T., & Smith, K. A. (2014). Cooperative learning: Improving university instruction by basing practice on validated theory. *Journal on Excellence in College Teaching*, 25(3–4), 85–118.
- Johnson, H. S., & Tisdall, J. (2002). Encouraging generic skills in science courses. *Electronic Journal of Biotechnology*, 5(2), 133–140. <https://doi.org/10.2225/vol5-issue2-fulltext-13>
- Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998a). Cooperative learning returns to college: What evidence is there that it works? *Change*, 30(4), 26–35.
- Kagan, S. (2021). The structural approach and Kagan structures. In N. Davidson (Ed.), *Pioneering perspectives in cooperative learning* (pp. 78–127). Routledge.
- Keane, T., Keane, W. F., & Blicblau, A. S. (2016). Beyond traditional literacy: Learning and transformative practices using ICT. *Education and Information Technologies*, 21(4), 769–781.
- Koffel, J. B., & Rethlefsen, M. L. (2016). Reproducibility of search strategies is poor in systematic reviews published in high-impact pediatrics, cardiology and surgery journals: A cross-sectional study. *PLoS One*, 11(9), e0163309. <https://doi.org/10.1371/journal.pone.0163309>
- Krumsvik, R. J., & Røkenes, F. M. (2016). Litteraturreview i ph.d.-avhandlingen [Literature reviews in the PhD dissertation]. In R. J. Krumsvik (Ed.), *En doktorgradsuddanning i endring: Et fokus på den artikkelbaserte ph.d.- avhandlingen* [A doctoral education in change: A focus on the article-based PhD dissertation] (pp. 51–91). Fagbokforlaget.
- Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review*, 10, 133–149.
- Ladyszewsky, R. K. (2013). The role of peers in feedback processes. In D. Boud & E. Molloy (Eds.), *Feedback in higher and professional education: Understanding it and doing it well* (pp. 174–189). Routledge.
- Leggett, M., Kinnear, A., Boyce, M., & Bennett, I. (2004). Student and staff perceptions of the importance of generic skills in science. *Higher Education Research and Development*, 23(3), 295–312. <https://doi.org/10.1080/0729436042000235418>
- Leung, A. C. K., Hashemi Pour, B., Reynolds, D., & Jerzak, S. (2017). New assessment process in an introductory undergraduate physics laboratory: An exploration on collaborative learning.

- Assessment and Evaluation in Higher Education*, 42(2), 169–181. <https://doi.org/10.1080/02602938.2015.1089977>
- Loh, R. C. Y., & Ang, C. S. (2020). Unravelling cooperative learning in higher education: A review of research. *Research in Social Sciences and Technology*, 5(2), 22–39. <https://doi.org/10.46303/fressat.05.02.2>
- Lou, Y., Abrami, P., Spence, J., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66, 423–458.
- Lou, Y., Abrami, P., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, 71, 449–521.
- Millis, B. J. (2010). Why faculty should adopt cooperative learning approaches. In B. J. Millis (Ed.), *Cooperative learning in higher education* (pp. 1–9). Stylus Publishing LLC.
- Millis, B. J., & Cottell, P. G. (1998). *Cooperative learning for higher education faculty*. American Council on Education/Oryx Press.
- Moher, D., Liberati, A., Altman, D. G., & The PRISMA group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Oakley, B., Felder, R. M., Brent, R., & Elhaji, I. (2004). Turning student groups into effective teams. *Journal of Student Centered Learning*, 2(1), 9–34.
- Oliver, S., Dickson, K., Bangpan, M., & Newman, M. (2017). Getting started with a review. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 71–91). Sage.
- Organization for Economic Co-operation and Development [OECD]. (2018). *The Future of Education and Skills 2030*. Retrieved November 08, 2022 from [https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf)
- Ott, L. E., Carpenter, T. S., Hamilton, D. S., & LaCourse, W. R. (2018a). Discovery learning: Development of a unique active learning environment for introductory chemistry. *The Journal of Scholarship of Teaching and Learning*, 18(4), 161–180. <https://doi.org/10.14434/josotl.v18i4.23112>
- Ott, L. E., Kephart, K., Stolle-McAllister, K., & LaCourse, W. R. (2018b). Students' understanding and perceptions of assigned team roles in a classroom laboratory environment. *Journal of College Science Teaching*, 47(4), 83–91. https://doi.org/10.2505/4/jcst18_047_04_83
- Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. University of Chicago Press.
- Pilcher, L. A., Riley, D. L., Mathabathe, K. C., & Potgieter, M. (2015). An inquiry-based practical curriculum for organic chemistry as preparation for industry and postgraduate research. *South African Journal of Chemistry*, 68, 236–244.
- Process Oriented Guided Inquiry Learning [POGIL]. (2019). *What is POGIL*. Retrieved November 08, 2022 from <https://pogil.org/about-pogil/what-is-pogil>
- Rattanamumma, T., & Puncreobutr, V. (2016). Assessing the effectiveness of STAD model and problem based learning in mathematics learning achievement and problem solving ability. *Journal of Education and Practice*, 7(12), 194–199.
- Rivera, N. (2013). *Cooperative learning in a community college setting: Developmental coursework in mathematics* [Doctoral dissertation, Arizona State University]. ProQuest Dissertations Publishing. Retrieved November 08, 2022 from <https://www.cpedinitiative.org/assets/docs/rivera.pdf>
- Romero, C. C. (2009). *Cooperative learning instruction and science achievement for secondary and early post-secondary students: A systematic review*. [Doctoral dissertation, Colorado State University]. ProQuest Dissertations Publishing. Retrieved November 08, 2022 from <https://www.proquest.com/docview/304862896>
- Sanders, J., & Munford, R. (2016). Fostering a sense of belonging at school—five orientations to practice that assist vulnerable youth to create a positive student identity. *School Psychology International*, 37(2), 155–171.
- Sandi-Urena, S., Cooper, M., & Stevens, R. (2012). Effect of cooperative problem-based lab instruction on metacognition and problem-solving skills. *Journal of Chemical Education*, 89(6), 700–706.
- Slavin, R. E. (1991). *Student team learning: A practical guide to cooperative learning*. National Education Association.
- Slavin, R. E. (1996). Research for the future. Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21, 46–69.
- Soler, M. G. & Dadlani, K. (2020, August 13). *Resetting the way we teach science is vital for all our futures*. Weforum. Retrieved November 08, 2022 from <https://www.weforum.org/agenda/2020/08/science-education-reset-stem-technology/>

- Springer, L., Donovan, S. S., & Stanne, M. E. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69, 21–51.
- Srougi, M. C., Miller, H. B., Witherow, D. S., & Carson, S. (2013). Assessment of a novel group-centered testing schema in an upper-level undergraduate molecular biotechnology course. *Biochemistry and Molecular Biology Education*, 41(4), 232–241. <https://doi.org/10.1002/bmb.20701>
- Stanford, C., Moon, A., Towns, M., & Cole, R. (2016). Analysis of instructor facilitation strategies and their influences on student argumentation: A case study of a process oriented guided inquiry learning physical chemistry classroom. *Journal of Chemical Education*, 93(9), 1501–1513. <https://doi.org/10.1021/acs.jchemed.5b00993>
- Taber, K. S. (2016). Learning generic skills through chemistry education. *Chemistry Education Research Practise*, 17, 225–228.
- Taylor, P. C. (2016) Why is a STEAM curriculum perspective crucial to the 21st century? In: *14th Annual conference of the Australian Council for Educational Research*, August 07-09, 2016, Brisbane (pp. 89-93). Retrieved November 08, 2022 from https://research.acer.edu.au/cgi/viewcontent.cgi?article=1299&context=research_conference
- Tinungki, G. M. (2015). The role of cooperative learning type team assisted individualization to improve the students' mathematics communication ability in the subject of probability theory. *Journal of Education and Practice*, 6(32), 27–31.
- United Nations Educational, Scientific and Cultural Organization [UNESCO]. (2016). *Unpacking Sustainable Development Goal 4: Education 2030; guide*. Retrieved November 08, 2022 from <https://unesdoc.unesco.org/ark:/48223/pf0000246300>
- Walker, L., & Warfa, A. R. M. (2017). Process oriented guided inquiry learning (POGIL®) marginally effects student achievement measures but substantially increases the odds of passing a course. *PLoS ONE*, 12(10), e0186203–e0186203. <https://doi.org/10.1371/journal.pone.0186203>
- Warfa, A. R. M., Nyachwaya, J., & Roehrig, G. (2018). The influences of group dialog on individual student understanding of science concepts. *International Journal of STEM Education*, 5(1), 46–14. <https://doi.org/10.1186/s40594-018-0142-3>
- Wilton, M., Gonzalez-Niño, E., McPartlan, P., Terner, Z., Christoffersen, R. E., & Rothman, J. H. (2019). Improving academic performance, belonging, and retention through increasing structure of an introductory biology course. *CBE Life Sciences Education*, 18(4), ar53. <https://doi.org/10.1187/cbe.18-08-0155>
- Winschel, G., Everett, R., Coppola, B., & Shultz, G. (2015). Using jigsaw-style spectroscopy problem-solving to elucidate molecular structure through online cooperative learning. *Journal of Chemical Education*, 92(7), 1188–1193.
- Yapici, I. M. (2016). Effectiveness of blended cooperative learning environment in biology teaching: Classroom community sense, academic achievement, and satisfaction. *Journal of Education and Training Studies*, 4(4), 269–280.
- Yimer, S. T., & Feza, N. N. (2020). Learners' conceptual knowledge development and attitudinal change towards calculus using Jigsaw cooperative learning strategy integrated with GeoGebra. *International Electronic Journal of Mathematics Education*, 15(1), em0554.