Cooperative Learning in Undergraduate STEM Education: Applications and Outcomes

Anja Møgelvang

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2023

UNIVERSITY OF BERGEN

Cooperative Learning in Undergraduate STEM Education: Applications and Outcomes

Anja Møgelvang

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 14.03.2023

Year: 2023

Title: Cooperative Learning in Undergraduate STEM Education: Applications and Outcomes

Name: Anja Møgelvang

Print: Skipnes Kommunikasjon / University of Bergen

Scientific Environment

This PhD thesis was conducted and funded under the bioCEED Centre for Excellence in Biology Education at the Department of Biological Sciences, Faculty of Mathematics and Natural Sciences, University of Bergen, Norway. The research was situated in the bioBERG research group, and I attended the research school Western Norway Graduate School of Educational Research II (WNGER II).

Acknowledgements

"A lot has derailed, but I'm on track!" is probably the phrase I have used most often when people have asked about my progress during this PhD thesis. It turns out that writing a PhD thesis on student cooperative learning in the midst of a pandemic which closed down campuses worldwide, comes with challenges. However, with challenges come unique opportunities and now, standing at the finish line and looking back I think the project in many ways grew stronger from what initially seemed to be a hopeless situation. I owe much of this to the wonderful people I have surrounded myself with, professionally and privately.

To my supervisors Sehoya Cotner, Vigdis Vandvik, and Jorun Nyléhn from bioCEED and the Department of Biological Sciences, thank you from the bottom of my heart. You have each contributed to this project in different ways – all equally valuable and when combined; Unbeatable. What a dream team! Sehoya, you came to my rescue at a critical time, offering the roots and wings I needed. Vigdis, I have learned so much from your sharp insights and authoritative style. Jorun, your quick and helpful feedback together with your availability have provided me with the support I have needed to progress and finish this project. I would also like to express my gratitude to Anne Grete Danielsen at the Department of Education. Your ideas and input have been crucial to the development of this project.

Warm thanks to my co-authors Jorun Nyléhn, Sehoya Cotner, Vigdis Vandvik, Ståle Ellingsen, and Christian Bianchi Strømme. My three papers would not have been what they are without your valuable feedback, perspectives, and experience. Your support has been very important to me. A special thanks goes to you, Ståle! For trusting me, being so positive, and letting me carry out my research in your course. It made all the difference! By extention, I wish to thank all the students who participated in my studies. You are at the heart of my research - it is all about you!

My profund appreciation goes to bioCEED for welcoming me into your home. I cannot imagine a warmer and more including place in all of academia. You are all truly wonderful. Jonathan, my office companion, and French gentleman (and frog), you are very special to me and sharing office with you these past years has been everything good in life - from balloons, Friday buns, and great talks to heartly laughter. Oddfrid, you are the backbone of bioCEED and my go-to-person – in times of need and in times of joy. We have in many ways set out on similar paths in life and I can always come to you for support. Kristin, the office ray of sunshine and one of the kindest persons in the world. During my time with bioCEED you and I have shared great experiences and I will never forget our concert in Århus. You, me, and Nick! Lucas, although no longer based at bioCEED, you have extended your help whenever I needed it. Thank you!

I also wish to thank the bioBERG research group. I have learned a lot from our discussions and your input.

To my dearest friends, thank you for always believing in me and giving me the necessary time-outs and the space to just be!

Last but not least, I wish to express my deepest graditude to my family. To Arve, my rock and eternal love. You are always there. Your interest in this project, your devotion and belief in me mean everything. To my dearest children, Olander, Regitze and Falk, Thank you for giving me perspective and reminding me of what is truly important in this world. You are my everything!

Abstract

Initiatives to strengthen psychosocial outcomes by, for example increasing students' perceived sense of belonging, generic skills, and science confidence, and reducing loneliness are important for several reasons. To successfully address current and complex global challenges, international cooperation is critical, and the world is trusting the scientific communities to step up to provide important knowledge, skills, and solutions. Science, Technology, Engineering, and Mathematics (STEM) knowledge is particularly important, but STEM students and teachers struggle to prepare for the task, facing long-time recruitment, retention, and learning outcome challenges. Psychosocial factors may play a vital role in overcoming these three challenges but seem to have been overlooked and under-researched in STEM higher education. A teaching and learning method hypothesized to promote cooperation and improve a range of psychosocial outcomes is a highly structured form of group work known as cooperative learning (CL). Thus, in this thesis I examine how CL is applied in undergraduate STEM education and how it is related to psychosocial outcomes, i.e., perceived sense of belonging, generic skills, science confidence, and loneliness, among the students. The purpose is to assess and further develop the evidence base of CL in undergraduate STEM education to: a) address existing STEM recruitment, retention, and learning outcome challenges, b) inform teaching practices, c) fill current knowledge gaps in the field, and d) guide future research.

The dissertation is article-based and consists of three papers and a synopsis. Paper I is a scoping review of 24 empirical studies on the applications and outcomes of CL in international mathematics and science education. Using a systematic five step approach, our 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. Paper II is a cross-sectional survey study examining the relationship between Norwegian

undergraduate STEM students' perceived levels of CL, sense of belonging, and generic skills. By means of correlation analyses and standard multiple regression analyses, we found that CL was positively associated with the development of both sense of belonging and generic skills, that sense of belonging and generic skills are positively interrelated, and that "promotive interaction" was the CL principle contributing most to the association with both sense of belonging and generic skills. The findings were discussed using social interdependence theory. Paper III reports on findings from a study conceived during the COVID-19 pandemic. Although posing challenges for students, teachers, and educational research, the pandemic also offered a unique opportunity to develop and examine digital teaching and learning methods. Thus, Paper III is a quasi-experiment testing the effect of digital CL compared to digital lectures on a range of psychosocial outcomes in a sample of Norwegian undergraduate biology students. The study was a one-group pretest-posttest design with a double pretest and follow-up. Our results showed that the students' self-reported psychosocial outcomes, i.e., sense of belonging, science confidence, and generic skills increased, and their loneliness decreased, after five weeks of digital CL compared to five weeks of digital lectures. In light of theory and previous research, we argue that these findings indicate that the effect of CL may be at least as substantial in digital settings as in face-to-face settings - and that psychosocial outcomes may be particularly vulnerable in digital settings and thus dependent on the digital teaching and learning strategy chosen.

As a whole, the thesis identifies and fills several knowledge gaps within the field of CL, ranging from (a) settings not previously researched, i.e., Norwegian undergraduate STEM education and digital teaching and learning settings, (b) identifying specific CL methods suited for these settings, (c) examining student outcomes other than academic success, (d) determining "promotive interaction" as the most important CL principle in some of the key student outcomes, and (e) providing a theoretical framework, i.e., social interdependence theory. The findings of the three papers indicate that in undergraduate STEM education, CL methods can be successfully implemented and lead to a range of positive psychosocial outcomes in both physical and digital teaching and learning settings. Further, the synopsis showcases how the three papers are connected, describes the alignment within the approach, and discusses the implications and limitations of the main findings.

Sammendrag

Tiltak for å styrke opplevde psykososiale faktorer slik som tilhørighet, generiske ferdigheter og vitenskapelig selvtillit samt tiltak for å redusere ensomhet hos studenter anses som viktig av flere grunner. I møtet med de pågående og komplekse globale utfordringene setter verden sin lit til internasjonalt samarbeid og forventer at forskningsmiljøene bidrar med de nødvendige kunnskaper, ferdigheter og løsninger. Kunnskap som utvikles innen realfag blir vurdert som særskilt viktig, men både studentene og lærerne innen realfag strever med å forberede seg på oppgaven. Gjennom lang tid har realfagsfagmiljøene hatt utfordringer med rekruttering, gjennomføring og læringsutbytte - og selv om psykososiale faktorer kan spille en viktig rolle i å løse disse utfordringene, har det blitt forsket lite på slike faktorer i høyere realfagsutdanning. En metode som fremmer samarbeid og som antas å ha en positiv innvirkning på psykososiale faktorer i utdanning, er samarbeidslæring. Samarbeidslæring er en svært strukturert form for gruppearbeid mellom studenter. I denne avhandlingen undersøker jeg hvordan samarbeidslæring kan anvendes i bachelorutdanninger innen realfag og hvordan samarbeidslæring er relatert til psykososiale faktorer slik som studenters opplevde tilhørighet, generiske ferdigheter, vitenskapelige selvtillit, og ensomhet, både internasjonalt og i Norge. Formålet med avhandlingen er å evaluere og videreutvikle evidensgrunnlaget for samarbeidslæring innen bachelorutdanninger i realfag når det gjelder: a) å møte eksisterende utfordringer med rekruttering, gjennomføring og læringsutbytte, b) belyse undervisningspraksis, c) fylle nåværende kunnskapshull og d) gi retning til fremtidig forskning.

Avhandlingen er artikkelbasert og består av tre vitenskapelige artikler og en kappe. Artikkel I er en omfattende systematisk litteraturgjennomgang som identifiserer 24 empiriske studier som tar for seg bruken og resultatene av samarbeidslæring i internasjonale bachelorutdanninger innen realfag. Den systematiske fem-trinns tilnærmingen som vi anvender i artikkel I viser at det finnes svært få slike studier, at slike studier sjeldent utføres utenfor USA og i andre disipliner enn kjemi. Videre ser vi at de elementene fra samarbeidslæring som anvendes hyppigst i de inkluderte 24

studiene er heterogen gruppedannelse, bruk av roller, og ulike samarbeidslæringsstrukturer. Avslutningsvis viser artikkel I at det overveiende flertall av de inkluderte studiene har knyttet samarbeidslæring opp mot økt læringsutbytte og/eller akademiske prestasjoner. Til sammenligning er studentenes holdninger, generiske ferdigheter og psykisk helse i liten grad undersøkt. Artikkel II er en kvantitativ tverrsnittsundersøkelse som tar for seg sammenhengen mellom samarbeidslæring, opplevd tilhørighet og generiske ferdigheter blant norske bachelorstudenter i realfag. Ved hjelp av korrelasjonsanalyser og standard regresjonsanalyser fant vi ut at (i) samarbeidslæring var sterkt positivt korrelert med både tilhørighet og opplevde generiske ferdigheter, (ii) tilhørighet og opplevde generiske ferdigheter var også positivt og sterkt korrelerte, og at (iii) prinsippet «støttende interaksjon» i samarbeidslæring bidro mest til sammenhengen med både tilhørighet og opplevde generiske ferdigheter. Alle tre funn ble i artikkel II drøftet ut fra teorien om sosial gjensidig avhengighet. Artikkel III rapporterer funn fra en studie som ble til under koronapandemien. På tross av en rekke utfordringer for både studenter, lærere og utdanningsforskning, ga pandemien oss en unik mulighet til å utvikle og undersøke digitale samarbeidslæringsmetoder. Derav er artikkel III et kvasi-eksperiment som tester effekten av digital samarbeidslæring sammenlignet med digitale forelesninger på en rekke psykososiale utfall hos et utvalg av norske bachelorstudenter i biologi. Studien ble designet som en såkalt pretest-posttest studie uten kontrollgruppe, men med bruk av dobbel pretest og posttest. Etter fem uker med digital samarbeidslæring sammenlignet med fem uker med digitale forelesninger økte studentenes opplevde følelse av tilhørighet, vitenskapelige selvtillit og generiske ferdigheter mens ensomhet avtok. I lys av teori og tidligere forskning argumenterer vi i artikkel III for at funnene kan indikere at effekten av samarbeidslæring kan være minst like betydningsfull i digitale som i fysiske arenaer – og at psykososiale faktorer kan være særskilt sårbare i digitale arenaer og dermed også avhengige av hvilken undervisnings- og læringsstrategi vi som lærere velger.

Samlet sett identifiserer og fyller avhandlingen en rekke kunnskapsfull innen samarbeidslæring slik som (a) kontekster som det ikke tidligere er forsket på, dvs.

bachelorutdanninger innen realfag i Norge og digitale undervisnings- og læringsarenaer, (b) å finne frem til spesifikke samarbeidslæringsmetoder som er egnet for disse kontekstene, (c) å undersøke andre faktorer enn akademiske prestasjoner, (d) å identifisere «støttende interaksjon» som det viktigste av samarbeidslæringsprinsippene for utvalgte faktorer og (e) å sette samarbeidslæring inn i et teoretisk rammeverk i form av teorien om sosial gjensidig avhengighet. Funnene i de tre delstudiene indikerer at flere samarbeidslæringsmetoder er godt egnet i internasjonale og norske bachelorutdanninger innen realfag og at disse kan føre til en rekke positive psykososiale gevinster i både fysiske og digitale undervisnings- og læringsmiljø. Videre viser kappen hvordan de tre delstudiene belyser hverandre og hvordan den overordnede tilnærmingen i avhandlingen henger sammen før funn, implikasjoner og begrensninger avslutningsvis diskuteres.

List of Publications

Paper I

Møgelvang, A. & Nyléhn, J. (2022). Co-operative learning in undergraduate mathematics and science education: A scoping review. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-022-10331-0

Paper II

Møgelvang, A., & Nyléhn, J. (In press). Interdependence between perceived cooperative learning, sense of belonging, and generic skills in undergraduate STEM education. *Nordic Journal of STEM Education.*

Paper III

Møgelvang, A., Vandvik, V., Ellingsen, S., Strømme, C. B., & Cotner, S. (2023). Cooperative learning goes online: Teaching and learning intervention in a digital environment impacts psychosocial outcomes in biology students. *International Journal of Educational Research, 117*, 102114. https://doi.org/10.1016/j.ijer.2022.102114

Author Contributions

Contribution	Paper I	Paper II	Paper III
Research design	**AM**	**AM**	**AM,** SE
Data collection	**AM**	**AM**	**AM**
Data analysis	**AM**, JN	**AM**	**AM**
Data interpretation	**AM,** JN	**AM,** JN	**AM,** VV, CBS, SHC
Writing of manuscript	**AM,** JN	**AM**, JN	**AM,** VV, SHC
Commenting on manuscript	JN	JN, SHC	VV, SE, CBS, SHC
Approved final version	**AM**, JN	**AM**, JN	**AM,** VV, SE, CBS, SHC
Revision of manuscript	**AM,** JN	**AM**, JN	**AM,** SHC
Corresponding author	**AM**	**AM**	**AM**

Note. Bolded initials led the task

AM = Anja Møgelvang; JN = Jorun Nyléhn; SHC = Sehoya Harris Cotner; VV = Vigdis Vandvik; SE = Ståle Ellingsen, CBS = Christian Bianchi Strømme

Contents

1. Introduction

In the face of serious global challenges, international cooperation and knowledge in science, technology, engineering, and mathematics (STEM) are particularly essential (Milgrom-Elcot, 2022; Schreiner et al., 2010; Taylor, 2016; UN, 2015). In an age characterized by climate crises, pandemics, increasing inequalities, and war, the motivation for learning is not simply increased productivity or national competitiveness (Cornford, 2005). Rather, the very goal of education is developing *values, attitudes, ways of doing and ways of being* (Biesta, 2020, p. 102) in a joint cooperative effort to achieve worldwide sustainability, health, equality, and peace. The reliance on STEM knowledge to achieve these goals places a specific responsibility on STEM higher institutions to educate a skilled, knowledgeable, and effective workforce (Donaldson, 2020, p. 722; Milgrom-Elcot, 2022; Schreiner et al., 2010; Taylor, 2016).

Despite the consensus on the importance of STEM knowledge for the greater good, STEM higher education struggles to recruit and retain students - and to provide the students with the learning outcomes needed (Canelas et al., 2017; Drew et al., 2015; Honey et al., 2020; Johnson et al., 2002; Leggett et al., 2004; Lisberg & Woods, 2018; Milgrom-Elcot, 2022; Schreiner et al., 2010; Taylor, 2016; Waite & McDonald, 2019). The STEM students themselves are not in doubt what they need to succeed. Their take-home message is loud and clear: STEM teachers, by creating learning environments that promote student belonging and belief in their own abilities, make all the difference (Milgrom-Elcot, 2022; unCommission, 2022). These student perceptions are supported by higher education research *in general* which does establish a strong association between psychosocial factors such as sense of belonging and belief in one's own abilities and desired outcomes such as student identity, retention, and success (Aurlien et al., 2019; Ayllón et al., 2019; Parkes, 2014; Sæthre, 2014; Thomas, 2012; Tinto, 1975, 1993; van Dinther et al., 2011). So far, most higher education research with a *specific* focus on STEM disciplines has targeted ways to improve "hard" student outcomes such as content knowledge and placed less emphasis on the promotion of psychosocial factors – although these are considered critical to explore (Cromley et al., 2016; Lytle et al., 2021; Singer et al., 2012; Talanquer, 2014; Trujillo & Tanner, 2014). Given the challenges faced by STEM higher education, the needs conveyed by the STEM students, and previous educational research, it is time STEM discipline-based educational research prioritizes the exploration of psychosocial factors – and how educators can promote these factors.

Psychosocial factors are developed through our interaction with our surroundings, and they may influence us both positively and negatively. Factors influencing us positively are also known as protective factors and those influencing us negatively as risk factors (Thomas et al., 2020). Increasing *sense of belonging*, a protective factor, seems of particular importance to STEM students in the US (Milgrom-Elcot, 2022; unCommission, 2022) and is considered vital to overcome challenges with student retention (Aurlien et al., 2019; Parkes, 2014; Sæthre, 2014; Thomas, 2012; Tinto, 1975, 1993). Increased sense of belonging may also benefit Norwegian STEM students as reports in Norwegian higher education reveal that students experience increasing loneliness (NOKUT [The Norwegian Agency for Quality Assurance in Education], 2021; Sivertsen, 2021). *Loneliness* is characterized as a risk factor with potential serious consequences and calls for initiatives to prevent and reduce student loneliness in Norwegian higher education have been voiced (Tekna, 2022; Trædal, 2020). Increasing a sense of belonging is one of the most effective means to do just that (Arslan, 2020; Baumeister & Leary, 1995; Palikara et al., 2021). Another important protective psychosocial factor for student success emphasized by the US students is the *belief* in one's own abilities. This belief is often characterized as confidence or self-efficacy and this factor is often related to increased student success, learning, and retention in STEM disciplines (Ballen et al., 2017; Bandura, 1997; Nissen & Shemwell, 2016; Rittmayer & Beier, 2008; Sawtelle et al., 2012; Talanquer, 2014).

Student *abilities* in higher education may roughly be divided into two categories: knowledge and skills. Both should be reflected in the outcomes acquired by higher education graduates. Traditionally, the primary purpose of STEM higher education has been to promote content knowledge outcomes over skills. This purpose, however, has not succeeded in preparing the students for life and work in the 21st century (Taber, 2016; Taylor, 2016), resulting in STEM higher education learning outcome challenges. More now than ever before, STEM higher education students need to develop a broad set of skills including critical thinking, creativity, problem solving, interpersonal understanding, collaboration, and communication (Badcock et al., 2010; Cornford, 2005; Donaldson et al., 2020; Leckey & McGuigan, 1997; Taber, 2016; Taylor, 2016). These skills are often referred to as *generic skills*, life skills, transferable skills or 21st century skills, and because they are developed through our interaction with our surroundings and may influence us positively or negatively in both life and work (Badcock et al., 2010; Taber, 2016; Taylor, 2016), they, too, may be characterized as psychosocial factors. Taken together, the evidence suggests that strengthening protective factors such as *sense of belonging*, *belief in one's own abilities*, and a variety of *generic skills* while reducing a risk factor such as *loneliness* can successfully address many of the challenges faced by the STEM higher education communities and their students.

The importance of psychosocial factors in higher education became especially evident during the COVID-19 pandemic. As the pandemic forced teachers and students to move both teaching and learning from traditional classrooms to online platforms, students worldwide reported a range of challenges, e.g., lack of belonging, increased loneliness, and difficulties in studying from home (Børve et al., 2021; Camfield et al., 2021; Deng et al., 2021; Lederer et al., 2021; Phillips et al., 2022; Sivertsen, 2021; Tice et al., 2021; Werner et al., 2021). These reports all shed a bright light on the importance of *student affect* (Beard et al., 2007) to retain and motivate students to fulfill their studies - and in many ways, this recognition may turn out to be a positive ramification of the pandemic in higher education. As the higher education communities are finally realizing the importance of psychosocial factors, the urgent question remains: *how* can STEM higher education strengthen protective factors and reduce risk factors to overcome the challenges it faces?

The need for teaching and learning strategies that promote cooperation and mitigate the challenges of recruitment, retention, and learning outcome in STEM higher education is well-established. In this thesis, I examine one such teaching and learning strategy, cooperative learning (CL), which is a highly structured form of group work (Millis & Cottell, 2010). In line with the dominating focus on “hard” student outcomes, CL has predominately been used to promote academic achievement in international undergraduate STEM education (Apugliese & Lewis, 2017; Springer et al., 1999). However, it is hypothesized - and to some extent empirically supported - that CL may also promote a range of psychosocial outcomes (Johnson & Johnson, 1989; Johnson et al., 2014; Millis & Cottell, 1998). Further, CL is based on a thorough theoretical framework, social interdependence theory. Social interdependence theory was developed in the 1930s and 1940s, a time when global challenges were, arguably, as critical and pressing as they are today. Back then, as today, cooperation was believed to be the solution to the global challenges of the time, and, according to social interdependence theory, a key to true, effective, and successful cooperation is positive interdependence. Thus, CL does not only provide us with a method to possibly strengthen student cooperation and protective factors while reducing risk factors in STEM higher education. CL may also provide us with a theoretical understanding of *how* and *why* (i) cooperation succeeds or fails, (ii) protective factors are strengthened, and (iii) risk factors are reduced—and do this on a historical background similar to ours. The decision to explore CL in international and Norwegian undergraduate STEM education seemed timely, and in this thesis, I examine the following psychosocial outcomes of CL: increased sense of belonging, generic skills, and science confidence, and decreased loneliness.

In addition to this introductory chapter, this PhD thesis consists of six other chapters contextualizing and discussing the applications and outcomes of CL in undergraduate STEM education. Chapter 2 elaborates on CL and the theoretical framework guiding this thesis, social interdependence theory. Chapter 3 positions the thesis through a review of previous research and knowledge gaps before introducing the aims and research questions of the thesis. Chapter 4 entails the methodological approach of the thesis and in chapter 5, a short overview of the results in each of the three studies is provided. These results, their contributions, and their limitations are discussed in chapter 6 before chapter 7 concludes and outlines both study implications and possibilities for future research.

2. Theoretical Framework

In this chapter, I will elaborate on the chosen theoretical framework for the thesis, social interdependence theory. First, I will explain cooperative learning (CL), the principles guiding CL, and select characteristics of CL elements before giving an account of social interdependence theory. Although social interdependence theory is the chosen theoretical framework for this thesis, other prominent theoretical perspectives exist. Thus, I will give a very short introduction to two of these alternative perspectives and argue why I ultimately chose social interdependence theory as the theoretical framework for my thesis. Lastly in the chapter, I will include a brief section which clarifies my use of CL terms throughout the thesis.

2.1 Cooperative Learning (CL)

Cooperative learning (CL) may be defined as: '...a highly structured form of group work' (Millis, 2010, p. 5) and '…the instructional use of small groups so that students work together to maximize their own and each other's learning' (Johnson & Johnson, 1989, p.121). For true cooperation to occur, both groups and group tasks should be structured according to five principles.

2.1.1 CL Principles

The five main principles in CL are positive interdependence, individual accountability, promotive interaction, appropriate use of social skills, and group processing (Johnson & Johnson, 2009).

Positive interdependence is the backbone and the most essential principle in CL (Stevahn, 2021). Positive interdependence is achieved by structuring the group and the group task in ways that make group members dependent on each other and that help create a common interest in co-working to successfully complete the task (Ballantine & McCourt Larres, 2007, p. 128). In other words, this mutually dependent interest among the group members becomes an incentive for each student to become an active participant and to engage in the solution of the problem task to reach their shared goal.

Individual accountability is an important principle in CL because it promotes responsibility and prevents social loafing, also known as free-riding behavior (Millis & Cottell, 1998). Individual accountability is achieved when the teacher includes a mechanism, e.g., individual tests, for holding group members accountable for learning the material and completing the group task (Ballantine & Larres, 2007, p. 128).

Promotive interaction takes place when group members encourage and ease each other's contributions through listening, exchanging ideas, offering explanations and constructive feedback (Gillies, 2014, p. 131). According to Johnson and Johnson (1990), such reciprocal actions may also lead to group members feeling more accepted and valued.

Appropriate use of social skills is the explicit training and negotiation of social inclusion, mutual respect, consideration, and assistance within the group (Gillies, 2016) to promote skills in leadership, decision-making, trust-building, communication, and conflict-management (Johnson et al., 2014, p. 94).

Group processing/reflection occurs in two steps: First the group members reflect on which group actions and strategies were useful and which were not and second, they decide which actions and strategies should be maintained and which need altering (Johnson & Johnson, 2009).

According to Millis and Cottell (1998), the two key CL principles in higher education are positive interdependence and individual accountability. More than the other three principles, these two should underpin all aspects of CL, including group features and CL structures. Thus, these are the principles most emphasized and discussed in the continuation of this thesis.

2.1.2 CL Group Features

Group size is an important feature of CL. 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 (Millis & Cottell, 1998, p. 13). When students work in relatively small groups, social loafing might be avoided, less forthright students can express their opinion, and in groups of four, pairing up is easy, and even if a person is missing, the group is still technically a group (Johnson et al., 1998a; 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).

Teacher assignment of heterogenous groups is another important feature. 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., 1998a; 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 are typically only used in schools as they last at least one year which is not tenable in most higher education programs. These types of groups are meant to provide long-term support in order to make academic progress and build committed relationships (Johnson et al., 1998a; Johnson, 1992).

2.1.3 CL Structures

CL structures are content-free strategies (Kagan, 1989; 2021) that organize the interaction of students by prescribing student behaviour step-by-step to complete the assignment (Johnson et al., 1998a). These structures are 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 cooperative behaviour and demonstrated more complex thinking and problem-solving skills. In a systematic review of secondary and post-secondary courses, Romero (2010) found that the effect on student achievement was greater for structured than unstructured CL interventions.

The CL principles and importance of highly structured elements in all aspects of CL, have resulted in several well-known CL methods, including both CL group features and CL structures (Box 1). The list of CL methods in Box 1 is by no means exhaustive and applications of CL may include one or more of these methods or other CL methods altogether.

Box 1.

Overview of selected well-known CL methods

Academic Controversy: Groups of four are divided into pairs and each pair given a pro or con position on a controversial subject. The pairs prepare supporting arguments, presents their position, and criticize the opposing position before changing sides and repeating the steps. Ultimately, the groups agree on a position and write a report giving the supporting evidence and rationale (Johnson et al., 1994).

Group Contract: A 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 (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).

Roundrobin & Roundtable: In response to a question or a task, the group members in turn orally provide thoughts and possible answers (*Roundrobin*) or write thoughts and possible answers with one pen and one piece of paper (*Roundtable*) or multiple pens and papers (*Simultaneous Roundtable*) are passed around in the group (Kagan, 1989).

STAD: A CL strategy where small groups of students with different levels of ability are working together to accomplish a shared goal. When all group members master the task, they take individual quizzes or exams (Slavin, 1991).

Think-Pair-Share/Square: A CL technique which 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 (*Think*). Next, the students are asked to pair up and discuss their thoughts or responses to the question (*Pair*) before they share their joint answer with the entire class (*Share*) or in their groups (*Square*) (Millis & Cottell, 1998, p. 73).

2.1.4 Cooperative Learning (CL) vs. Collaborative Learning

The conscious and structured approach to every element of group work - from group features, shared goals, tasks, resources, roles, and structures, to rewards - separates cooperative learning (CL) *sensu stricto* from other forms of small-group learning, e.g., collaborative learning in the broader sense. In contrast to CL, collaborative learning is characterised by looser structures and relies mainly on very few elements, except for task and goal, to guide the collaborative process (Davidson, 2021; Millis & Cottell, 1998, p. 7-10). Thus, collaborative learning teachers never or only rarely consider and theoretically motivate design elements such as group features, role assignments, team-building activities, cooperative structures, equal participation, or activist interventions in their teaching (Davidson, 2021). According to the CL literature, the success of CL - and hence the reasons educators should consider implementing CL in their teaching – lies precisely in this conscious, structured, and theoretically founded approach. Through this approach educators will ensure that student groups do not succeed solely due to chance and increase the probability that all students will experience successful group work processes and outcomes.

2.1.5 CL Outcomes

Traditionally, most studies on the outcomes of cooperative learning (CL) in higher education have compared cooperative, competitive, and individualistic approaches and divided the outcomes into three reciprocal categories: 1) efforts to achieve (e.g., academic motivation, persistence, productivity, and performance), 2) positive relationships (e.g., social skills, promoting each other's success, and forming academic and personal relationships), and 3) good psychological health (e.g., personal ego-strength, self-confidence, and autonomy) (Johnson & Johnson, 1999). These three outcomes may according to Johnson and Johnson (1999) be explained due to several group processes arising from positive interdependence. CL promotes efforts to achieve because the group members work together to achieve a shared goal. Positive relationships develop because CL groups lead to friendship which in turn increases personal commitment to the success and growth of peers. Further, working cooperatively leads to good psychological health because committed and caring group members provide opportunities to share and solve problems, offer support, and install a feeling of self-value. The most studied student outcome in higher education falls within the first category and seems to be academic performance or achievement. 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. More recent meta-analyses and systematic reviews (Kyndt et al., 2013; Romero, 2009) support these findings. However, according to CL literature (Johnson et al., 1998a, 2014), studies examining student outcomes in the second and third category do exist and generally these seem to find that cooperative approaches compared to competitive and individualistic approaches improve both the quality of relationships and psychological health among university students in general. Previous research on the relationship between CL and student outcomes in STEM higher education specifically, is outlined in chapter 3.

2.2 Social Interdependence Theory

The approach to cooperative learning (CL) I have chosen in this thesis, i.e., the Johnson and Johnson approach, stems from social interdependence theory. Social interdependence theory states that we are socially interdependent when our individual outcomes are influenced by other people's actions and was first introduced by Morton Deutsch in 1949 (Johnson & Johnson, 2005). Both the background and premise of social interdependence comprise relevant reasons for choosing the Johnson and Johnson CL approach and social interdependence theory as the theoretical foundation for the thesis.

2.2.1 Background

Historically, the formulation of social interdependence theory took place in an era of global challenges during the 1930s and 1940s, and Deutsch himself underlined many times how much these challenges affected him and his work (Deutsch, 2012). Deutsch was born in 1920 to Jewish parents and grew up during the Great Depression, amid the rise of Nazism and other totalitarian regimes. During World War II he served in England for the United States Air Force, witnessing the destructiveness of the war (Stevahn, 2021). These historical events shaped the early intellectual work of Deutsch. Upon his return, he embarked on his doctoral dissertation, sparked by a general interest in issues of war and peace and with a particular attentiveness to whether the nations composing the newly formed United Nations Security Council would cooperate or compete. His interest in cooperative and competitive political processes gradually changed into a lifelong endeavor to understand the fundamental characteristics of the theoretical relationship between human cooperative and competitive behavior (Deutsch, 2012).

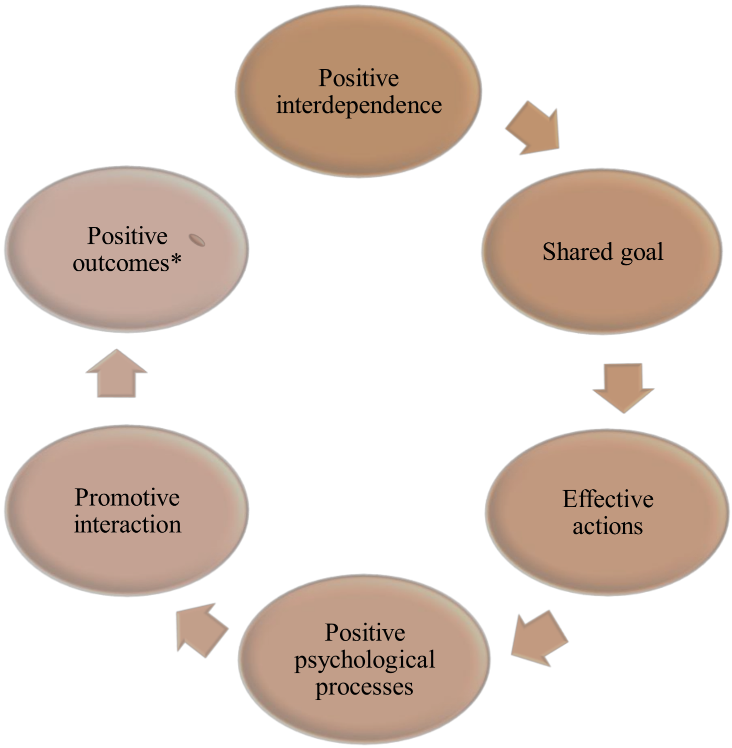
Theoretically, Morton Deutsch and social interdependence theory were heavily influenced by American Gestalt psychology (Johnson & Johnson, 2005). Gestalt means “whole” and thus, Gestalt psychology states that humans perceive their world as meaningful wholes rather than a sum of parts (Koffka, 1935). Inspired by this notion, the German American psychologist Lewin (Lewin, 1948) proposed that groups, too, are “dynamic wholes”. What Lewin meant by this is that group members are interdependent and that a change in the state or behavior of any member changes the state or behavior of any other member (Johnson & Johnson, 2002, p. 120). Deutsch, who was one of Lewin’s students, explored these group dynamics further and theorized that interdependence in groups might be both positive and negative. Positive interdependence would initiate cooperation, and negative interdependence would lead to competition (Deutsch, 2012).

2.2.2 Premise

The premise of social interdependence theory is that actions, psychological processes, interaction, and subsequently outcomes of individuals, are dependent on how social interdependence in groups is structured. There are three ways of structuring social interdependence: positive interdependence, negative interdependence, and no interdependence (Johnson & Johnson, 2002). According to Deutsch, positive interdependence arises when there is a positive correlation among the goal achievement of individuals in a group, i.e., when these individuals think that the only way, they can reach their own goals is if other individuals reach their goals (Johnson & Johnson, 2009, p. 366). Therefore, as shown in Figure 1, the individuals within the group engage in effective actions to try to reach their shared goal and in the process, they are likely to experience positive psychological processes. Deutsch (2012) points to three psychological processes and these are cathexis, substitutability, and inducibility. Cathexis concerns the human innate predisposition to respond positively to stimuli that are beneficial for us and negatively to those that are harmful. Substitutability is a term used to describe the degree to which an individual’s actions can satisfy another individual's intentions. Substitutability is key in our modern society, whether it be the functioning of social institutions, division of labor or role specialization. Inducibility refers to the readiness to accept or reject doing what another individual wants us to do. Positive interdependence is likely to affect these three psychological processes in a positive manner and because of that, the next step in Figure 1, promotive interaction, i.e., when individuals engage, support, and cheer each other on, will follow. Ultimately, the entire process will lead to positive outcomes for the individuals in the group. As previously stated, outcomes in this respect are often divided into the three reciprocal categories: 1) efforts to achieve (e.g., academic motivation, persistence, productivity, and performance), 2) positive relationships (e.g., social skills, promoting each other’s success, and forming academic and personal relationships), and 3) good psychological health (e.g., personal ego-strength, self-confidence, and autonomy) (Johnson et al., 2014).

Figure 1.

Positive interdependence processes in groups. See text for details and explanations.



Note. * Positive outcomes can include 1) efforts to achieve, 2) positive relationships and 3) good psychological health (Johnson et al., 2014).

According to Deutsch, negative interdependence arises when there is a negative correlation among the goal achievement of individuals in a group, i.e., when these individuals think that the only way, they can reach their goals is if the other individuals fail to reach their goals (Johnson & Johnson, 2009, p. 366). Therefore, the individuals within the group engage in bungling actions to try and prevent each other from reaching their goal. As a result, it is likely that individuals will experience negative psychological processes such as negative cathexis, non-substitutability, and resistance to influence. Hence, negative interdependence also results in what Deutsch calls *contrient* interaction patterns, where individuals discourage and obstruct each other (Deutsch, 2012). Ultimately, the entire process will lead to negative outcomes for the individuals in the group, whether it be efforts to achieve, relationships, or psychological health (Johnson et al., 2014).

No interdependence arises when there is no correlation among the goal achievement of individuals in a group and these individuals think that they are completely independent of other individuals when it comes to reaching their goals (Johnson & Johnson, 2009, p. 366). Hence, the incentives to work in groups disappear and so do the actions, psychological processes, interactions, and outcomes otherwise emerging because of interdependence (positive or negative) in groups (Deutsch, 2012; Johnson et al., 2014).

Deutsch (2012) stressed that few situations are characterized by purely positive, negative, or a complete lack of interdependence. In most situations, different types of interdependence occur simultaneously, unconsciously, and interchangeably. Thus, to facilitate positive student outcomes, i.e., high efforts to achieve, positive relationships, and high psychological health, it is important to create *deliberate* structures leading to positive interdependence between students in groups – hence cooperative learning (CL).

2.3 Other Theoretical Perspectives on Cooperative Learning (CL)

2.3.1 Dewey and Pragmatism

One of the theoretical perspectives most often linked to CL is Dewey and his educational philosophy known as pragmatism (Davidson, 2021). Pragmatism derives from “pragma” meaning action or practice and emphasizes human beings as active participants (Egelandsdal & Ness, 2020, p. 62). This view is clearly expressed in the famous saying “Learning by doing” associated with Dewey’s philosophy and reflects many parallels to CL. Dewey introduced many other features with parallels to CL and these include: inquiry, student involvement, reflection on experience and action, the importance of interaction, problem-solving, and the idea that learning activities should be democratic, relevant, meaningful, and valuable to the students (Davidson, 2021; Dewey, 2011; Egelandsdal & Ness, 2020; Säljö, 2013).

2.3.2 Vygotsky and Social Constructivism

Another theoretical perspective frequently associated with CL is the sociocultural perspective advanced by Vygotsky (Davidson, 2021). The sociocultural perspective underlines that human beings learn and develop through their interactions with their surrounding communities and cultures (Säljö, 2013; Woolfolk, 2001). Two of the most important concepts in Vygotsky's theory are *scaffolding* and the *zone of proximal development,* both of which illustrate how we are dependent on others to learn and develop (Säljö, 2013; Vygotskij, 1986; Woolfolk, 2001). Scaffolding constitutes different types of support for learning such as clues, breaking problems down into steps, and encouragement (Woolfolk, 2001, p. 49). The zone of proximal development is the phase or area in which the student is dependent on guidance from or cooperation with a more knowledgeable other to successfully solve a problem (Säljö, 2013). It is not difficult to see the parallels between the sociocultural perspective, including these concepts, and CL. Not only do they both emphasize the notion of social interaction in learning, but they also stress the importance of tools or, in CL terminology *structures*, to support learning.

2.4 Why Social Interdependence Theory?

First and foremost, I considered it important that the teaching and learning method applied in the thesis was grounded in theory, and specifically that the theory on which Johnson & Johnson based their approach to CL was social interdependence theory. In Johnson & Johnson's own words, they developed their method through four consecutive steps:

> (a) practical procedures should be derived from theory, (b) the theory must be validated by research, (c) operational procedures should be formulated from the validated theory, and (d) the implementation of the operational procedures will reveal shortcomings in the theory, which results in revisions in the theory, a new round of validating research studies, modified procedures, and more fruitful implementation, which illuminates new shortcomings in the theory and so forth. (Johnson & Johnson, 2021, p. 47)

Second, the background and premise of social interdependence also appealed to me. Social interdependence theory was formulated in a previous era of severe global challenges and may possibly provide valuable insight into ways of understanding and maybe even ways of solving the global challenges of our time. Just as we did then, now we need to cooperate to solve critical global challenges, and social interdependence theory provides a framework on what it may take to obtain true cooperation: the need for positive interdependence. The notion of gestalt in social interdependence theory is also interesting to this thesis as the examined outcome variables may best be understood as dynamic wholes. Loneliness, sense of belonging, and science confidence seem to be dependent on agreement between the self and the surrounding structures (Allen et al., 2021) and generic skills appear to be holistic in nature (Taber, 2016). Further, a theoretical framework and method preoccupied with outcomes belonging to categories such as "positive relationships" and "good psychological health" aligned well with the focus of the thesis.

Finally, social interdependence theory is consistent with the chosen methodology of this thesis. Deutsch himself describes his philosophical stance with these words: "I wanted my theory and research to be relevant to important social issues, but I also wanted my work to be scientifically rigorous and tough-minded" (Deutsch, 2012, p. 4). Thus, it is fair to say that the work of Deutsch is developed within the ontology, epistemology, and methodology of the post-positivist paradigm, as is this thesis (see section 4.8).

2.5 Clarifications

The application and understanding of cooperative learning (CL) in this thesis are mainly based on the approach by David W. Johnson and Roger T. Johnson (Johnson & Johnson, 1989; 2009; 2021; Johnson et al., 2014). Throughout the thesis, I have consciously chosen to divide CL elements into two main categories, CL group features and CL structures, because this division may constitute an intuitive and simplified way to acquire an overall understanding of CL and how to facilitate it. This division is also found in Cottell (2010) and in Millis and Cottell (1998) on CL in higher education. Similarly to Millis and Cottell, I also understand the concept “CL group features” as ways of composing groups and “CL structures” as general ways of providing structure in both group tasks and the group work processes (Cottell, 2010; Millis & Cottell, 1998). Thus, the latter, i.e., CL structures, in this thesis include but are not confined to the traditional understanding of structures associated with Spencer Kagan and his structural approach (Kagan, 1989; 2021). The main reason for this adoption of the concept is that “structures” align well with the definition of CL on which I have based this thesis. Lastly, I do not assign different meanings to the concepts or terms “CL elements” and “CL methods”. The main reason for using both lies in the formulation of the research questions and hypotheses guiding Papers I and III. As “CL elements” is the term used in the research questions in Paper I, I predominantly use “CL elements” when referring to or discussing Paper I. Similarly, since “CL methods” is the term used in the hypothesis stated in Paper III, I predominantly use “CL methods” when referring to or discussing Paper III.

3. Previous Research

As in higher education in general, the evidence indicates that the most studied cooperative learning (CL) outcomes in science, technology, engineering, and mathematics (STEM) higher education falls within the first category of CL outcomes, i.e., higher efforts to achieve (Apugliese & Lewis, 2017; Springer et al., 1999). In their meta-analysis Springer et al. (1999) demonstrated that small-group learning promotes greater academic achievement and increased persistence in undergraduate STEM students, and in a more recent meta-analysis (Apugliese & Lewis, 2017) the positive impact of CL on students' chemistry understanding was found to be both significant and robust. Due to the dominating focus on this first category of outcomes, and academic achievement in particular - and not least due to the current challenges faced by STEM higher education - I have primarily examined psychosocial outcomes of CL in undergraduate STEM education in this thesis. Thus, this chapter will only entail previous STEM higher education studies on the effect of CL on perceived sense of belonging, generic skills, science confidence and academic self-efficacy, and loneliness. As a conclusion, I will briefly sum up the state of the field, identify existing knowledge gaps, and comment on how this thesis contributes to the field.

3.1 Student Outcomes of CL

To stress the importance of the surroundings, e.g., the teaching and learning strategies students face, and how these may influence the students' perceived sense of belonging, generic skills, science confidence, and loneliness in both positive and negative ways, I have adopted the term "psychosocial" outcomes to describe the student outcomes examined in this thesis. In this section, the psychosocial constructs examined in this thesis are first defined and explained and second, related to previous research on CL outcomes in undergraduate STEM education.

3.1.1 Sense of Belonging

To belong may be regarded a basic human need (Baumeister & Leary, 1995; Maslow, 1968), possibly just as important for our health and survival as are basic physical needs (Baumeister & Leary, 1995). We attempt to fulfill our need to belong by engaging in meaningful interpersonal relationships and social interactions (Baumeister & Leary, 1995), and *sense of belonging* may be defined as our experience of being an integral part of our surrounding systems or environment (Hagerty et al., 1992, p. 173).

In their recent review on belonging, Allen et al. (2021) provide an explanation of how sense of belonging or its opposite, alienation, may develop. They underline that our sense of belonging develops in the interaction between the self and our surroundings but may be regarded as a subjective feeling. Thus, the surrounding structures appear to provide an orientation for the self to decide who and what is acceptable and what is right and wrong (Allen et al., 2021, p. 88). If the ongoing interaction between the self and surrounding structures reinforces the perceptions, identity, culture, and experiences of the self, a sense of belonging may be facilitated. If the opposite scenario plays out in the interaction between the self and their surroundings, and reinforcement fails to occur, a sense of alienation may result (Allen et al., 2021). Alienation among young people is on the rise (Destin, 2019; Holland, 2019; Katartzi & Hayward, 2020; Leyshon, 2011) and may lead to declining mental health, low well-being, and increased loneliness (Arslan, 2020; Palikara et al., 2021).

Considering the importance of belonging and the increase in alienation, educational institutions should implement intentional and systematic practices to become *places of belonging* (Murdock-Perriera et al., 2019). Central to such practices is that components such as perceptions of belonging and opportunities to belong reinforce and affect one another continuously in the development of belonging (Allen et al., 2021). Thus, it seems essential to acknowledge that students enter with different perceptions of belonging informed by past experiences - and to add to an existing sense of belonging and create new experiences that may remedy past experiences of alienation. A way to achieve this may be to create learning and social settings where students are given opportunities to belong on different levels, e.g., with peers, teachers, disciplines, and the institution. These settings should be structured in systematic ways that reduce competition and ensure that all students may fulfill the need for belonging through social interactions and meaningful relationships (Allen et al., 2021; Baumeister & Leary, 1995, Murdock-Perriera et al., 2019), e.g., through groups and group work.

Sense of belonging has been examined as one of many student outcomes of CL, including in undergraduate STEM education (Furuto, 2017; Wilton et al., 2019; Yapici, 2016). In a study in undergraduate mathematics, Furuto (2017) reported that implementing the CL method STAD (Box 1, 2.1.3) increased a sense of belonging among the students. Similar CL methods led to similar results in an undergraduate biology course (Yapici, 2016). In yet another undergraduate biology study, Wilton et al. (2019) introduced structured in-class student-student/student-teacher interactions and peer-led discussions and found that the students reported greater sense of belonging than did students in a similar course with traditional teaching. Taken together, these studies suggest a potential for a positive relationship between CL and sense of belonging in undergraduate STEM education.

3.1.2 Generic Skills

Generic skills, holistic skills that operate across wide ranges of contexts (Taber, 2016, p. 226), are developed through our interaction with our surroundings, and often predict success in life and work (Heckman & Kautz, 2012, p. 2). Well-developed generic skills influence us positively in both life and work because they are (i) important tools for lifelong learning (Bourn, 2018) and coping strategies (Leckey & McGuigan, 1997), (ii) highly desired by employers (Davey et al., 2018), and (iii) believed essential to navigate in an unpredictable future (Badcock et al., 2010; Taber, 2016; Taylor, 2016). Thus, generic skills may be categorized as psychosocial skills that are also known as “life skills”, “transferable skills”, and “21st century skills” (UN, 2015; UNICEF, 2021). These skills include a) ways of thinking: e.g., creativity, critical thinking, problem solving, and meta-cognition, b) ways of working: e.g., collaboration and communication, c) tools for working: e.g., information and communication literacy, and d) living in the world: e.g., citizenship and life and career (Binkley et al., 2012, pp. 18-19).

In educational settings, generic skills seem to be developed through an integration of content knowledge and active learning methods, particularly group work (Ballantine & McCourt Larres, 2007; Kember et al., 2007; Smith & Bath, 2006; Tynjälä & Gijbels, 2012; Virtanen & Tynjälä, 2019, 2022). Integrating content and active collaborative learning methods is believed to facilitate a dual process: When students use theoretical knowledge and concepts to discuss and solve practical problems, they learn how to solve practical problems in new ways and also how to conceptualize their practical experiences in relation to the theory (Virtanen & Tynjälä, 2019, p. 882), hence improving both their theoretical and practical learning.

Johnson and Johnson (2002; 2016), Millis and Cottell (1998), and Slavin (1996) have all suggested that CL may lead to improved generic skills. Multiple empirical studies have tested this proposition, based on subjective student reports (Canelas et al., 2017; Carson & Glaser, 2010; Cheruvelil et al., 2020), objective assessment (Rattanatumma & Puncreobutr, 2016; Sandi-Urena et al., 2012) or both (Díaz-Vázquez et al., 2012; Pilcher et al., 2015; Srougi et al., 2013; Winschel et al., 2015) and find that CL generally does enhance undergraduate STEM students' acquisition of generic skills. The generic skills examined in these studies were teamwork skills (4 of the studies listed above included this skill), problem-solving skills (4 studies), critical thinking skills (2 studies), communication skills (1 study), and metacognitive skills (1 study).

In undergraduate chemistry, Canelas et al. (2017) compared two similar courses, one employing traditional lectures and one employing the CL method Process Oriented Guided Inquiry Learning (POGIL), see Box 1 in section 2.1.3, and found that the latter resulted in increased acquisition of key transferable skills such as problem-solving and collaboration significantly. Other studies from undergraduate chemistry (Sandi-Urena et al., 2012; Winschel et al., 2015) found that different cooperative lab instructions related to an increase in the students' problem-solving skills. These findings are supported by a study from undergraduate mathematics (Rattanatumma & Puncreobutr, 2016), in which the students taught using the CL method Student Team Achievement Division (STAD), see Box 1 in section 2.1.3, achieved higher scores in problem-solving ability compared to the students taught through problem-based learning. Likewise, a qualitative study replacing recipe-based lab teaching by the jigsaw strategy (Box 1, 2.1.3) found a substantial increase in the students' metacognitive skills (Pilcher et al., 2015). A few studies on the effects of *online* CL on generic skills in STEM higher education exist. Using objective assessment, one of these studies showed that cooperative reciprocity, triggered by asymmetric learning contents, led to significantly increased critical thinking skills development (Lee et al., 2016). Another study, by means of both objective assessment and subjective student reports, found that a mobile learning application containing jigsaw-based cooperative and interactive learning approaches greatly enhanced information evaluation skills compared to traditional learning (Parsazadeh et al., 2018).

Of the psychosocial outcomes addressed in this thesis, student perceptions of generic skills are the most frequently examined in previous research. All the previous studies above examining the relationship between CL and students' self-reported generic skills point to a positive relationship and hence, improved perceived generic skills may be characterized as the most established of the CL outcomes included in this thesis.

3.1.3 Science Confidence and Academic Self-Efficacy

Science confidence may be understood as a student's perception of their own abilities to learn and complete tasks specific to science, such as articulating a testable hypothesis, designing an experiment, or explaining scientific concepts to peers (Cotner et al., 2020; Seymour et al., 2004). Thus, science confidence seems to be closely related, but not identical to, the more general *academic self-efficacy* (Ainscough et al., 2016; Ballen et al., 2017; Cotner et al., 2020; Nissen & Shemwell, 2016; Rittmayer & Beier, 2008). Self-efficacy is the belief in one's capacity to organize and execute the courses of action required to manage prospective situations (Bandura, 1997) and consequently *academic self-efficacy* captures students' beliefs that they are capable enough to master academic tasks and obtain successful results. Bandura (1997) theorized that an individual develops their own self-efficacy for a specific task through social and personal experiences. These experiences fall into four categories: mastery experiences, vicarious learning experiences, social persuasion experiences, and an individual's physiological and affective state. Students draw on all four of these categories as sources of information in building their self-efficacy. Generally, science confidence and self-efficacy are positively associated with desirable student outcomes such as better performance and retention in a discipline (Ballen et al., 2017; Rittmayer & Beier, 2008). The relationship between confidence and performance may be understood through the lens of social cognitive career theory (Bandura, 1986; Lent et al., 1994) - whereby a perceived reduced capacity, and possibly a consequent *lack of belonging*, learned through social and personal experiences in a discipline informs an individual's self-evaluation and sense of a future in that discipline. Science confidence has, to the best of my knowledge, never been examined as one of the typical student outcomes of CL in science higher education. However, academic self-efficacy (Bandura, 1986) has – and since science confidence, like academic self-efficacy, is believed to develop through social and personal experiences (Bandura, 1986; 1997), CL may provide the necessary conditions to increase students' science confidence.

In recent STEM higher education studies, self-efficacy and specific types of academic self-efficacy have been examined as a potential outcome of CL (Espinosa et al., 2019; Furuto, 2013; 2017; Rivera, 2013). In a study in introductory algebra (Rivera, 2013), the implementation of a range of CL strategies such as roles, think-pair-share, and jigsaw (Box 1, 2.1.3) improved the students' mathematics self-efficacy significantly. Similarly in higher education algebra, the introduction of the CL strategy STAD (Box 1, 2.1.3) led to significant improvements in the students' mathematics self-efficacy (Furuto, 2017). Further, 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 and group composition in an introductory physics class based on CL principles.

Although academic self-efficacy seems to be a prominent psychosocial outcome of CL and academic self-efficacy and science confidence are closely related, we cannot assume that science confidence is equally affected by CL. In this thesis, I set out to examine if a relationship between CL and science confidence may exist and do this for the first time, not only in general but also in a digital setting.

3.1.4 Loneliness

Loneliness is a subjective feeling of distress due to deficiencies in an individual's social relationships (Peplau & Perlman, 1982, p. 3). Moving away from the safety of home and suddenly having to manage on their own, university students may be particularly vulnerable to loneliness (Cutrona, 1982; Stewart-Brown et al., 2000). Student loneliness is on the rise (Knapstad et al., 2018) and increased substantially during the COVID-19 pandemic (Phillips et al., 2022; Sivertsen, 2021; Werner et al., 2021). Furthermore, loneliness among university students may cause health problems (Hayley et al., 2017; Richardson et al., 2017) and loneliness in general is perceived as a risk factor to increased mortality (Holt-Lunstad et al., 2015). To prevent and counteract loneliness, universities may facilitate initiatives (both within and beyond the classroom, physically and digitally) to decrease loneliness (Adriansen & Madsen, 2012), and increasing opportunities for social interaction is particularly promising (Hawkley & Cacioppo, 2010).

CL is an example of an initiative offering increased opportunities for social interactions. To the best of my knowledge only one study in university populations has examined the relationship between CL and loneliness (Kocak, 2008) and this study concluded that CL heterogenous groups and CL structures led to a significant decrease in loneliness compared to traditional teaching (Kocak, 2008). In this thesis, I seek to validate these findings, albeit among undergraduate biology students in Norway and in a digital setting.

3.2 Research Summary and Research Gaps

Although academic achievement is the most established student outcome of CL in undergraduate STEM education (Apugliese & Lewis, 2017; Springer et al., 1999), the studies reviewed in this chapter indicate that CL may increase protective psychosocial outcomes such as sense of belonging, generic skills, and scientific confidence/academic self-efficacy, and decrease risk factors such as loneliness. However, only few of these studies theoretically or empirically examined what it is about CL that may explain this impact. Further, the relationships between CL and science confidence and CL and loneliness have not been established in previous research and certainly not in an undergraduate STEM setting. Finally, the existing studies cover a limited range of sociocultural settings and none of these, except for a few studies on the relationship between CL and generic skills, have been performed in a digital setting. Thus, in this thesis I attempt to explore how and why CL may affect psychosocial outcomes among undergraduate STEM students, and I do this through the lenses of a thorough theoretical framework, i.e., social interdependence theory, and through empirically identifying and measuring components (principles) in CL attributing to its success. I also examine the effect of CL on two psychosocial outcome variables not previously researched in undergraduate STEM education, i.e., science confidence and loneliness. Lastly, my thesis adds to the existing knowledge base by examining settings not previously researched, i.e., undergraduate STEM students in Norway and a digital teaching and learning platform.

3.3 Aims and Research Questions of the Thesis

Based on the outlined background rationale, theoretical framework, and previous research, the aims of this PhD thesis were to assess and develop the evidence base of CL as a means to promote student cooperation and desirable psychosocial outcomes among students in international and Norwegian undergraduate STEM education to: a) address current STEM challenges, b) inform teaching practices, c) fill current knowledge gaps in the field, and d) guide future research. Thus, the main research question guiding the thesis was:

How is cooperative learning applied and how is it related to psychosocial outcomes among students in undergraduate science, technology, engineering, and mathematics education?

This main research question was explored through several sub-research questions and hypotheses, and these were:

In Paper I:

1. *Which disciplines, countries, and research methods are prevalent in studies of cooperative learning elements in undergraduate mathematics and science education?*
2. *What are the characteristics of the cooperative learning elements used in, and principles guiding, undergraduate mathematics and science education?*
3. *What are the student outcomes of cooperative learning elements in undergraduate mathematics and science education?*
4. *How are the various cooperative learning elements associated with student outcomes?*

In Paper II:

How is cooperative learning related to sense of belonging and generic skills among students in Norwegian undergraduate STEM education?

In Paper III:

Digital cooperative learning methods lead to beneficial changes in biology students' self-reported psychosocial outcomes, i.e., sense of belonging, science confidence, generic skills, and loneliness, compared to traditional digital lectures.

To achieve these aims and answer these research questions, first we assessed the evidence base of and identified knowledge gaps in the characteristics, applications, and outcomes of CL in international undergraduate mathematics and science education in a scoping review (Paper I). Second, we examined the relationships between CL and two psychosocial outcomes, i.e., sense of belonging and generic skills, in a range of Norwegian undergraduate STEM disciplines in a cross-sectional survey study (Paper II). Third, we implemented and compared the effect of digital CL and digital lectures on a range of psychosocial outcomes, i.e., sense of belonging, science confidence, generic skills, and loneliness among Norwegian undergraduate biology students in a quasi-experimental study (Paper III).

4. Methods

The purpose of this thesis was to examine the applications and outcomes of cooperative learning (CL) in undergraduate science, technology, engineering, and mathematics (STEM) education. Thus, this chapter entails the methodological approach adapted to achieve this purpose, including study designs, samples and procedures, measures, analytical strategies, validity, reliability, research ethics, and philosophical stance.

4.1 Study Designs

Three different study designs were employed in the thesis. To obtain an overview of the applications and outcomes of CL in international undergraduate mathematics and science education and to identify possible knowledge gaps, we decided to conduct a scoping review (Arksey & O'Malley, 2005) in Study I. Based on the findings and knowledge gaps of this review, in Study II, we employed a cross-sectional design using surveys to examine and theoretically explain possible relationships between CL and select psychosocial outcomes, i.e., sense of belonging and generic skills, in a Norwegian undergraduate STEM setting. Based on the findings of both Study I and II, we adopted a quasi-experimental design (Shadish et al., 2002) in Study III. This design allowed a digital CL intervention in a Norwegian undergraduate biology course and made use of many design controls, e.g., a double pre-test and follow-up, to determine the hypothesized effect of carefully selected CL methods on a range of psychosocial outcomes, i.e., sense of belonging, science confidence, generic skills, and loneliness.

4.2 Samples and Procedures

The samples of the three studies in this thesis consisted of students in international and national undergraduate STEM education. In Paper I, the sample consisted of 24 empirical studies on the applications and effects of CL on students in international undergraduate mathematics and science education. This sample was identified through systematic search and screening strategies guided by research questions, inclusion and exclusion criteria, and the PRISMA protocol (Moher et al., 2009). In Papers II and III, we were interested in examining CL in undergraduate STEM (Paper II) and biology (Paper III) students in Norway, but due to restraints in the use of CL, time restraints within the PhD frame, data collection challenges during the pandemic, and course enrollment in general, it was not possible to target a nationally representative sample. Thus, the participants in both papers were drawn from a convenience sample, i.e., from sample units that were readily accessible to me as a researcher (Crano et al., 2015). In Paper II the sample consisted of students (n = 401, response rate = 92%) from different STEM disciplines at one of the major Norwegian universities. Only undergraduate STEM courses which had a minimum enrollment of 30 students, and which had implemented some sort of CL elements were invited to participate. The data was collected by means of surveys, some digital using SurveyXact (Rambøll, 2021) and some on paper, during the fall 2020. In Paper III, the participants (n = 71, response rate = 83%) were recruited from a mandatory undergraduate biology course at the same university studied in Paper II. The digital CL methods implemented were based on the findings on applications of CL in undergraduate mathematics and science education identified in Paper I and the data was collected by means of digital surveys in SurveyXact (Rambøll, 2021) at four different time points during the spring 2021.

4.3 Measures

Both Studies II and III have employed scales to measure a range of constructs. The scales measuring sense of belonging and generic skills have been applied in both studies whereas the scale measuring cooperative learning (CL) was only used in Study II. The scales measuring science confidence and loneliness were only used in Study III. For more details on scales, see Table 1 in section 4.4 and for details on items, internal consistency, validation etc., see Papers II and III.

All the included scales have been validated in previous international studies (Alkan, 2016; Atxurra et al., 2015; Byrne & Flood, 2003; Freeman et al., 2007; Goodenow, 1993; Hughes et al., 2004; Jansen et al., 2013; Matthews-Ewald & Zullig, 2013; Ramsden & Entwistle, 1981; Walker et al., 2008) and the scales measuring generic skills, science confidence, and loneliness had also been previously validated in Norwegian studies (Cotner et al., 2020; Espeland & Indrehus, 2003; Knapstad et al., 2018). However, as the scales measuring sense of belonging and CL had never been used in a Norwegian setting, several steps were taken to validate these. First the items in the scales were translated and back-translated by two different sets of researchers using the International Test Commission (ITC) test translation and adaptation guidelines (Hambleton, 2001). Second, the translation of each item was discussed with a group of five Norwegian undergraduate STEM students to ensure that the students' understanding of the items reflected the meaning of the items. Third, the translation and number of items in the scale were subject to change after student feedback and the statistical findings of an extensive pilot study (n=253). For further details on the validation processes of the scales measuring sense of belonging and CL, see Paper II.

4.4 Analytical Strategies

In Paper I, the aim was to obtain knowledge on the scope, applications, and outcomes of CL in international undergraduate science and mathematics discipline education. Thus, the included studies in Paper I were identified, screened, and analyzed using the five steps of the scoping review: 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). Furthermore, the screening process followed the PRISMA protocol (Moher et al., 2009) and was managed using the review tool, Rayaan (Rayaan, 2022). For further details, see Paper I and Paper I, Supplemental material 1+2.

The analyses in Papers II and III were conducted using the statistical program IBM SPSS 25 (IBM, 2017) and 27 (IBM, 2021) and included reliability analyses, factor analyses, Pearson's correlations, multiple regression, independent samples T-test, and one-way ANOVA analyses. In both studies, normality was assessed by reviewing standard deviations, skewness, and kurtosis. As only participants who had answered all questions on all measurement points were included in the data set in Paper III, no strategy for missing data was needed. However, in Paper II, some missing data existed. Generally, the missing rate was very low, i.e., under 5% which characterizes as inconsequential (Schafer, 1999). The missing rate of the CL variable was a little higher, 8%. Still, these values should not pose any particular problems as data is not likely to be biased if less than 10% is missing (Bennett, 2001). Within educational science, missing rates of 15 to 20% are not uncommon, indicating that the missing rates in Study II may be considered acceptable (Enders, 2010). Still, we decided to use an "Exclude cases pairwise" strategy to deal with missing data. This strategy only excludes the participants from specific analyses that are in need of all data (Pallant, 2016).

In Paper II, we were mainly interested in exploring relationships between CL and sense of belonging and generic skills in a Norwegian undergraduate STEM setting. Thus, we first conducted Pearson's correlations to explore and assess the relationships between CL, sense of belonging, and generic skills. Next, we conducted standard multiple regression analyses to determine how much unique variance each of the CL subscales, i.e., CL principles, would explain in the prediction of sense of belonging and generic skills. As multiple regression analyses make several assumptions about the data (Tabachnick & Fidell, 2014), preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity, and homoscedasticity. For more details, see Paper II and Paper II, Supplementary materials.

In Paper III, we examined the effect of implemented digital CL methods compared to digital lectures on sense of belonging, science confidence, generic skills, and loneliness in undergraduate biology students in Norway. Thus, our primary confirmatory analysis was One-way repeated measures ANOVA. We used Wilks' Lambda to assess significance and adjusted for multiple comparisons through Bonferroni correction (Field, 2018). Effect sizes were measured in Partial Eta squared, which may be considered small (η_p^2 >.01 to .05), moderate (η_p^2 >.06 to .13), or strong (η_p^2 >.14) in magnitude (Cohen, 2013). We also ran exploratory analyses, i.e., Independent-samples t-tests and One-way between-groups ANOVA with Tukey post-hoc tests to detect whether the outcomes varied by student gender and generation in college. Important assumptions such as Levine's test for homogeneity of variances were also reviewed.

An overview including study characteristics, data, measures, and analyses used in the three studies in the thesis is presented in Table 1.

Table 1.

Overview of the methods used in the three papers in the thesis

	Paper I	Paper II	Paper III
Study characteristics			
Program	International undergraduate mathematics and science education	Norwegian undergraduate STEM education	Norwegian undergraduate biology education
Sample size	24 empirical studies	401 students	71 students
Design	Scoping review	Cross-sectional convenience sample	Quasi-experiment convenience sample
Geography	Global	Norwegian university	Norwegian university
Data	Database searches Keywords Inclusion/exclusion criteria	Survey	Survey
Measures			
CLAS		x	
PSSM		x	x
CEQ (Generic skills subscale)		x	x
TILS			x
Science confidence			x
Analyses	PRISMA Charting, mapping, and organizing data	Pearson's correlations Standard Multiple Regression	One-Way repeated measures ANOVA Independent-samples t-tests One-way between-groups ANOVA

Note. CLAS = Cooperative Learning Application Scale; PSSM = Psychological Sense of School Membership; CEQ = Course Experience Questionnaire; TILS = Three-Item Loneliness Scale; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses

4.5 Validity

According to Shadish et al. (2002, p. 34) validity refers to: “the approximate truth of an inference. When we say something is valid, we make a judgment about the extent to which relevant evidence supports that inference as being true or correct.” An important notion in this understanding of validity is the use of the word “approximate” which aligns with the postpositivist paradigm (see section 4.8). As expressed in this understanding, to be valid our knowledge claims must be based on evidence that demonstrate the truth of the claim. Relevant to the studies in this thesis are different types of validity and these are identified and explained in each paper below. See the individual papers for a further discussion of validity in each case.

In Paper I, the scoping review, several steps to ensure validity in general were taken. First, we paid careful attention to the search strategy. To identify relevant studies in your searches, 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). Second, the screening process was guided by clearcut inclusion and exclusion criteria and followed a very systematic approach guided by the PRISMA protocol (Moher et al., 2009). Both titles and abstracts and full-text articles were screened independently, systematized, and documented in the review tool Rayaan (Rayaan, 2022). Systematic approaches such as these are crucial to the validity of reviews as they ensure relevance and create transparency while preventing both selection bias and publication bias (Booth et al., 2016). However, as a scoping review does not contain a critical appraisal, it cannot make any assessment of the validity of the included studies themselves.

In Paper II, a cross-sectional study, validity was mainly ensured through validation of the instruments, also known as construct validity, and the relevance of the instrument to users, known as face validity. Construct validity is obtained if an instrument succeeds in measuring the underlying theoretical construct it is supposed to measure (Kleven et al., 2011) and face validity refers to the relevance of an instrument as it appears to the respondents (Muijs, 2013). In Paper II, construct validity was ensured by using previously validated scales. Further, the translation of the two scales which had not been previously applied in a Norwegian setting underwent many validation procedures (see Paper II) to establish both construct and face validity.

In Paper III, as in Paper II, construct validity and face validity were established using previously validated measures and procedures. In addition, since this was a quasi-experimental intervention study, internal validity also came into play. Internal validity is the degree to which you can infer cause-effect relationships in a study (Shadish et al., 2002). To be able to infer cause-effect relationships in Study III, a range of design controls were adopted. The two most important of these design controls were the use of an extra pretest and follow-up measurements which reduce the risk of errors due to student maturation and regression (Shadish et al., 2002).

4.6 Reliability

Reliability is a term used to indicate if you can trust or rely on research to be accurate and consistent. There are many ways to test if you can rely on research and relevant in this thesis are reproducibility, inter-rater consistency, and internal consistency. Reproducibility means that other researchers when repeating the study obtain the same or similar results (Resnik & Shamoo, 2017). Inter-rater consistency is the degree to which two independent researchers are consistent in their judgements while internal consistency is a reliability measurement determining if the components of an instrument are interrelated (Crano et al., 2015).

In Paper I, reliability was primarily linked to reproducibility. There are several ways researchers can increase reproducibility in literary reviews and these include (i) complete reporting of all the search strategies in each database, (ii) assistance from university librarians in the development and execution of the search strategies, and (iii) reporting guidelines (e.g. PRISMA) (Brunton et al., 2017; Koffel & Rethlefsen, 2016). With reproducibility in mind, Paper I included all these steps. Furthermore, screening based on clear inclusion- and exclusion criteria ensured consistency in decision-making (Arksey & O’Malley, 2005) and the fact that the entire screening process was conducted by two independent researchers enabled inter-rater consistency, too (Crano et al., 2015).

In Papers II and III, reliability was assessed through measures of internal consistency, specifically Cronbach's alpha (α). Guidelines specify that alpha coefficients of .70 are acceptable, values of .80 are very good, while values of .90 are excellent (Kline, 2016). In Paper II, alpha coefficients ranged from .77 to .93 and in Paper III from .76 to .97, thus displaying overall good internal consistency. Furthermore, as measures, data, and thorough descriptions of the implemented CL elements were made available, both studies should be reproducible.

4.7 Ethics

Research ethics gives us a systematic approach to moral and normative questions (Befring, 2007). These questions deal with proper scientific practice, the protection of individuals and society, and publishing (Ringdal, 2009).

Proper scientific practice as one of the key ethical principles guided all three studies. In Study I, proper scientific practice was primarily linked to relevance, transparency, reproducibility, and avoidance of bias through elaborate search strategies (Brunton et al., 2017; Koffel & Rethlefsen, 2016), systematic screening guided by PRISMA (Moher et al., 2009), and clear inclusion and exclusion criteria. Measures to meet proper scientific practice in Study II and III included validated scales, validation through student interviews and a pilot study, design controls, and available data.

There are several ways to protect individuals, and one is to conduct an anonymous study. As the goal of Study II could be obtained without collecting and processing personal data, the participants were spared of unnecessary identification. Rather, as proper scientific practice requires, the students were informed of the purpose of the study and that participation was voluntary. As anonymous studies are not evaluated by the Norwegian Centre for Research Data (Norwegian Centre for Research Data

(NSD), 2021), Study II was registered directly in UiB's own system for risk and compliance, RETTE (UiB, 2020).

As Study III was longitudinal and it was necessary to be able to trace and follow the development of each participant, personal data in the form of student ID number was collected and processed. Thus, the data collection procedures followed the regulation of the General Data Protection Regulation (European Commision, 2022) and was approved by the Norwegian Centre for Research Data (Norwegian Centre for Research Data (NSD), 2021). To protect the individuals, each student ID number was replaced with another, not traceable, ID number. Overviews matching student ID numbers and the other ID numbers were stored on a password protected UiB device, separated from the questionnaire data. Throughout all steps in the research process, I was the only person with access to the personal data. Due to the psychological variable "loneliness", the Regional Committee for Medical and Health Research Ethics (Regional Committees for Medical and Health Research Ethics (REK), 2020) was also consulted to ensure that the data collection adhered to the health research ethics regulations. The participants were informed of the purpose of the study, that their participation was voluntary and that they could withdraw from the study at any time. Furthermore, they were informed that any personal data would be de-identified, treated confidentially, and deleted after the completion of the study. A further ethical consideration in relation to the protection of individuals was to ensure that the intervention would not harm or impair students unnecessarily, i.e., beyond the scope or the purpose of the intervention. Thus, in Study III the participating students were given equal conditions and the intervention was relatively short.

To accommodate the last of the important questions of research ethics, publishing your results, all three papers have been submitted to scientific peer-reviewed journals. Two of these, Paper I and Paper III, have been published and Paper II has been accepted for publication.

4.8 Philosophical Stance

Proper research practice requires all the different aspects in the research process to be coherent and aligned (Hatch, 2002). Examples of such aspects may be the research questions, hypotheses, methods, data analyses, and findings. In this thesis, the main research question entailed the application and outcomes of CL. To be able to answer this main research question, I identified and examined testable sub-questions and hypotheses mainly through quantitative methods and presented the findings in tables and diagrams. Inherent in this approach lies a notion of science in the traditional sense.

Science in the traditional sense is also known as the classic paradigm of science. A paradigm prescribes a certain nature of reality (ontology), knowledge of what can be known (epistemology), and knowledge of how knowledge is gained (methodology). This thesis is placed within the post-positivist scientific paradigm as described by Kuhn (Hatch, 2002). In the post-positivist paradigm, the overarching ontology is that reality exists but is never fully apprehended, only approximated. Therefore, the closest we can come to understanding reality is through generalizations, patterns, and descriptions. Suited methodologies to exploring these ontological and epistemological beliefs are scientific quantitative methods such as experiments, quasi-experiments, and surveys (Hatch, 2002).

Post-positivist beliefs and methodologies are easily transferable to this thesis. Using quantitative methods such as systematic searches, surveys, and quasi-experiments that are as controlled as possible, I have attempted to identify and infer causal relationships between CL and selected student outcomes. The research questions and hypotheses, the design of the surveys and the quasi-experiment, and the interpretation of the data are guided by and presupposed by theory. This process is very similar to the hypothetico-deductive method in which research findings will either support or weaken the hypothesis set forth initially - and therefore also an important step in identifying the “fittest” theories (Chalmers, 1999), i.e., the theories that will come as close to reality as possible. The overarching theory I have chosen in understanding the “reality” between CL and selected student outcomes in this thesis, is social interdependence theory.

Coherence and alignment throughout the research process do not only secure proper scientific practice. It also enables the individual researcher to navigate within their own and others’ research. When we as researchers choose a path, we must inevitably leave others and every path has its possibilities and its limitations (Ringdal, 2007). I believe that this thesis, by aligning with the post-positivist tradition, offers many possibilities. Thus, this thesis should constitute a piece in the puzzle and guide future research. Also, it should enable STEM higher education environments to make research-based assessments of CL as a teaching method. Such assessments should include if CL may be considered a suitable method in terms of ensuring recruitment, retention, and desired student learning outcomes in their respective fields. On the other hand, the overall paradigmatic approach of the project will also have its limitations. When one uses the hypothetico-deductive method, one will only capture the answer to the hypotheses set forth initially, and thereby potentially neglect alternative interpretations, contextualization, or in-depth human understandings of the phenomena in question. However, these limitations may be subject to future research.

5. Results

A fundamental premise for this thesis was to identify specific knowledge gaps (Paper I) and fill these gaps in different ways (Papers II and III), ensuring a systematic knowledge growth and a solid consistency between all three studies. Thus, this chapter presents the results from each of the three studies in the thesis underlining which knowledge gaps have been identified and how they have been filled.

5.1 Paper I

The purpose of Paper I was to to assess the evidence base of cooperative learning (CL) in undergraduate mathematics and science education to inform teaching practices and to identify potential knowledge gaps to inform future research, in general and specific to the thesis. Thus, four review questions were posed:

1. *Which disciplines, countries, and research methods are prevalent in studies of cooperative learning elements in undergraduate mathematics and science education?*
2. *What are the characteristics of the cooperative learning elements used in, and principles guiding, undergraduate mathematics and science education?*
3. *What are the student outcomes of cooperative learning elements in undergraduate mathematics and science education?*
4. *How are the various cooperative learning elements associated with student outcomes?*

Following the five steps systematic approach of the scoping review (Arksey & O'Malley, 2005), we identified 24 studies. Our findings showed that studies examining CL elements in undergraduate education in mathematics and science were relatively few, primarily quantitative in nature, almost non-existent outside the North American continent, and mainly conducted in chemistry. In terms of CL elements, we discovered that a vast majority of the reviewed studies met the recommendations of group size - but not the recommendations of teacher-selected and heterogeneous group composition from previous CL literature and research (Johnson & Johnson, 1999; Millis, 2010; Millis & Cottell, 1998). Despite of this gap, we found no evidence that teacher-selected and/or heterogenous groups compared to student-selected and/or homogenous groups, resulted in more outcome success. Most of the groups in the reviewed studies were formal CL groups and when we compared these to informal CL groups (see section 2.1.2), we found that the long-lasting groups seemed to be associated with more positive results in outcome, particularly in terms of academic success i.e., content knowledge and/or academic achievement and psychological health than the short-lasting groups. Further, we found that relatively few studies included a thorough theoretical framework and reported on both of the guiding CL principles: positive interdependence and individual accountability. The most applied CL structures in the reviewed studies were roles (n=6), Process Oriented Guided Inquiry Learning (POGIL) (n=6) (Box 1, 2.1.3) or both (n=4). With regard to outcome, the majority (n=21) the 24 reviewed CL studies examined academic success, i.e, content knowledge and/or academic achievement, followed by attitudes (n=10), generic skills (n=7), and different types of psychological outcomes (n=4) . We found that most of the included studies reported only positive results of the implemented CL elements (n=19) while a few studies identified both positive results and negative results (n=3) or no positive results at all (n=2). Negative results were mainly found in studies relating CL to academic achievement success and attitudes and were attributed to lack of purpose in the use of roles, lack of required elements in the POGIL method, and deficient control of confounding variables. Reviewed studies examining psychosocial outcomes relevant to this thesis, i.e., generic skills, sense of belonging and academic self-efficacy (related to science confidence), were fewer, but we discovered that almost all of these found improvement following the implementation of one or more CL elements.

In conclusion, we identified a range of knowledge gaps. First, CL studies in undergraduate STEM education are rare outside the US and not explored in many STEM disciplines, except for chemistry. Also, the application of CL elements in undergraduate STEM education is only partly in alignment with CL theory and principles - and is often lacking a thorough theoretical framework. Lastly, there are few studies examining psychosocial outcomes compared to outcomes measuring academic success, i.e., content knowledge and/or academic achievement.

5.2 Paper II

To fill some of the knowledge gaps identified in Paper I, the purpose of Paper II was to examine how self-reported cooperative learning (CL) relates to the development of sense of belonging and perceived generic skills among a sample of Norwegian students in a range of undergraduate science, technology, engineering, and mathematics (STEM) disciplines - and explain these relationships in light of a thorough theoretical framework, i.e., social interdependence theory. Thus, the research question guiding Study II was: *How is cooperative learning related to sense of belonging and generic skills among Norwegian students in undergraduate STEM education?*

Our findings in Paper II showed that students' perceived sense of belonging and generic skills were increased with self-reported CL, and both correlations were strong (Table 2). Students who experienced high levels of CL in their STEM courses also experienced high levels of belonging and perceived generic skills such as problem-solving, communication, and teamwork skills. We also found that sense of belonging and generic skills were positively and strongly correlated. Undergraduate STEM students in Norway who experienced high levels of belonging also reported high levels of generic skills and vice versa (Table 2). Further, we discovered that the CL principle "promotive interaction" was the CL principle contributing most significantly to the association with both the students' self-reported sense of belonging and generic skills (Table 3). Promotive interaction arises from positive interdependence (Figure 1, 2.2.2) and is characterized by members encouraging and easing each other's contributions through listening, exchanging ideas, offering explanations, and constructive feedback (Gillies, 2014, p. 131). The remaining CL principles such as "group work reflection", "positive interdependence", and "tutoring" did not predict sense of belonging at all. The CL principles "group work reflection" and "tutoring" significantly predicted generic skills, but not nearly as strongly as "promotive interaction". The CL principle "positive interdependence" did not predict generic skills (Table 3). In other words, when facilitating group work based on CL principles, it seems particularly important to include promotive interaction, especially if the goal is to improve the STEM students' experienced sense of belonging and generic skills.

The findings were primarily discussed in light of social interdependence theory. Specifically, we argued that shared goals and the actions, processes, and promotive interaction brought about by positive interdependence (Figure 1, 2.2.2) may facilitate positive student perceptions of belonging and strengthen generic skills among undergraduate STEM students.

Table 2.

Pearson correlation matrix of the study variables

		1	2	3	4	5	6	7	8	9	10
1	Cooperative learning (CL)	-									
2	Positive interdependence	.73	-								
3	Interaction	.83	.57	-							
4	Group work reflection	.80	.34	.53	-						
5	Tutoring	.89	.51	.64	.68	-					
6	Sense of belonging (SoB)	**.56**	.41	**.57**	.39	.47	-				
7	General sense of belonging	.47	.27	.49	.38	.37	.83	-			
8	Social sense of belonging	.41	.35	.43	.22	.32	.83	.51	-		
9	Academic sense of belonging	.52	.38	.47	.37	.47	.80	.49	.54	-	
10	Generic skills (GS)	**.52**	.33	**.46**	.42	.45	**.56**	.55	.35	.46	-

Note. All correlations were significant ($p<.01$).

Table 3.

Standard multiple regression results of the subscales of CLAS in predicting sense of belonging and generic skills

Variable	Unstandardized coefficients *B*	Unstandardized coefficients *SE*	Standardized coefficients *β*	*t*	95% CI *LL*	95% CI *UL*
Sense of belonging						
Constant	2.43	.13		19.32	2.19	2.68
Group work reflection	.01	.01	.07	1.25	-.01	.03
Positive interdependence	.02	.01	.09	1.76	-.00	.03
Interaction	.07	.01	**.40*****	6.68	.05	.09
Tutoring	.01	.01	.11	1.63	-.00	.03
Generic Skills						
Constant	11.25	1.02		11.09	9.26	13.25
Group work reflection	.20	.07	**.18****	2.91	.06	.33
Positive interdependence	.07	.07	.06	.98	-.07	.20
Interaction	.30	.08	**.24*****	3.63	.14	.46
Tutoring	.14	.07	**.14***	2.00	.00	.28

Note. CLAS = Cooperative Learning Application Scale; CI = Confidence interval; *LL* = lower limit; *UL* = upper limit.
$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$

5.3 Paper III

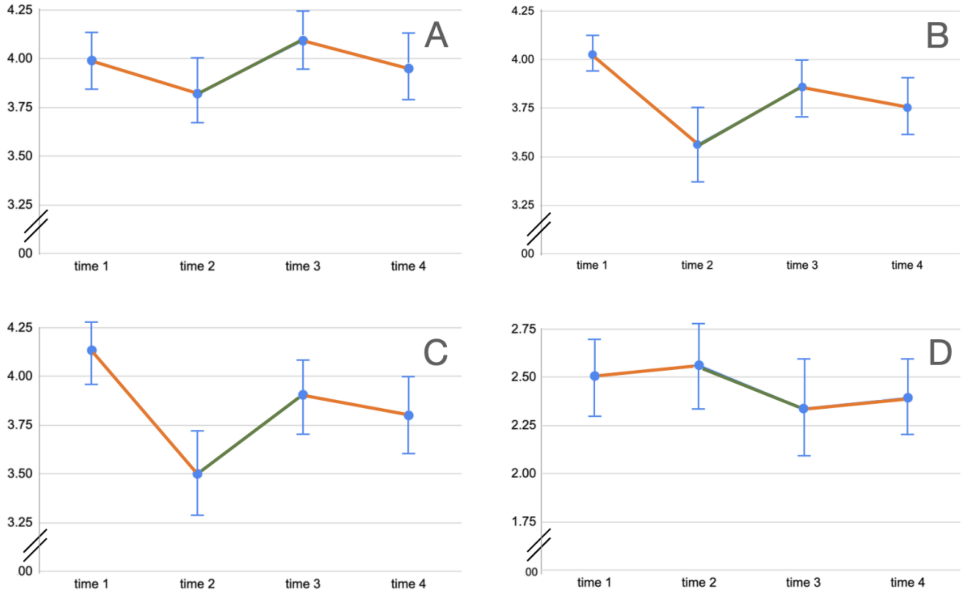
To fill yet other knowledge gaps identified in Paper I and build on the findings of Paper II, in Paper III, we implemented a selection of cooperative learning (CL) methods and explored if such methods were associated with changes in the students' sense of belonging, science confidence, generic skills, and loneliness. Due to the COVID-19 pandemic, we gained a unique possibility to conduct the study digitally, and although not targeted and therefore not initially identified in Paper I, this design element would fill yet another knowledge gap, i.e., digital CL (Davidson, 2021). To be able to compare Paper III to previous research on CL in undergraduate STEM education (Paper I), we compared digital CL to digital lectures and posed the following hypothesis: *Digital cooperative learning methods lead to beneficial changes in biology students' self-reported psychosocial outcomes compared to traditional digital lectures.* Specifically, we predicted that a digital cooperative learning intervention would increase the students' self-reported sense of belonging, science confidence, and generic skills, and lead to a decrease in loneliness.

Our results showed that coming into the course, the students scored relatively high (around 4.0 on the Likert scale) on expected sense of belonging, scientific confidence, and self-reported generic skills, and relatively low on loneliness (2.5) (Figure 2, time point 1). After five weeks of digital lectures, student's self-reported sense of belonging (-.17 [95% CI, -.30 to -.04] $p< .05$), science confidence (-.46 [95% CI, -.67 to -.25] $p< .05$), and especially generic skills (-.63 [95% CI, -.92 to -.35] $p< .05$) had decreased significantly (values given are mean [CI]) (Figure 2, time point 2). Loneliness did not change. After the next five weeks, during which students were engaged in the CL module consisting of cooperation as a learning objective, small, heterogenous, and formal CL groups, group contract, and jigsaw structures, these trends were reversed. Now sense of belonging (.27 [95% CI, .14 to .41] $p< .05$), science confidence (.30 [95% CI, .08 to .51] $p< .05$), and particularly generic skills (.41 [95% CI, .14 to .67] $p< .05$), were significantly increased, and loneliness (-.23 [95% CI, -.42 to -.03] $p< .05$) significantly decreased (Figure 2, time point 3). After the last five weeks of another series of digital lectures only sense of belonging (-.14 [95% CI, -.26 to -.02] $p< .05$) was significantly decreased. Science confidence and generic skills did not change significantly and neither did loneliness (Figure 2, time point 4). Taken together, all outcome variables changed over time, with large effect sizes for sense of belonging (η_p^2 = .34), science confidence (η_p^2 = .34), and generic skills (η_p^2 = .44), and a more moderate effect for loneliness (η_p^2 = .13) (Figure 2 and Paper III, Appendix A).

Except for one instance, i.e., at time point 1, where females (M=4.34) reported significantly higher perceived generic skills than males (M=3.88), we found no statistically significant differences in mean scores when running exploratory analyses, based on either gender or generation in college, on any of the measured outcomes (Paper III, Appendix B).

Figure 2

Close-up four-panel figure illustrating the change in students' expressed sense of belonging (A), science confidence (B), generic skills (C), and loneliness (D) including confidence intervals measured at four time points using Likert scales from 1 (strongly disagree/never) to 5 (strongly agree/very often)



6. Discussion

The knowledge and skills of science, technology, engineering, and mathematics (STEM) students in higher education, together with international cooperation, are important for addressing global challenges. However, international STEM higher education communities are struggling to live up to this task, facing pervasive challenges to insufficient student recruitment, low retention, and poor learning outcomes. Previous research and the STEM higher students themselves emphasize the importance of psychosocial outcomes such as sense of belonging and science confidence to overcome these challenges (Aurlien et al., 2019; Ballen et al., 2017; Bandura, 1997; Milgrom-Elcot, 2022; Nissen & Shemwell, 2016; Parkes, 2014; Rittmayer & Beier, 2008; Sawtelle et al., 2012; Sæthre, 2014; Talanquer, 2014; Thomas, 2012; Tinto, 1975, 1993; unCommission, 2022). Cooperative learning (CL) is a teaching and learning strategy focused on the art of true cooperation and evidence suggests that CL leads to a range of important student psychosocial outcomes.

The aims of this PhD thesis were to assess and develop the evidence base of CL to promote student cooperation and much needed student psychosocial outcomes in international and Norwegian undergraduate STEM education to a) address existing STEM challenges, b) inform teaching practices, c) fill current knowledge gaps in the field, and d) guide future research. Thus, the discussion in this chapter focuses on the main research question: *How is cooperative learning applied and how is it related to psychosocial outcomes among students in undergraduate science, technology, engineering, and mathematics education?* The discussion will concentrate first on applications of CL and second on psychosocial outcomes of CL - positioning both in theory and previous research. Central to the discussion is how the findings of the three articles build on each other and how they collectively contribute to the advancement of the field. The chapter will conclude with a discussion of the limitations of the thesis.

6.1 The Applications of CL

In Papers I and III, we examined the application of CL elements, i.e., CL group features and CL structures, in international and Norwegian undergraduate mathematics and science education. In regard to CL group features, we found that a clear majority of the CL studies included in Paper I employed groups of four members, thus meeting the recommendations on group size of CL literature (Johnson & Johnson, 1999; Kagan, 2021; Millis & Cottell, 1998). Another of our findings was that less than half of the groups were formed by the teacher and were heterogeneous and thus, not meeting the recommendations on group composition of most CL literature (Johnson & Johnson, 1999; Kagan, 2021; Millis & Cottell, 1998). However, when group composition was held up against the findings, we found no evidence that teacher-selected heterogenous groups led to more positive outcomes than did the student-selected groups. That quite a few of the studies did not meet CL recommendations and yet, that this seemingly had no impact on the results of the studies is worth mentioning - but it is hard to identify a reason for this apparent gap. It may be that the population, i.e., international undergraduate mathematics and science education, differs from other populations. It may also be that the effect of group composition lessened in combination with other CL structures. Or it may be due to other reasons. If causality is to be determined here, more stringently controlled research comparing various group composition strategies in STEM higher education is needed. It is, however, concerning that the studies, we identified in Paper I, examine outcomes of CL while not necessarily following the recommendations given in the CL literature (Johnson et al., 1998a; Kagan, 2021; Millis & Cottell, 1998). This lack of coherence and alignment with theory may lead to invalid results as it may become unclear what these studies are investigating.

Further, in Paper I, we discovered that formal CL groups may be better suited than informal CL groups to improve academic success, sense of belonging, and academic self-efficacy. Formal CL groups last from one class to several weeks or months and informal CL groups last from few minutes to one class. This finding suggests that undergraduate mathematics and science students need to partake in well-designed and theoretically motivated group learning experiences (Bandura, 1986) that last over some time in order to improve their academic success, sense of belonging, and academic self-efficacy in the process. The evidence of formal CL groups for successful psychosocial outcomes among undergraduate science (biology) students identified in Paper I was supported in Paper III in this thesis. In addition, Paper III contributed with new knowledge as this evidence was extended to a new setting, i.e., a digital setting, and new outcomes, i.e., science confidence and loneliness. Consequently, formal CL groups may constitute a prime example of how educators can facilitate positive interdependence and individual accountability in student groups, the two most important CL principles in higher education (Johnson & Johnson, 1999; Millis & Cottell, 1998). When CL groups last a certain amount of time, group members become increasingly and positively dependent on each other and it becomes increasingly difficult to not contribute.

In Paper I, we found out that the most popular CL structures in studies on CL in undergraduate mathematics and science education were (rotating) roles and POGIL (Box 1, 2.1.3), and both structures were primarily associated with increased academic success (Byrne, 2015; Daniel, 2016; Díaz-Vázquez et al., 2012; Hein, 2012; Ott et al., 2018a; Stanford et al., 2016; Warfa et al., 2018) and secondly with improved attitudes (Chase et al., 2013; Ott et al., 2018b; Pilcher et al., 2015). Some of these studies did not, however, find that CL led to increased academic success or improved attitudes. Suggested reasons for these findings included deficiency of perceived purpose and contribution in role allocation (Ott et al., 2018b), the use of study activities which did not contain all required elements as prescribed by the POGIL method (Chase et al., 2013), and lectures incorporating some student-centered activities and thus possibly reducing differences between control and experimental groups (Canelas et. al, 2017). Thus, following prescribed CL elements, isolating CL elements, and controlling for confounding variables were highly prioritized study design features in Paper III in this thesis – and through the use of stringently CL approaches and a range of design controls (see Paper III), we were able to infer an effect of the incorporated CL elements on the measured outcomes.

Other CL structures identified in Paper I were jigsaw, Student Team Achievement Division (STAD), and Think-Pair-Share (Box 1, 2.1.3). Like POGIL and rotating roles, the studies employing these CL structures all led to enhanced academic success, but also to increased attendance (Daniel, 2016), generic skills (Pilcher et al., 2015), sense of belonging, and academic self-efficacy (Furuto, 2017; Wilton et al., 2019; Yapici, 2016). Our summary of these findings suggests that rotating roles, POGIL, jigsaw, STAD, and Think-Pair-Share all seem to be suitable structures to increase academic success, in particular, in undergraduate mathematics and science education – and that the last three structures may improve other student outcomes as well. As our findings in Paper I support previous research on CL structures in other subject disciplines in both higher education and elsewhere (Gilles, 2003, 2008; Johnson & Johnson, 1999; Johnson et al., 1998a; Romero, 2009), we chose jigsaw as the main CL structure in Study III in this thesis. Our findings in Paper III both supported this previous research on jigsaw structures while conveying new knowledge. The CL jigsaw structures we implemented led to increased sense of belonging among the students, as it has done in previous studies (Furuto, 2017; Wilton et al., 2019; Yapici, 2016), thus strengthening the existing evidence base. Further, the implemented jigsaw structures led to new outcomes, i.e., increased scientific confidence and generic skills and reduced loneliness –in settings not previously explored, i.e., Norwegian higher education and digital platforms, thus extending the existing evidence base.

Although 19 of the 24 studies we reviewed in Paper I exclusively found a positive increase in measured outcomes, 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. To explore these uncertainties, more research comparing different CL structures is needed. Further, 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 we searched grey literature (Booth et al., 2016) for Paper I, in an attempt to avoid publication bias, we cannot exclude that it has played a role.

Identifying and assessing applications of CL within a particular population such as international and Norwegian undergraduate mathematics and science education is important as it enables research-based teaching approaches to CL. Such identification and assessment may give insight into “what works” and “what is appropriate” in a certain population and thus reduce the dangers of trial and error in one’s own teaching. Hence, our findings in Paper I can convey valuable knowledge for mathematics and science educators, embarking on CL teaching and/or research – as it did for me in this thesis. Based on the findings in Paper I, Paper III in this thesis exemplifies a research-based approach to CL. As many CL methods may be transferable to a synchronous online context (Davidson, 2021), the digital CL methods employed in Study III were carefully selected on the basis of the findings of both group features and CL structures in Paper I. Thus, the CL methods ranged from formal CL groups of four teacher-selected, and heterogeneous students (Johnson & Johnson, 1999; Kagan, 2021; Millis & Cottell, 1998), cooperation as a learning objective in the course description to motivate the intervention and its assessment (Cheruvelil et al., 2020), group contracts (Cheruvelil et al., 2020; Oakley et al., 2004; Aakre & Mørkve, 2021), and jigsaw (Box 1, 2.1.3) (Daniel, 2016; Pilcher et al., 2015; Yimer & Feza, 2020). By implementing CL group features and CL structures which had been empirically tested in previous research in similar populations, the dangers of trial and error were avoided, and the target outcomes achieved. Further, by employing these CL methods in a digital setting, this thesis adds valuable and new knowledge to the extent of the applications of CL methods in digital higher education and certainly in undergraduate STEM education. Not only do the results in Paper III suggest that CL methods from physical settings may be transferable to digital settings. When the results of these CL methods are compared to the face-to-face CL studies in Paper I (Canelas et al., 2017; Furuto, 2017; Kocak, 2008; Pilcher et al., 2015; Rattanatumma & Puncreobutr, 2016; Wilton et al., 2019; Yapici, 2016), the added value of CL methods in digital settings may exceed those in physical settings. However, as in

Paper I, it is not possible to know from Study III alone if the positive outcomes are due to one specific digital CL method over another or the combination of digital CL group features and structures. Still, with digital teaching and learning on the rise, the knowledge gained from the digital nature of Study III is valuable and deserves further exploration. Thus, future digital CL studies should both attempt to validate the findings in Paper III, preferably in STEM populations elsewhere, and implement CL elements not included in Paper III. Based on the findings from Paper I, other promising CL structures in STEM populations that may be transferred to digital settings include roles, POGIL, and STAD (Box 1, 2.1.3).

6.2 The Outcomes of CL

The outcomes of CL were examined in all three papers. As a scoping review, Paper I sought to identify all existing student outcomes of CL in international undergraduate mathematics and science education – and point to knowledge gaps within these outcomes. These findings and knowledge gaps regarding student outcomes of CL elements were then further explored in Papers II and III. Specifically, Paper II and in part Paper III examined outcomes which Paper I found to be promising, but which had not been studied a lot in previous international undergraduate STEM education. Thus, Paper II, through a cross-sectional survey study, sought only to examine if a relationship between self-reported CL and sense of belonging and generic skills exists in Norwegian undergraduate STEM education – and give a theoretical explanation to such relationship. Paper III, on the other hand, sought to validate the findings of Paper II by (i) including sense of belonging and generic skills as CL outcomes in a different population, i.e., among Norwegian undergraduate biology students, (ii) identify causality between CL and sense of belonging and generic skills by employing a different study design, i.e., a longitudinal quasi-experimental intervention study, and (iii) explore if the effect of CL on sense of belonging and generic skills would extend to a different setting, i.e., digital teaching and learning. Further, Paper III included two outcomes not identified in Paper I, i.e., science confidence and loneliness. These outcomes were primarily selected due to COVID-19 pandemic reports on declining student belonging which may lead to loneliness

(Arslan, 2020; Palikara et al., 2021) and to reduced science confidence (Bandura, 1986; Lent et al., 1994), see section 3.1.1 and 3.1.3.

6.2.1 Academic Success

The primary student outcome of CL in international undergraduate mathematics and science education we identified in Paper I was academic success, i.e., either enhanced content knowledge or academic achievement or both. In fact, of the 24 reviewed studies 21 studies included content knowledge and/or academic achievement as the outcome measure, and 17 of these reported an improvement following the implementation of CL elements. Thus, our findings in this review add to the extensive evidence base regarding positive relationships between CL and academic success in STEM higher education and in general (Apugliese & Lewis, 2017; Johnson et al., 1998b; Kyndt et al., 2013; Romero, 2010). Academic success may be placed in the first of three categories of CL outcomes identified by Johnson and Johnson, called "Efforts to achieve" (Johnson & Johnson, 1989). An explanation for the positive relationship between CL and academic success may be found in social interdependence theory (Deutsch, 2012) and CL literature (Johnson & Johnson, 1989; 1999) which underline that academic success improves when students work together to achieve a common goal, i.e., when they are positively interdependent. When we compared the four studies in Paper I which found no improvement in academic success (Canelas et al., 2017; Chase et al., 2013; Harlow et al., 2016; Leung et al., 2017), we identified a common trait in three of the four studies. In these three studies, CL structures were only carried out for a short amount of time, from one test (Canelas et al., 2017), or one laboratory session (Leung et al., 2017) to three discussion sessions (Chase et al., 2013). This finding may indicate that duration is important to obtain enhanced academic success from CL elements in mathematics and science undergraduate education and thus, that formal CL groups should last a certain time and certainly longer than one class.

6.2.2 Attitudes

Our findings from Paper I show that the second most examined student outcome of CL elements in undergraduate mathematics and science education was attitudes.

Attitudes was examined in ten of the 24 reviewed studies. Attitudes may be placed in the second of the three categories of CL outcomes called “positive relationships” (Johnson & Johnson, 1989). Applying social interdependence theory, improved attitudes towards university life, the discipline, the learning process, or group work itself will develop through positive interdependence in groups. Of particular importance here is discussions, i.e., promotive interaction, where students learn and model the norms and values of university life (Johnson et al., 2014). Thus, CL makes up an effective tool for improving student attitudes (Johnson et al., 2014; Petty & Briñol, 2010) and this relationship has also been established in previous research (Johnson et al., 1998). Two of the ten studies examining attitudes in Paper I (Chase et al., 2013; Ott et al., 2018b) found no improvement in student attitudes following implemented CL methods, and we discovered that the common denominator for these two studies was the use of roles. This does not necessarily mean that roles are not suited to improve student attitudes as hypothesized by Johnson & Johnson (2014). However, if they are to do that in undergraduate mathematics and science education, it may according to Ott et al. (2018b) 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 (Chase et al., 2013).

6.2.3 Psychosocial Outcomes

The term “psychosocial outcomes” in this thesis was adopted to stress the importance of the surroundings, e.g., the teaching and learning strategies students face, and how these may influence the students’ perceived sense of belonging, generic skills, science confidence, and loneliness in both positive and negative ways.

6.2.4 Sense of Belonging

Sense of belonging was included as a CL outcome in all three papers in this thesis. In Paper I, we identified three studies in which sense of belonging was the measured outcome and all reported positive findings (Furuto, 2017; Wilton et al., 2019; Yapici, 2016). Two of these studies were set in the US (Furuto, 2017; Wilton et al., 2019), the third in Turkey (Yapici, 2016), and all three were characterised by CL formal groups lasting for a minimum of 10 weeks and based on well documented CL structures such as STAD or Think-Pair-Share (Box 1, 2.1.3). In Paper II, when establishing this relationship in a range of STEM disciplines in a setting not previously examined, i.e., among Norwegian undergraduate students, we found that the levels of sense of belonging increased with the levels of self-reported CL. We also found that the CL principle contributing most to this positive relationship, was "promotive interaction". According to social interdependence theory, promotive interaction arises from positive interdependence and leads to psychological health (Deutsch, 2012), including sense of belonging. Thus, it is the particular characteristics of promotive interaction, i.e., interaction that takes place when group members encourage each other's contributions through listening, exchanging ideas, offering explanations, and constructive feedback (Gillies, 2014, p. 131), which may bring about a feeling of personal acceptance and value among peers (Johnson & Johnson, 1990) - and ultimately fulfill our need to belong (Baumeister & Leary, 1995).

In Paper III, where we went a step further and conducted a quasi-experiment in a digital setting to test the effect of digital CL methods on sense of belonging among undergraduate biology students in Norway, we found that the effect of digital CL on sense of belonging was even greater than studies from the physical settings we identified in Paper I (e.g., Yapici, 2016; Wilton et al., 2019). Both the effect size (d = 0.5 to 0.8) (own calculations) of CL methods in Yapici (2016) and the influence (r = .30 to .49) of enhanced structure in Wilton et al. (2019) on sense of belonging were moderate in magnitude (Cohen, 2013) while the effect size of digital CL on sense of belonging in Study III was large (η_p^2 = .34). Taken together with the drastic decrease in sense of belonging after the first period of digital lectures in Study III, we argue that sense of belonging might be particularly vulnerable to traditional teacher-centered instruction in a digital setting. If that is the case, it is an additional indication that teachers need to carefully consider how to teach digitally, and, encouragingly, that in a digital setting the added value of student-centered learning methods, exemplified by CL, may exceed those in physical settings.

This thesis has primarily contributed new knowledge on the relationship between CL and sense of belonging in four ways: (a) the setting where the research took place, i.e., in Norwegian undergraduate STEM education and among undergraduate biology students in a digital setting, (b) the type of interaction identified in the explanation of this relationship, i.e., “promotive interaction”, (c) the theoretical underpinning of the relationships between CL and sense of belonging through social interdependence theory, and (d) the effect of digital CL on sense of belonging was greater than studies from physical settings.

6.2.5 Generic Skills

We found that generic skills, both self-reported and objectively measured, was the third most studied student outcome of CL elements in Paper I, and self-reported generic skills were also included as outcomes in both Papers II and III in this thesis. That CL elements may lead to the development of a range of generic skills has been hypothesized by several CL authors (Johnson & Johnson, 2002; 2016; Millis & Cottell, 1998; Slavin, 1996) and the findings in this thesis seem to support this hypothesis.

In Paper I, we discovered that all of the reviewed studies (n=7) examining generic skills as a CL outcome, found a positive relationship among the two. Further, we identified several generic skills, i.e., teamwork, problem-solving, critical thinking, communication, and metacognition, as positive results of implementing different CL elements in international undergraduate mathematics and science studies. Based on this identification, in Paper II, we established a positive and significant relationship between self-reported CL and generic skills across a variety of STEM disciplines in a Norwegian setting. Further, we found a strong correlation between both outcomes measured in Paper II, i.e., generic skills and sense of belonging. Theoretically, it is likely that sense of belonging may lead to increased generic skills. According to Baumeister and Leary (1995) we will strive to fulfill our need to belong through interaction, and previous research shows that generic skills are developed by way of student interaction in groups (Ballantine & McCourt Larres, 2007; Kember et al., 2007; Smith & Bath, 2006; Tynjälä & Gijbels, 2012; Virtanen & Tynjälä, 2019;

2022). On the other hand, it is also likely that students who master generic skills at an early stage will experience an enhanced sense of belonging. Demonstrating solid generic skills may lead to recognition from peers and educators alike and is also likely to be reflected in good grades, which in turn may affect sense of belonging positively. A third explanation may be that sense of belonging and generic skills reinforce each other. This third explanation is in accordance with CL and social interdependence theory which states that there is a reciprocal relationship among CL outcomes (Johnson & Johnson, 1989, p. 9). If this third explanation rings true, it is yet another indication that there may be a lot to be gained by implementing CL methods in one's teaching. Last, but not least we identified "promotive interaction" as the CL principle contributing most to the relationship between CL and generic skills. In social interdependence theory, promotive interaction is considered to arise from positive interdependence and an important step leading to positive student outcomes (Deutsch, 2012). While previous studies have hypothesized and demonstrated that group work can promote both self-reported and objectively measured generic skills (Ballantine & McCourt Larres, 2007; Kember et al., 2007; Smith & Bath, 2006; Tynjälä & Gijbels, 2012; Virtanen & Tynjälä, 2019, 2022), they have not empirically examined the characteristics in group work that promote these outcomes, and as such this thesis adds to the existing knowledge base.

In Paper III, we tested the relationship between CL and generic skills experimentally in a digital setting. Here, we found that the effect of digital CL methods on self-reported generic skills among undergraduate biology students was significant and just as strong (η_p^2 = .44) as the effect in physical STEM undergraduate settings elsewhere (Canelas et al., 2017; Pilcher et al., 2015) and digital settings in other populations (Lee et al., 2016; Parsazadeh et al., 2018). Thus, a previous study from a physical setting did, like Study III, find significant increases and very large effect sizes ($d >$ 0.8) (Cohen, 2013) in a range of self-reported generic skills following a CL intervention versus traditional lectures (Canelas et al., 2017). Likewise, a qualitative study replacing recipe-based lab teaching by jigsaw methods found a substantial increase in the students' metacognitive skills (Pilcher et al., 2015). The few studies on the effects of online CL on generic skills in STEM higher education, we were able to find, showed that the effect size of digital CL on information evaluation skills (Parsazadeh et al., 2018) was very large ($d > 0.8$) (Cohen, 2013) and that digital CL led to significantly increased critical thinking skills development (Lee et al., 2016). The results in Paper III supports their general findings and adds new knowledge as it was set in a sociocultural settings not previously examined.

Taken together, this thesis has supported both the hypothesis and previous research on the positive relationship between CL and generic skills. It has also contributed with new knowledge by examining these relationships in new settings, i.e., in undergraduate STEM education in Norway and on a digital platform. Paper II adds to the explanation of previous research and theory stating that generic skills develop through a combination of content knowledge and active learning methods, particularly group work (Ballantine & McCourt Larres, 2007; Kember et al., 2007; Smith & Bath, 2006; Tynjälä & Gijbels, 2012; Virtanen & Tynjälä, 2019, 2022). It does so by *empirically* identifying the CL principle most important to and the t*ype of interaction* needed to promote generic skills, i.e., promotive interaction. Further, Paper II provides a theoretical framework for understanding the relationships between CL and generic skills. Lastly, Paper III demonstrates how the effect of digital CL methods on self-reported generic skills among undergraduate Norwegian biology students was significant and just as strong as the effect in physical STEM undergraduate settings elsewhere (Canelas et al., 2017; Pilcher et al., 2015) and digital settings in other populations and sociocultural settings (Lee et al., 2016; Parsazadeh et al., 2018).

6.2.6 Science Confidence and Loneliness

Although not identified as outcomes in Paper I, science confidence and loneliness were examined as CL outcomes in Paper III motivated by the emergent need to mitigate psychosocial concerns among higher education students during the COVID-19 pandemic (Børve et al., 2021; Camfield et al., 2021; Deng et al., 2021; Lederer et al., 2021; Phillips et al., 2022; Sivertsen, 2021; Tice et al., 2021; Werner et al., 2021). In the case of scientific confidence, we found that it increased over time in response to the CL intervention, with a large effect size (η_p^2 = .34). Although it seems as if Study III is the first from STEM higher education to associate CL with science confidence, previous research exists on the relationship between CL and different types of academic self-efficacy (Furuto, 2017; Rivera, 2013), which may resemble science confidence. These studies found significant increases in student academic self-efficacy following CL interventions employing heterogenous grouping, roles, and jigsaws (Box 1, 2.1.3) compared to traditional teaching, and the effect varied from moderate (d = 0.5 to 0.8) (Rivera, 2013) to very large (d > 0.8) (Furuto, 2017). Thus, our findings indicate for the first time that science confidence may be heavily enhanced by CL and that the positive effect of CL on science confidence may be minimum as substantial as the effect on academic self-efficacy. The reasons why science confidence appears to be heavily enhanced by CL may be due to certain experiences that are necessary for developing a student's positive self-evaluation (Bandura, 1986: 1997; Lent, 1994). These include social and personal experiences such as working with specific scientific tasks, interaction, and immediate feedback and these experiences all constitute important parts of CL and are also found in social interdependence theory. For example, in social interdependence theory, working with specific tasks may be exemplary of the effective actions that group members engage in to reach their shared goal (Figure 1, 2.2.2) while interaction and immediate feedback are characteristic of promotive interaction in which group members encourage each other's contributions through listening, exchanging ideas, offering explanations, and constructive feedback (Gillies, 2014, p. 131).

In the case of loneliness, we found that it decreased with a moderate effect size (η^2 = .13) in response to the CL intervention. The effect of digital CL on loneliness in Paper III may according to Cohen (2013) be considered in the upper range of moderate (η_p^2 >.06 to .13), while the effect of traditional CL on loneliness in the one study we have managed to identify (Kocak, 2008), was in the lower range of moderate (η^2 = .06). This may indicate that the effect of CL on loneliness in digital settings is at least as substantial as in traditional settings, perhaps more. Increasing opportunities for social interaction is particularly promising to prevent and counteract loneliness (Hawkley & Cacioppo, 2010) and thus, the reduction in loneliness measured after the CL intervention compared to digital lectures could be explained by the type of interaction and the types of student groups. Promotive interaction, with its mutual encouragement and support due to the inherent positive interdependence among group members, as posited by social interdependence theory, is believed to lead to positive psychological health (Deutsch, 2012; Johnson & Johnson, 2018). Further, belonging to highly structured, small, heterogeneous, and formal CL groups (Johnson et al., 1994; Johnson, 1992; Kagan, 2021; Millis & Cottell, 1998) may, to a larger degree than other types of groups, facilitate the necessary personal safety to interact and promote inclusion (Adriansen & Madsen, 2012; Baumeister & Leary, 1995). In a digital setting where the threshold to initiate dialog seems higher than in a physical setting, such personal safety may be considered particularly important. With so few studies on the relationship between CL and loneliness, with loneliness on the rise in higher education (Knapstad et al., 2018; Phillips et al., 2022; Sivertsen, 2021; Werner et al., 2021), and the severe consequences of loneliness (Hayley et al., 2017; Holt-Lunstad et al., 2015; Richardson et al., 2017), Paper III does not only cast light on an important issue, but it exemplifies how this issue may be addressed.

As neither science confidence nor loneliness has been examined as typical CL outcomes in previous research, Paper III in this thesis contributes new knowledge on the effects of CL - and it does so in a setting which has not previously been explored and which is undoubtably becoming increasingly relevant, a digital undergraduate STEM setting.

6.3 Summary of the Findings

In sum, the findings in this thesis, which have been obtained through a diversity of methods and which continuously build on each other, have contributed to the advancement of the field in several ways. First, the focus is on psychosocial outcomes of CL rather than academic success. In light of the recruitment, retention, and learning outcomes challenges faced by the international STEM community (Canelas et al., 2017; Honey et al., 2020; Johnson et al., 2002; Leggett et al., 2004;

Lisberg & Woods, 2018; Taylor, 2016; Waite & McDonald, 2019), this shift in focus is both important and warranted, and the research in my thesis has demonstrated how to promote such positive psychosocial outcomes through small changes. Second, the research has been set in a population and a mode not previously researched, i.e., among undergraduate STEM students in Norway and in a digital teaching and learning setting. Research results are not necessarily transferable to other continents, countries, or cultures and thus, empirical research and knowledge in a variety of settings, as demonstrated in this thesis, is necessary. Only then, do we obtain the knowledge foundation we need to advance teaching in our particular sociocultural environment. Further, initiatives to implement digital learning in higher education have only gained strength in the wake of the pandemic (European Commision, 2021) and thus, there is a need to identify digital evidence-based and student-centered teaching and learning strategies that facilitate learning and well-being. Paper III in this thesis exemplifies one such strategy. Third, this thesis has explored a theoretical framework, social interdependence theory, to explain and advance the relationship between CL and select psychosocial outcomes in undergraduate STEM education. As emphasized by Johnson and Johnson (2021, pp. 48-49), any practical procedure should be derived from theory – and any theory should be validated by research. Theoretical underpinnings are not only necessary to understand empirical relationships, but they are also what drive research forward. However, as indicated in Paper I, theory and the application of theory appear to be somewhat deficient in the STEM discipline-based educational communities. By including a solid theoretical framework, this thesis addresses this gap and advances the field. Fourth, one of the most prominent characteristics of social interdependence theory, promotive interaction, has through empirical investigations been identified to be of particular significance to psychosocial student outcomes. Promotive interaction arises due to positive interdependence in groups. Thus, implementing elements underpinning positive interdependence reduces trial and error and gives valuable direction to our teaching of group work. Examples of such elements demonstrated in this thesis are group features e.g., group size, composition, and duration, and CL structures, e.g., roles, group contracts, and jigsaw (Box 1, 2.1.3). Fifth, the application of such CL elements in Paper III of this thesis has highlighted that these methods are transferable to online synchronous settings – and showcases how one may do it. Sixth, this work has shown that the effect of CL methods in digital settings seems to be just as, if not more, substantial in promoting positive psychosocial outcomes as in traditional physical settings. That the added value of student-centered learning methods, exemplified by CL, may exceed those in physical settings constitutes a real incentive to implement such methods. Seventh, the effects of digital CL compared to digital lectures may indicate that methods to explicitly enhance psychosocial outcomes are particularly valuable in digital settings. Thus, we as teachers need to carefully consider how to teach digitally, and new digital learning methods should be developed, tested, and communicated.

6.4 Limitations

When interpreting the results of this thesis, it is necessary to consider both limitations and strengths of the research project. One of the limitations in all three studies is selection bias. In Paper I, very strict inclusion and exclusion criteria limited to undergraduate mathematics and science *discipline* education 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 mathematics and science *professional* studies, particular in study programs for pre-service teachers which might be transferable to mathematics and science discipline studies. As the primary purpose of Paper I was to inform and assess the CL applications and outcomes in mathematics and science *discipline*-based courses, the aforementioned limitation in the inclusion and exclusion criteria may be considered reasonable. Also, there may be great differences in the organization of pre-service teacher education and solely discipline-based education internationally. Some places pre-service teacher students attend courses or even colleges consisting only of professional studies while others attend mixed courses and colleges – and that may affect the applications and outcomes of CL in discipline-based mathematics and science studies. Pre-service mathematics and science teacher students are not necessarily identical to students in solely mathematics and science discipline-based courses, either. In fact, as CL is a well-known teaching and learning method in schools, it is fair to argue that pre-service teacher students might favour such strategies more than students not preparing to become nor interested in becoming teachers – and including such findings in the review might have skewed the conclusions.

In Papers II and III, the participants were all drawn from convenience samples. In the case of Paper II, only courses implementing variants of student cooperation were invited to participate and in Paper III, the participating biology course was primarily selected because of the positive attitude to cooperation by the educator in charge. In both cases these circumstances might have affected the students' answers and as selection bias cannot be controlled for by statistical procedures, it may pose a threat to the internal validity of the studies. However, it is often difficult to avoid some degree of selection bias in educational research and in fact random assignment is often considered to be unethical (Adelson, 2013). Further, there may be other good reasons to limit one's sample selection. In Paper II, it was difficult to include courses not engaged in student cooperation. To determine the level of CL and the relationship between CL on the one hand and sense of belonging and generic skills on the other, it was necessary to measure CL. If participants from random sampling not having experience with group work had been included, the survey items would not have made any sense. In Paper III, selection bias extends to the setting of the study. As previously stated, the intervention was conducted in the middle of the COVID-19 pandemic when higher education in Norway was subjected to severe restrictions and the Norwegian population showed increasing signs of de-motivation and frustration with the seemingly endless situation. This backdrop may have influenced the answers of the students. Perhaps they would not have felt that discouraged from digital lectures nor that positive towards CL methods under different circumstances. Even if the setting in Paper III may cause threats to internal validity, it may be argued that the existing situation at that point was extreme for all and that in itself should not stand in the way of research. Rather, we should seize the opportunity to obtain knowledge in times of extremities, especially if the research may benefit the individuals. During the period of the intervention study, all higher education educators were recommended implementing initiatives to mitigate student concerns (Børve et al., 2021) and thus, the situation at hand, including both recommendations and the students' immediate situation, guided the decision to carry out the research the way, we did. Further, to strengthen causal claims and internal validity, many design controls were included. These will be elaborated on in the next paragraph.

A second limitation is weaknesses in study designs. In Paper I, one may argue that the lack of critical appraisal may pose a threat to the validity of the reviewed studies and as such also pose a limitation to the knowledge base obtained through the scoping review. Nevertheless, scoping reviews are one of the most systematic of the review types and when a full systematic review is not feasible due to time or resource limits, as e.g. typical of PhD-projects, a scoping review is fit for many purposes and a product in its own right (Gough & Thomas, 2017). In Paper II, a cross-sectional study design was employed and thus, no causal relationships could be claimed. However, one may argue that just knowing whether a relationship between the studied variables exists or not is valuable. These relationships had never been examined in this setting previously and the findings of the cross-sectional study paved the way and informed Study III in valuable ways. In Paper III, the quasi-experimental design lacked the advantages provided by randomization, and a control group could have strengthened the causal claims (Shadish et al., 2002). To mitigate these weaknesses in Study III, many design controls were employed to strengthen causal claims (Shadish et al., 2002). These design controls and in particular the use of a double pre-test and a follow-up indicated a real and substantial difference in the students' psychosocial outcomes following the implemented digital CL methods compared to digital lectures. Further, a repeated measures design enables us to detect within-person change over time and has high statistical power (Guo et al., 2013, p. 100) - and the effect sizes of the intervention were all very large, except for the impact on loneliness which was in the upper range of moderate. Thus, it is reasonable to infer that the changes in student scores in the sample were linked to the CL intervention rather than to confounding variables (Shadish et al., 2002).

Third, the measures in Papers II and III may constitute a limitation. Self-reported measures may be deemed unreliable, e.g., due to social desirability bias (Czaja & Blair, 2005), compared to more objective measures. Further, applying measures that have not previously been validated in a Norwegian setting may pose limitations to the studies, and being a relatively new scale, the *Cooperative Learning Application Scale* (CLAS) may have added to these uncertainties. However, as we were primarily interested in the students' own perceived psychosocial outcomes, self-reported measures may be considered the best choice. Further, sense of belonging, loneliness, and different types of confidence are all subjective and internal perceptions that may be difficult to measure objectively. As neither CL nor sense of belonging had been thoroughly researched in Norwegian higher education, it was deemed necessary to employ measures not previously used. To minimize the risks and ensure validity, many steps to validate the measures were taken, see section 4.3 and Paper II. Finally, the internal consistency of all measures was good or excellent (Cronbach, 1951).

One last, but important, limitation to this PhD thesis is that the external validity is low. The samples in Papers II and III were relatively small convenience samples from STEM populations at only one Norwegian university, making it difficult to generalize the findings to other populations. However, that does not mean that these papers are not valuable in their own right. Rather, they both disseminate findings on under-studied and important relationships relevant to the STEM students themselves, to the international STEM community, and by extension to those seeking cooperative solutions to contemporary challenges.

7. Conclusion

The main research question guiding this thesis was: *How is cooperative learning applied and how is it related to psychosocial outcomes among students in undergraduate science, technology, engineering, and mathematics education?* To address this main question, I have conducted three studies which have informed each other. The findings indicate that cooperative learning (CL) in undergraduate science, technology, engineering, and mathematics (STEM) education is applied in various ways, but not always in alignment with CL theory. The most popular and promising CL applications range from heterogenous, formal CL groups of four students to POGIL, rotating roles, and jigsaw (Box 1, 2.1.3). CL in undergraduate STEM education has traditionally been mainly associated with academic success, but the studies that do examine the effect of CL on psychosocial outcomes report positive findings. This was also the case in Papers II and III in this thesis. Paper II established a positive relationship between perceived CL and perceived sense of belonging and generic skills among a sample of undergraduate STEM students in Norway. Paper III found the effect of digital CL, compared to digital lectures, on psychosocial outcomes (i.e., self-reported sense of belonging, generic skills, science confidence, and loneliness) to be positive, significant, and substantial among undergraduate biology students. Although important in addressing the challenges faced by STEM higher communities, psychosocial outcomes in STEM higher education seem to be under-researched. With this backdrop and the magnitude in effect sizes found in Paper III, psychosocial outcomes among students (i) should be explored further, (ii) may be particularly vulnerable in digital settings, and (iii) are important to be aware of and address. There may be many ways to do just that. Based on the findings in this thesis, I suggest that CL is a suitable way to enhance psychosocial outcomes among undergraduate STEM students.

7.1 Implications for Teaching

At first glance it might seem challenging to implement student-centered approaches such as CL in Norwegian undergraduate STEM education. Barriers may include lack of resources, large introductory courses, extensive use of lectures, and testing through summative exams rather than formative assessments. However, the findings in this thesis indicate that student affect matters and it is time we consider the psychosocial dimension of our teaching. Psychosocial factors are vital to overcoming many of the challenges faced by STEM higher education (Aurlien et al., 2019; Ayllón et al., 2019; Ballen et al., 2017; Bandura, 1997; Nissen & Shemwell, 2016; Parkes, 2014; Rittmayer & Beier, 2008; Sawtelle et al., 2012; Sæthre, 2014; Thomas, 2012; Tinto, 1975, 1993; Trujillo & Tanner, 2014; van Dinther et al., 2011) and in turn solid STEM student recruitment, retention, and learning outcomes are vital to solving many global challenges (Milgrom-Elcot, 2022; Schreiner et al., 2010; Taylor, 2016; UN, 2015). With this background in mind, it is the *obligation* of STEM educators to identify and deliver teaching and learning conditions that will ensure that our students succeed. This thesis exemplifies how we can do just that, and it does not take that much. Small changes may make a great impact and there are many ways we can go about it as long as our chosen group approach centers around the students and positive interdependence.

Educators contemplating CL elements in their undergraduate STEM education should be aware to include CL group features and CL structures underpinned by CL principles and make sure that these CL group features and CL structures are of a certain duration and/or intensity. POGIL and roles (Box 1, 2.1.3) are the most applied CL elements in undergraduate STEM education and both may lead to positive student outcomes given the POGIL prescriptions are followed, and roles are perceived as purposeful and contributing to team processes and outcomes. Further, the use of jigsaw and group contracts (Box 1, 2.1.3), both in traditional and digital STEM higher education settings, seem to be beneficial to students' psychosocial outcomes. The findings of this thesis in particular stress the importance of the CL principle "promotive interaction" - and thus, I recommend that educators provide students opportunities to participate in groups where they can help and support each other, exchange ideas, communicate thoughts, and offer explanations and constructive feedback (Gillies, 2014, p. 131; Johnson et al., 2014). Educators developing digital courses should take care to include student-centered approaches in their courses. My thesis suggests the positive impacts of this engagement can be meaningful and far-reaching, and colleagues in higher education are urged to seek and implement digital strategies - such as digital CL - that promote student well-being.

7.2 Future Research

Based on our findings in Paper I, there is a need to design studies which explore CL using qualitative methods, in other countries than the US, and in a greater variety of undergraduate mathematics and science disciplines. If the impact of CL elements is to be measured in additional quantitative educational studies, it is important 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 include student outcomes other than enhanced academic success. Research will also benefit from addressing group features, CL structures, and CL principles corresponding to CL theory. In Paper II, a strong correlation between the outcome variables sense of belonging and generic skills was found. As Paper II was a cross-sectional study, no causal relationships could be determined and thus, our understanding of the influence of CL on these variables requires further examination. To address this relationship, future research should address this topic and/or employ designs that may explain causal relationships, e.g., longitudinal studies. The longitudinal study of this thesis, Paper III, did identify causal relationships between CL and a range of psychosocial outcomes, including sense of belonging and self-reported generic skills. As a novelty, Paper III also established causal relationships between CL and science confidence and CL and loneliness. According to previous research and the STEM higher students themselves, enhanced sense of belonging and science confidence are vital to overcome the challenges of recruitment and retention faced by the STEM community - and the increased loneliness faced by the students. Further, improved generic skills may mitigate many of the learning outcome challenges in STEM education. Given *how* important STEM students are viewed for the future and for the global challenges we face, research on more teaching and learning strategies contributing to increased belonging, science confidence and generic skills while reducing loneliness deserve more attention and exploration in STEM higher education. A novelty in this thesis is the exploration of CL and psychosocial outcomes in a fully digital setting, and it would be interesting to see if future research will support or weaken these findings. Thus, more research on the effect of CL on psychosocial outcomes (and indeed other types of outcomes) in digital settings is warranted. Lastly, future CL studies in STEM higher education may advance the field by providing more thorough theoretical frameworks to advance our understanding of the effect of CL on selected student outcomes.

References

Adelson, J. L. (2013). Educational research with real-world data: Reducing selection bias with propensity score analysis. *Practical Assessment, Research & Evaluation, 18*, 15.

Adriansen, H. K., & Madsen, L. M. (2012). Facilitation: A novel way to improve students' well-being. *Innovative Higher Education, 38*(4), 295-308. https://doi.org/10.1007/s10755-012-9241-0

Ainscough, L., Foulis, E., Colthorpe, K., Zimbardi, K., Robertson-Dean, M., Chunduri, P., & Lluka, L. (2016). Changes in biology self-efficacy during a first-year university course. *CBE—Life Sciences Education, 15*(2), ar19. https://doi.org/10.1187/cbe.15-04-0092

Alkan, N. (2016). Psychological sense of university membership: An adaptation study of the PSSM scale for Turkish university students. *Journal of Psychology, 150*(4), 431-449. https://doi.org/10.1080/00223980.2015.1087373

Allen, K.-A., Kern, M. L., Rozek, C. S., McInerney, D. M., & Slavich, G. M. (2021). Belonging: A review of conceptual issues, an integrative framework, and directions for future research. *Australian Journal of Psychology ,73*(1), 87-102. https://doi.org/10.1080/00049530.2021.1883409

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.

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19-32.

Arslan, G. (2020). School belongingness, well-being, and mental health among adolescents: Exploring the role of loneliness. *Australian journal of psychology*. https://doi.org/https://doi.org/10.1111/ajpy.12274

Atxurra, C., Villardón-Gallego, L., & Calvete, E. (2015). Design and validation of the cooperative learning application scale (CLAS). *Revista de Psicodidactica, 20*(2), 339-357. https://doi.org/https://doi.org/10.1387/RevPsicodidact.11917

Aurlien, I., Bjørnsrud, K. M., Haugom, E., & Thrane, C. (2019). Betydningen av målbevissthet og sosial tilhørighet for frafall i høyere utdanning [The importance of goal determination and social belonging for dropout in higher education]. *Uniped 42*(2), 125-138. https://doi.org/10.18261/issn.1893-8981-2019-02-02

Ayllón, S., Alsina, Á., & Colomer, J. (2019). Teachers' involvement and students' self-efficacy: Keys to achievement in higher education. *Plos One, 14*(5), e0216865-e0216865. https://doi.org/10.1371/journal.pone.0216865

Badcock, P. B. T., Pattison, P. E., & Harris, K.-L. (2010). Developing generic skills through university study: A study of arts, science and engineering in Australia. *Higher Education, 60*(4), 441-458. https://doi.org/10.1007/s10734-010-9308-8

Ballantine, J., & McCourt Larres, P. (2007). Cooperative learning: A pedagogy to improve students' generic skills? *Education & Training, 49*(2), 126-137. https://doi.org/10.1108/00400910710739487

Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K. R. (2017). Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning. *CBE—Life Sciences Education, 16*(4), ar56.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology, 4*(3), 359-373. https://doi.org/10.1521/jscp.1986.4.3.359

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497-529. https://doi.org/10.1037/0033-2909.117.3.497

Beard, C., Clegg, S., & Smith, K. (2007). Acknowledging the affective in higher education. *British Educational Research Journal, 33*(2), 235-252. http://www.jstor.org/stable/30032743

Befring, E. (2007). *Forskningsmetode med etikk og statistikk [Research methods with ethics and statistics]*. Det norske samlaget.

Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health, 25*(5), 464-469.

Biesta, G. (2020). Risking ourselves in education: Qualification, socialization, and subjectification revisited. *Educational Theory, 70*(1), 89-104.

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 https://doi.org/10.1007/978-94-007-2324-5_2

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). Sage.

Bourn, D. (2018). *Understanding global skills for 21st century professions*. Springer International Publishing. https://doi.org/10.1007/978-3-319-97655-6_3

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.

Byrne, M., & Flood, B. (2003). Assessing the teaching quality of accounting programmes: An evaluation of the Course Experience Questionnaire. *Assessment and Evaluation in Higher Education, 28*(2), 135-145. https://doi.org/10.1080/02602930301668

Børve, K., Straume, K., Smith, R., Surén, P., Samdal, O., & Hågensen, E. (2021). *Tiltak for oppfølging av studenter ved fagskoler, høyskoler og universiteter under koronoapandemien [Initiatives to follow up students in higher education during the Corona pandemic]* https://www.regjeringen.no/contentassets/459df2260de74e57a9f1317bd0bb492f/tiltak-for-oppfolging-av-studenter-ved-fagskoler-hoyskoler-og-universiteter-under-pandemien.pdf

Camfield, E. K., Schiller, N. R., & Land, K. M. (2021). Nipped in the bud: COVID-19 reveals the malleability of STEM student self-efficacy. *CBE--Life Sciences Education, 20*(2), ar25-ar25. https://doi.org/10.1187/cbe.20-09-0206

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 & Evaluation in Higher Education, 35*(4), 381-402. https://doi.org/http://dx.doi.org/10.1080/02602930902862826

Chalmers, A. F. (1999). *What is this thing called science?* (3rd ed). Open University Press.

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

Cheruvelil, 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. https://doi.org/http://dx.doi.org/10.1525/abt.2020.82.1.18

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic Press.

Cohen, M. W. (2010). Cooperative learning in educational psychology. Modelling success for future teachers. In B. J. Millis (Ed.), *Cooperative learning in higher education* (pp. 69-89). Stylus.

Constable, N. (2021). Migrant mothers, rejected refugees and excluded belonging in Hong Kong. *Population Space and Place, 27*(5), n/a. https://doi.org/10.1002/psp.2475

Cornford, I. (2005). Challenging current policies and policy makers' thinking on generic skills. *Journal of Vocational Education & Training ,57*(1), 25-45. https://doi.org/10.1080/13636820500200273

Cotner, S., Jeno, L. M., Walker, J. D., Jørgensen, C., & Vandvik, V. (2020). Gender gaps in the performance of Norwegian biology students: The roles of test anxiety and science confidence. *International Journal of STEM Education, 7*(1), 1-10. https://doi.org/10.1186/s40594-020-00252-1

Cottell, P. G. (2010). Cooperative learning in accounting. In B. J. Millis (Ed.), *Cooperative learning in higher education* (pp. 11-33). Stylus.

Crano, W. D., Brewer, M. B., & Lac, A. (2015). *Principles and methods of social research* (3rd ed.). Routledge.

Cromley, J. G., Perez, T., & Kaplan, A. (2016). Undergraduate STEM achievement and retention: Cognitive, motivational, and institutional factors and solutions. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 4-11. https://doi.org/10.1177/2372732215622648

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. https://doi.org/10.1007/BF02310555

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.

Cutrona, C. E. (1982). Transition to college: Loneliness and the process of social adjustment. In L. A. Peplau & D. Perlman (Eds.), *Loneliness: A sourcebook of current Theory, research and therapy* (pp. 291-309). Wiley.

Czaja, R., & Blair, J. (2005). *Designing surveys: A guide to decisions and procedures* (2nd ed.). Pine Forge.

Daniel, K. L. (2016). Impacts of active learning on student outcomes in large-lecture biology courses. *American Biology Teacher, 78*(8), 651-655. https://doi.org/10.1525/abt.2016.78.8.651

Davey, T., Meerman, A., Galán-Muros, V., Orazbayeva, B., & Baaken, T. (2018). *The state of university-business cooperation in Europe (Final report).* https://www.ub-cooperation.eu/

Davidson, N. (2021). Synthesis of CL approaches and a multifaceted rationale for CL - past, present, and future. In N. Davidson (Ed.), *Pioneering perspectives in cooperative learning. Theory, research, and classroom practice for diverse approaches to CL* (pp. 234-254). Routledge.

Deng, J., Zhou, F., Hou, W., Silver, Z., Wong, C. Y., Chang, O., Drakos, A., Zuo, Q. K., & Huang, E. (2021). The prevalence of depressive symptoms, anxiety symptoms and sleep disturbance in higher education students during the COVID-19 pandemic: A systematic review and meta-analysis. *Psychiatry Research, 301*, 113863-113863. https://doi.org/10.1016/j.psychres.2021.113863

Destin, M. (2019). Socioeconomic mobility, identity, and health: Experiences that influence immunology and implications for intervention. *American Psychologist, 74*(2), 207-217. https://doi.org/10.1037/amp0000297

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* (Vol. 2, pp. 275). SAGE Publications Ltd. https://doi.org/10.4135/9781446249222.n40

Dewey, J. (2011). *Democracy and education*. Simon & Brown.

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). https://doi.org/10.20429/ijsotl.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

Drew, J. C., Oli, M. W., Rice, K. C., Ardissone, A. N., Galindo-Gonzalez, S., Sacasa, P. R., Belmont, H. J., Wysocki, A. F., Rieger, M., & Triplett, E. W. (2015). Development of a distance dducation program by a land-grant university augments the 2-year to 4-year STEM pipeline and increases diversity in

STEM. *Plos One, 10*(4), e0119548-e0119548. https://doi.org/10.1371/journal.pone.0119548

Egelandsdal, K., & Ness, I. J. (2020). «Læring som praksis» – John Deweys pragmatisme [Learning as practice» - John Dewey's pragmatism]. In A. G. Danielsen (Ed.), *Til elevens beste – pedagogiske perspektiver* (pp. 61-78). Gyldendal.

Ekholm, E., & Chow, J. (2018). Addressing publication bias in educational psychology. *Translational Issues in Psychological Science, 4*(4), 425-439. https://doi.org/https://doi.org/10.1037/tps0000181

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Publications.

Espeland, V., & Indrehus, O. (2003). Evaluation of students' satisfaction with nursing education in Norway. *Journal of Advanced Nursing, 42*(3), 226-236. https://doi.org/10.1046/j.1365-2648.2003.02611.x

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). https://doi.org/doi.org/10.1103/PhysRevPhysEducRes.15.010132

European Commision. (2021). *Digital Education Action Plan (2021-2027): Resetting Education and Training for the Digital Age [Online].* Retrieved 16.06.22 from https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan_en

European Commision. (2022). *Complete guide to GDPR compliance*. Retrieved 20.04.22 from https://gdpr.eu/

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE edge.

Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. *Psychonomic Bulletin & Review, 19*(6), 975-991. https://doi.org/https://doi.org/10.3758/s13423-012-0322-y

Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education ,75*(3), 203-220. https://doi.org/10.3200/JEXE.75.3.203-220

Furuto, M. (2013). *A case study on a diverse college algebra classroom: Analyzing pedagogical strategies to enhance students' mathematics self-efficacy* [Oregon State University].

Furuto, M. A. (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*(1-2), 35-49.

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. (2014). Developments in cooperative learning: Review of research. *Anales de Psicología 30*(3), 792-801. https://doi.org/10.6018/analesps.30.3.201191

Gillies, R. M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education, 41*(3), 39-54.

Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in Schools, 30*(1), 79-90.

Gough, D., & Thomas, J. (2017). Commonality and diversity in reviews. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 43-70). Sage.

Hagerty, B. M. K., Lynch-Sauer, J., Patusky, K. L., Bouwsema, M., & Collier, P. (1992). Sense of belonging: A vital mental health concept. *Archives of Psychiatric Nursing, 6*(3), 172-177. https://doi.org/10.1016/0883-9417(92)90028-H

Hambleton, R. K. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment, 17*(3), 164-172. https://doi.org/10.1027//1015-5759.17.3.164

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/https://doi.org/10.1080/09645292.2013.813908

Hatch, J. A. (2002). *Doing qualitative research in education settings*. State University of New York Press.

Hawkley, L. C., & Cacioppo, J. T. (2010). Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of Behavioral Medicine, 40*(2), 218-227. https://doi.org/10.1007/s12160-010-9210-8

Hayley, A. C., Downey, L. A., Stough, C., Sivertsen, B., Knapstad, M., & Øverland, S. (2017). Social and emotional loneliness and self-reported difficulty initiating and maintaining sleep (DIMS) in a sample of Norwegian university students. *Scandinavian Journal of Psychology, 58*(1), 91-99. https://doi.org/10.1111/sjop.12343

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics, 19*(4), 451-464. https://doi.org/10.1016/j.labeco.2012.05.014

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

Holland, D. G. (2019). The struggle to belong and thrive. In E. Seymour & A.-B. Hunter (Eds.), *Talking about leaving revisited: Persistence, relocation, and loss in undergraduate STEM education* (pp. 277-327). Springer International Publishing. https://doi.org/10.1007/978-3-030-25304-2_9

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk Factors for mortality: A meta-analytic review. *Perspectives on Psychological Science, 10*(2), 227-237. https://doi.org/10.1177/1745691614568352

Honey, M., Alberts, B., Bass, H., Castillo, C., Okhee, L., Strutchens, M. M., Vermillion, L., & Rodriguez, F. (2020). *Stem Education for the Future: A Visioning Report*.

https://www.nsf.gov/edu/Materials/STEM%20Education%20for%20the%20Future%20-%202020%20Visioning%20Report.pdf

Hughes, M. E., Waite, L. J., Hawkley, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys: Results rrom two population-based studies. *Research on Aging, 26*(6), 655-672. https://doi.org/10.1177/0164027504268574

IBM. (2017). *IBM SPSS statistics for Macintosh, version 25.0.* In IBM Corp.

IBM. (2021). *IBM SPSS Statistics for Macintosh, Version 27.0.* In IBM Corp.

Jacobs, G. M., McCafferty, S. G., & Iddings, A. C. D. (2006). Roots of cooperative learning in general education. In G. M. Jacobs, S. G. McCafferty, & A. C. D. Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 9-17). Cambridge University Press.

Jansen, E., van der Meer, J., & Fokkens-Bruinsma, M. (2013). Validation and use of the CEQ in The Netherlands. *Quality Assurance in Education ,21*(4), 330-343. https://doi.org/10.1108/QAE-11-2012-0041

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction Book Company.

Johnson, D. W., & Johnson, R. T. (1990). Cooperative learning and achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 23-37). Praeger.

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. (2002). Social interdependence theory and university instruction - theory into practice. *Swiss Journal of Psychology, 61*(3), 119-129. https://doi.org/10.1024//1421-0185.61.3.119

Johnson, D. W., & Johnson, R. T. (2005). New developments in social interdependence theory. *Genetic Social and General Psychology Monographs, 131*(4), 285-358. https://doi.org/10.3200/MONO.131.4.285-358

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher ,38*(5), 365-379. https://doi.org/10.3102/0013189X09339057

Johnson, D. W., & Johnson, R. T. (2016). Cooperative learning and teaching citizenship in democracies. *International Journal of Educational Research, 76*, 162-177. https://doi.org/https://doi.org/10.1016/j.ijer.2015.11.009

Johnson, D. W., & Johnson, R. T. (2021). Learning together and alone. The history of our involvement in cooperative learning. In N. Davidson (Ed.), *Pioneering perspectives in cooperative learning. Theory, research, and classroom practice for diverse approaches to CL* (pp. 44-62). Routledge.

Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1994). *Cooperative learning in the classroom*. Association for Supervision and Curriculum Development.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998a). *Active learning: Cooperation in the college classroom* (2nd ed.). Interaction Book Company.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998b). Cooperative learning returns to college: What evidence is there that it works? *Change, July/August*, 27-35.

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, D. W., Johnson, R. T., & Holubec, E. J. (1992). *Advanced cooperative learning* (2 ed.). Interaction Book Company.

Johnson, E., Herd, S., & Tisdall, J. (2002). Encouraging generic skills in science courses. *Electronic Journal of Biotechnology, 5*, 22-23. http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-34582002000200004&nrm=iso

Kagan, S. (1989). The structural approach to cooperative learning. *Educational Leadership, 47*(4), 12.

Kagan, S. (2021). The structural approach and Kagan structures. In N. Davidson (Ed.), *Pioneering perspectives in cooperative learning. Theory, research, and classroom practice for diverse approaches to CL* (pp. 78-127). Routledge.

Katartzi, E., & Hayward, G. (2020). Transitions to higher education: The case of students with vocational background. *Studies in Higher Education (Dorchester-on-Thames), 45*(12), 2371-2381. https://doi.org/10.1080/03075079.2019.1610866

Kember, D., Leung, D. Y. P., & Ma, R. S. F. (2007). Characterizing learning environments capable of nurturing generic capabilities in higher education. *Research in Higher Education ,48*(5), 609-632. https://doi.org/10.1007/s11162-006-9037-0

Kleven, T. A., Tveit, K., & Hjardemaal, F. (2011). *Innføring i pedagogisk forskningsmetode: en hjelp til kritisk tolking og vurdering [Introduction to educational research methods: Help to critical interpretation and assessment]*. Fagboklaget.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guildford.

Knapstad, M., Heradstveit, O., & Sivertsen, B. (2018). *Students' Health and Wellbeing Study 2018*. SiO. https://www.studenthelse.no/tema/sosiale_forhold

Kocak, R. (2008). The Effects of Cooperative Learning on Psychological and Social Traits among Undergraduate Students. *Social behavior and personality, 36*(6), 771-782. https://doi.org/10.2224/sbp.2008.36.6.771

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

Koffka, K. (1935). *Principles of gestalt psychology*. Routledge & Kegan Paul.

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. https://doi.org/10.1016/j.edurev.2013.02.002

Leckey, J. F., & McGuigan, M. A. (1997). Right tracks - wrong rails. The development of generic skills in higher education. *Research in Higher Education, 38*(3), 365-378. https://doi.org/10.1023/a:1024902207836

Lederer, A. M., Hoban, M. T., Lipson, S. K., Zhou, S., & Eisenberg, D. (2021). More than inconvenienced: The unique needs of U.S. college students during the COVID-19 pandemic. *Health Education Behavior, 48*(1), 14-19. https://doi.org/10.1177/1090198120969372

Lee, H., Parsons, D., Kwon, G., Kim, J., Petrova, K., Jeong, E., & Ryu, H. (2016). Cooperation begins: Encouraging critical thinking skills through cooperative reciprocity using a mobile learning game. *Computers & Education, 97*, 97-115. https://doi.org/https://doi.org/10.1016/j.compedu.2016.03.006

Leggett, M., Kinnear, A., Boyce, M., & Bennett, I. (2004). Student and staff perceptions of the importance of generic skills in science. *Higher Education Research & Development, 23*(3), 295-312. https://doi.org/10.1080/0729436042000235418

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45*(1), 79-122. https://doi.org/10.1006/jvbe.1994.1027

Lewin, K. (1948). *Resolving social conflicts: Selected papers on group dynamics*. Harper and brothers.

Leyshon, M. (2011). The struggle to belong: Young people on the move in the countryside. *Population Space and Place, 17*(4), 304-325. https://doi.org/10.1002/psp.580

Lisberg, A., & Woods, B. (2018). Mentorship, mindset and learning strategies: cn integrative approach to increasing underrepresented minority student retention in a STEM undergraduate program. *Journal of STEM Education, 19*(3).

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-Class grouping: A meta-analysis. *Review of Educational Research 66*(4), 423-458.

Lytle, A., De Rosa, A. J., & Fisher, F. T. (2021). A review of psychosocial factors associated with undergraduate engagement and retention in STEM. *American Society for Engineering Education*. https://peer.asee.org/a-review-of-psychosocial-factors-associated-with-undergraduate-engagement-and-retention-in-stem

Maslow, A. H. (1968). *Toward a psychology of being* (2nd ed.). D. Van Nostrand.

Matthews-Ewald, M. R., & Zullig, K. J. (2013). Evaluating the performance of a short loneliness scale among college students. *Journal of College Student Development, 54*(1), 105-109. https://doi.org/10.1353/csd.2013.0003

Milgrom-Elcot, T. (2022, March 23). Rising to the challenge of providing all students with high-quality STEM education. *Brookings*. https://www.brookings.edu/blog/brown-center-chalkboard/2022/03/23/rising-to-the-challenge-of-providing-all-students-with-high-quality-stem-education/

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.

Millis, B. J., & Cottell, P. G. (1998). *Cooperative learning for higher education faculty*. American Council on Education/Oryx.

Moher, D., Liberati, A., Altman, D., & The PRISMA group. (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA statement. *PLoS, Med, 6*(7), e1000097. https://doi.org/https://doi.org/10.1371/journal.pmed.1000097

Muijs, D. (2013). *Doing quantitative research in education with SPSS*. Sage.

Murdock-Perriera, L. A., Boucher, K. L., Carter, E. R., & Murphy, M. C. (2019). Places of belonging: Person- and place-focused interventions to support belonging in college. In M. B. Paulsen & L. W. Perna (Eds.), *Higher education: Handbook of theory and research* (pp. 291-323). Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-030-03457-3_7

Nissen, J. M., & Shemwell, J. T. (2016). Gender, experience, and self-efficacy in introductory physics. *Physical review. Physics Education Research, 12*(2), 020105. https://doi.org/10.1103/PhysRevPhysEducRes.12.020105

NOKUT [The Norwegian Agency for Quality Assurance in Education]. (2021). *The Student Survey*. Kunnskapsdepartmentet [Ministry of Education and Research]. https://www.nokut.no/globalassets/studiebarometeret/2021/hoyere-utdanning/studiebarometeret-2020_hovedtendenser_1-2021.pdf

Norwegian Centre for Research Data (NSD). (2021). *Data Protection Services*. Retrieved 20.04 from https://www.nsd.no/en/data-protection-services

Oakley, B., Felder, R. M., Brent, R., & Elhajj, I. (2004). Turning student groups into effective teams. *Journal of Student Centered Learning, 2*(1), 9-34.

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. *Journal of the Scholarship of Teaching and Learning, 18*(4), 161-180. http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1200478&site=ehost-live

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.

Palikara, O., Castro-Kemp, S., Gaona, C., & Eirinaki, V. (2021). The mediating role of school belonging in the relationship between socioemotional well-being and loneliness in primary school age children. *Australian Journal of Psychology, 73*(1), 24-34. https://doi.org/10.1080/00049530.2021.1882270

Pallant, J. (2016). *SPSS survival manual* (6 ed.). McGraw-Hill Education.

Parkes, S. (2014). Fostering a sense of belonging: Supporting the early formation of student identity as successful learners in Higher Education. *Journal of Learning Development in Higher Education*(7). https://doi.org/10.47408/jldhe.v0i7.229

Parsazadeh, N., Ali, R., & Rezaei, M. (2018). A framework for cooperative and interactive mobile learning to improve online information evaluation skills. *Computers & Education, 120*, 75-89. https://doi.org/https://doi.org/10.1016/j.compedu.2018.01.010

Peplau, L. A., & Perlman, D. (1982). Perspectives on loneliness. In L. A. Peplau & D. Perlman (Eds.), *Loneliness: A sourcebook of current theory, research and therapy* (pp. 1-18). Wiley.

Petty, R., & Briñol, P. (2010). Attitude change. In R. F. Baumeister & E. J. Finkel (Eds.), *Advanced social psychology: The state of the science* (pp. 217-259). Oxford.

Phillips, R., Seaborne, K., Goldsmith, A., Curtis, N., Davies, A., Haynes, W., McEnroe, R., Murphy, N., O'Neill, L., Pacey, C., Walker, E., & Wordley, E. (2022). Student loneliness through the pandemic: How, why and where? *The Geographical journal, 188*(2), 277-293. https://doi.org/10.1111/geoj.12438

Pilcher, L. A., Darren, L. R., Kgadi, C. M., & 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.

Rambøll. (2021). *SurveyXact*. Retrieved 19.10.21 from https://www.surveyxact.no

Ramsden, P., & Entwistle, N. J. (1981). Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology, 51*(3), 368-383. https://doi.org/10.1111/j.2044-8279.1981.tb02493.x

Rattanatumma, 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.

Rayaan. (2022). *Rayaan. Faster systematic reviews*. https://www.rayyan.ai/

Regional Committees for Medical and Health Research Ethics (REK). (2020). *Guidance and Support*. Retrieved 20.04.22 from https://www.norcrin.no/en/regional-committees-for-medical-and-health-research-ethics-rek/

Resnik, D. B., & Shamoo, A. E. (2017). Reproducibility and research integrity. *Accountability in Research, 24*(2), 116-123. https://doi.org/10.1080/08989621.2016.1257387

Richardson, T., Elliott, P., & Roberts, R. (2017). Relationship between loneliness and mental health in students. *Journal of Public Mental Health, 16*(2), 48-54. https://doi.org/10.1108/JPMH-03-2016-0013

Ringdal, K. (2009). *Enhet og mangfold* [Unity and Diversity](2nd ed.). Fagboklaget.

Rittmayer, A. D., & Beier, M. E. (2008). Overview: Self-efficacy in STEM. *SWE-AWE CASEE Overviews, 1 (3)*, 12.

Rivera, N. (2013). *Cooperative learning in a community college setting: Developmental coursework in mathematics.* [Arizona State University].

Romero, C. C. (2010). Cooperative learning instruction and science achievement for secondary and early post-secondary students: A systematic review. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 70*(8-A), 2943.

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. https://doi.org/10.1021/ed1011844

Sawtelle, V., Brewe, E., & Kramer, L. H. (2012). Exploring the relationship between self-efficacy and retention in introductory physics. *Journal of Research in Science Teaching, 49*(9), 1096-1121. https://doi.org/10.1002/tea.21050

Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research, 8*(1), 3-15.

Schreiner, C., Henriksen, E. K., Sjaastad, J., Jensen, F., & Løken, M. (2010). *Vilje-con-valg: Valg og bortvalg av realfag i høyere utdanning [Will-vs-choice: Opt-in or opt-out of STEM disciplines in higher education]*. https://www.naturfagsenteret.no/binfil/download.php?did=6473

Seymour, E., Hunter, A.-B., Laursen, S. L., & Deantoni, T. (2004). Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. *Science Education, 88*, 493-534.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Singer, S. R., Nielsen, N. R., Schweingruber, H. A., National Research Council . Board on Science, E., National Research Council . Division of, B., Social Sciences, E., & National Academies, P. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. National Academies Press.

Sivertsen, B. (2021). *Students' Health and Wellbeing Study 2021*. SiO. https://sioshotstorage.blob.core.windows.net/shot2018/SHOT2021.pdf

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.

Smith, C., & Bath, D. (2006). The role of the learning community in the development of discipline knowledge and generic graduate outcomes. *Higher Education, 51*(2), 259-286. https://doi.org/10.1007/s10734-004-6389-2

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research, 69*(1), 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

Stevahn, L. (2021). The legacy of Morton Deutsch. Theories of cooperation, conflict, and justice. In N. Davidson (Ed.), *Pioneering perspectives in cooperative learning. Theory, research, and classroom practice for diverse approaches to CL* (pp. 17-43). Routledge

Stewart-Brown, S., Evans, J., Patterson, J., Petersen, S., Doll, H., Balding, J., & Regis, D. (2000). The health of students in institutes of higher education: An important and neglected public health problem? *Journal of Public Health, 22*(4), 492-499. https://doi.org/10.1093/pubmed/22.4.492

Säljö, R. (2013). Støtte til læring - tradisjoner og perspektiver [Support for learning - traditions and perspectives]. In R. J. Krumsvik & R. Säljö (Eds.), *Praktisk-*

pedagogisk utdanning. En antologi. [Postgraduate certificate in education. An anthology.] (pp. 53-79). Fagboklaget.

Sæthre, H. Å. (2014). *Å tilrettelegge for at studentene skal lykkes: en verktøykasse for utdanningsledere, studieadministrasjon, studenttillitsvalgte [Facilitating student succes: A toolbox for educational leaders, study administration, student representatives]*. Haas1.

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th, new international ed.). Pearson.

Taber, K. S. (2016). Learning generic skills through chemistry education. *Chemistry Education Research and Practice Journal, 17*, 225-228.

Talanquer, V. (2014). DBER and STEM education reform: Are we up to the challenge? *Journal of Research in Science Teaching, 51*(6), 809-819. https://doi.org/10.1002/tea.21162

Taylor, P. C. (2016). *Why is a STEAM curriculum perspective crucial to the 21st century?* 14th Annual conference of the Australian Council for Educational Research, Brisbane, August 7-9.

Tekna. (2022). Bekymret for ensomhet - Krever planer for økt studentengasjement [Worried about loneliness - Demands action for increased student engagement]. *Tekna*. Retrieved 17.0822 from https://www.tekna.no/aktuelt/bekymret-for-ensomhet---krever-planer-for-okt-studentengasjement1/

Thomas, K., Nilsson, E., Festin, K., Henriksson, P., Lowén, M., Löf, M., & Kristenson, M. (2020). Associations of psychosocial factors with multiple health behaviors: A population-based study of middle-aged men and women. *International Journal of Environmental Research and Public Health, 17*(4). https://doi.org/10.3390/ijerph17041239

Thomas, L. (2012). *Building student engagement and belonging in higher education at a time of change.* P. H. Foundation. https://www.heacademy.ac.uk/system/files/what_works_final_report_0.pdf

Tice, D., Baumeister, R., Crawford, J., Allen, K.-A., & Percy, A. (2021). Student belongingness in higher education: Lessons for professors from the COVID-19 pandemic. *Journal of University Teaching & Learning Practice, 18*(4), 2.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research ,45*(1), 89-125. https://doi.org/10.3102/00346543045001089

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). The University of Chicago Press.

Trujillo, G., & Tanner, K. D. (2014). Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity. *CBE--Life Sciences Education, 13*(1), 6-15. https://doi.org/10.1187/cbe.13-12-0241

Trædal, T. (2020). Bekymring for ensomme studenter over hele landet [Lonely students throughout the country are a concern]. *Khrono*. Retrieved 31.10.22 from https://khrono.no/bekymring-for-ensomme-studenter-over-hele-landet/527538

Tynjälä, P., & Gijbels, D. (2012). Changing world: Changing pedagogy. In P. Tynjälä, M.-L. Stenström, & M. Saarnivaara (Eds.), *Transitions and transformations in learning and education* (pp. 205-222). Springer Netherlands. https://doi.org/10.1007/978-94-007-2312-2_13

UiB. (2020). *RETTE. System for risk and compliance. Processing of personal data in research and student projects at UiB.* UiB. Retrieved 25.08 from https://rette.app.uib.no/

UN. (2015). *The 2030 agenda for sustainable development*. UN. Retrived 21.04.22 from https://www.un.org/sustainabledevelopment/education/

unCommission. (2022). *Elevating the voices of our future innovators*. Retrieved 29.11.22 from https://theuncommission.org/stories/

UNICEF. (2021). *Adolescent education and skills*. Retrieved 21.04.22 from https://www.unicef.org/education/skills-development

van Dinther, M., Dochy, F., & Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educational Research Review, 6*(2), 95-108. https://doi.org/https://doi.org/10.1016/j.edurev.2010.10.003

Virtanen, A., & Tynjälä, P. (2019). Factors explaining the learning of generic skills: A study of university students' experiences. *Teaching in Higher Education, 24*(7), 880-894. https://doi.org/10.1080/13562517.2018.1515195

Virtanen, A., & Tynjälä, P. (2022). Pedagogical practices predicting perceived learning of social skills among university students. *International Journal of Educational Research, 111*, 101895. https://doi.org/https://doi.org/10.1016/j.ijer.2021.101895

Vygotskij, L. S. (1986). *Thought and language* (A. Kozulin, Ed.). MIT Press.

Waite, A. M., & McDonald, K. S. (2019). Exploring challenges and solutions facing STEM careers in the 21st century: A human resource development perspective. *Advances in Developing Human Resources, 21*(1), 3-15. https://doi.org/10.1177/1523422318814482

Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A delicate balance: Integrating active learning into a large lecture course. *CBE--Life Sciences Education, 7*(4), 361-367. https://doi.org/10.1187/cbe.08-02-0004

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), 1-14. https://doi.org/http://dx.doi.org/10.1186/s40594-018-0142-3

Werner, A. M., Tibubos, A. N., Mülder, L. M., Reichel, J. L., Schäfer, M., Heller, S., Pfirrmann, D., Edelmann, D., Dietz, P., Rigotti, T., & Beutel, M. E. (2021). The impact of lockdown stress and loneliness during the COVID-19 pandemic on mental health among university students in Germany. *Scientific Reports, 11*(1), 22637-22637. https://doi.org/10.1038/s41598-021-02024-5

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-ar53. https://doi.org/10.1187/cbe.18-08-0155

Winschel, G. A., Everett, R. K., Coppola, B. P., Shultz, G. V., & Lonn, S. (2015). Using jigsaw-style spectroscopy problem-solving to elucidate molecular structure through online cooperative learning. *Journal of Chemical Education, 92*(7), 1188-1193. https://doi.org/10.1021/acs.jchemed.5b00114

Woolfolk, A. (2001). *Educational psychology* (8th ed.). Allyn and Bacon.

Yapici, I. U. (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.

Papers I, II, III

Paper I

Møgelvang, A. & Nyléhn, J. (2022). Co-operative learning in undergraduate mathematics and science education: A scoping review. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-022-10331-0

Paper II

Møgelvang, A., & Nyléhn, J. (In press). Interdependence between perceived cooperative learning, sense of belonging, and generic skills in undergraduate STEM education. *Nordic Journal of STEM Education.*

Paper III

Møgelvang, A., Vandvik, V., Ellingsen, S., Strømme, C. B., & Cotner, S. (2023). Cooperative learning goes online: Teaching and learning intervention in a digital environment impacts psychosocial outcomes in biology students. *International Journal of Educational Research, 117*, 102114. https://doi.org/10.1016/j.ijer.2022.102114

Paper I

Møgelvang, A. & Nyléhn, J. (2022). Co-operative learning in undergraduate mathematics and science education: A scoping review. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-022-10331-0

International Journal of Science and Mathematics Education
https://doi.org/10.1007/s10763-022-10331-0

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

Anja Møgelvang[1] · **Jorun Nyléhn**[2]

Received: 14 June 2022 / Accepted: 24 October 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 (Ladyshewsky, 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 (**P**opulation, **I**ntervention, **C**omparison, and **O**utcome) (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: **P**opulation = Students in undergraduate mathematics, physics, chemistry, biology, and geology, **I**ntervention = Exposure to CL learning elements, **C**omparison = Not relevant in this review, and **O**utcome = 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	*(cooperati* NEAR/2 learn* OR collaborati* NEAR/2 learn* OR team* NEAR/2 learn*)*
2nd line	*("higher education" OR college* OR universit* OR postsecondary OR "post secondary" OR graduate* OR undergraduate* OR tertiary OR bachelor*")*
3rd line	*(stem OR math* OR physic* OR chemi* OR biolog* OR geolog*)*
4th line	*(Outcome* OR result* OR effect OR effects OR skill* OR competenc* OR knowledge OR achievement* OR performanc*OR benefit* OR impact*)*
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).

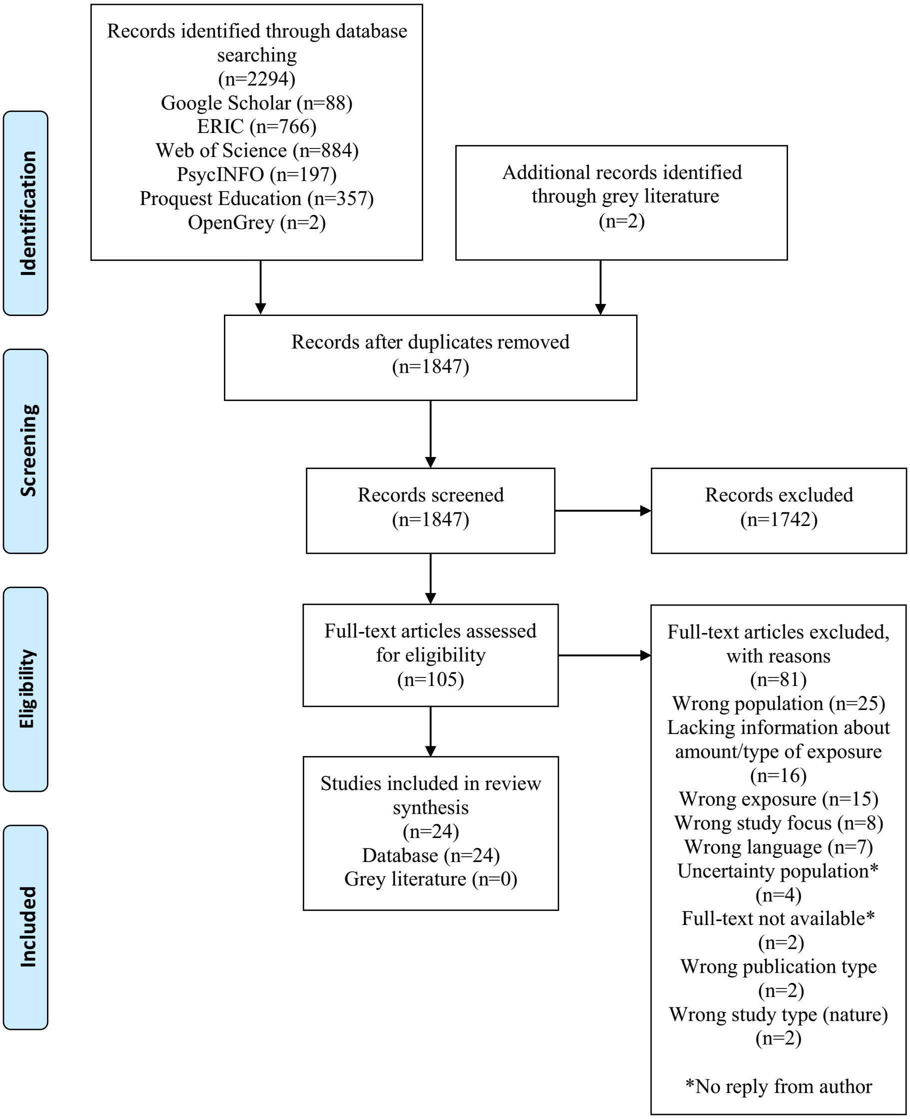


Fig. 1 PRISMA flow diagram (Moher et al., 2009)

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) Brewe 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 Team contract	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) Díaz-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) Díaz-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 knowledge Generic skills	Participation Psychological health	Content knowledge Attitudes	Academic achievement Content knowledge	Academic achievement Content knowledge	Academic achievement Content knowledge 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) Tinungki (2015)	(21) Warfa et al. (2018)	(22) Wilton et al. (2019)	(23) Yapici (2016)	(24) Yimer and Feza (2020)
Discipline	Chemistry	Biological sciences 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 recordings Group discussions Discourse analysis	Test Observation Interview	Audio recordings 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) Tinungki (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 knowledge Generic skills Attitudes	Academic achievement Generic skills Attitudes	Content knowledge	Content knowledge	Content knowledge	Academic achievement Retention Psychological health	Academic achievement Psychological health Attitudes	Content knowledge 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 (Rattanatumma & 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 predicator 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.

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
Cheruvelil, 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/ijsotl.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

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 doktorgradsutdanning 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.

Ladyshewsky, 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/ressat.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., & Elhajj, 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
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
Rattanatumma, 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.

Supplemental Material 1

Full Overview of the Component Search Strings, Fields, and Thesaurus Used in Each Database

Database	Search Field	Search String	Field Selection**	Thesaurus***
ERIC	1	((cooperati* N2 learn*) OR (collaborati* N2 learn*) OR (team* N2 learn*))	TI, AB, KW	DE = “Cooperative learning”
	2	(“higher education” OR college* OR universit* OR postsecondary OR “post secondary” OR graduate* OR undergraduate* OR tertiary OR bachelor*)	TI, AB, KW	DE = “Higher Education” “Collages” “Undergraduate “Study” “College Students” “Universities” “Postsecondary Education” “Graduates” “Bachelor’s Degrees” “Undergraduate Students”
	3	(stem OR math* OR physic* OR chemi* OR biolog* OR geolog*)	TI, AB, KW	DE = “STEM Education” “Mathematics” “Physics” “Chemistry” “Biology” “Molecular Biology” “Radiation Biology” “Marine Biology” “Biological Sciences” “Genetics” “Zoology” “Ecology” “Botany” “Geology”
	4	(Outcome* OR result* OR effect OR effects OR skill* OR competenc* OR knowledge OR achievement* OR performanc* OR benefit* OR impact*)	TI, AB, KW	DE = “Outcome Measures” “Outcomes of Education” “Test Results” “Scores” “Skills” “Thinking Skills” “Writing Skills” “Speech Skills” “Basic Skills” “Science Process Skills” “Communication Skills” “Listening Skills” “Information Skills” “Study Skills” “Research Skills” “Reading Skills” “Competence” “Academic Ability” “Knowledge Level” “Learning” “Achievement” “Academic Achievement” “Achievement Gains” “Performance” “Performance Tests” “Educational Benefits”
Proquest Education	1	(((cooperati* NEAR/2 learn*) OR (collaborati* NEAR/2 learn*) OR (team* NEAR/2 learn*)))	TI, AB, SU	MAINSUBJECT = (“Cooperative learning”)
	2	(("higher education" OR college* OR universit* OR postsecondary OR "post secondary" OR graduate* OR undergraduate* OR tertiary OR bachelor*))	TI, AB, SU	MAINSUBJECT = (“Higher education”) (“University graduates”) (“University students”) (“College students”)
	3	((stem OR math* OR physic* OR chemi* OR biolog* OR geolog*))	TI, AB, SU	MAINSUBJECT = (“STEM education”) (“Mathematics education”) (“Physics”) (“Chemistry”) (“Biology”) (“Geology”)

	4	((Outcome* OR result* OR effect OR effects OR skill* OR competenc* OR knowledge OR achievement* OR performanc* OR benefit* OR impact*))	TI, AB, SU	MAINSUBJECT = (“Effects”) (“Skills”) (“Educational Tests & measurements”) (“Academic achivement”) (“Performance evaluation”) (“Human influences”)
PsycINFO	1	((cooperati* ADJ3 learn*) OR (collaborati* ADJ3 learn*) OR (team* ADJ3 learn*))	TI, AB, ID	Subject heading = Cooperative learning/ or Collaborative learning/
	2	(“higher education” OR college* OR universit* OR postsecondary OR “post secondary” OR graduate* OR undergraduate* OR tertiary OR bachelor*)	TI, AB, ID	Subject heading = higher education/ graduate education/ undergraduate education/colleges/ graduate schools/
	3	(stem OR math* OR physic* OR chemi* OR biolog* OR geolog*)	TI, AB, ID	Subject heading = stem/mathematics/physics/chemistry/ biology/botany/ ecology/ genetics/ neurobiology/ zoology/
	4	(Outcome* OR result* OR effect OR effects OR skill* OR competenc* OR knowledge OR achievement* OR performanc* OR benefit* OR impact*)	TI, AB, ID	Subject heading = treatment process and outcome measures/ treatment outcomes/knowledge management/knowledge level/ knowledge transfer/generation effect (learning)/ learning/ cognitive processes/ memory/skill learning/ communication skills training/ social skills training/ critical thinking/competence/ ability/achievement/ performance/academic achievement/ school learning/
Web of Science	1	(cooperati* NEAR/2 learn* OR collaborati* NEAR/2 learn* OR team* NEAR/2 learn*)	TS	No thesaurus
	2	(“higher education” OR college* OR universit* OR “postsecondary” OR “post secondary” OR graduate* OR undergraduate* OR “tertiary” OR bachelor*”)	TS	No thesaurus
	3	(“stem” OR math* OR physic* OR chemi* OR biolog* OR geolog*)	TS	No thesaurus
	4	(Outcome* OR result* OR “effect” OR “effects” OR skill* OR competenc* OR “knowledge” OR “achievement*” OR “performanc*” OR benefit* OR impact*)	TS	No thesaurus

Database	Component	Search String	Field**	Thesaurus***
Google Scholar****	Swedish search string	"kooperativt lärande"/samarbetslärande/ teamlärande effekt biologi matematik fysik kemi geologi elever	(med **alle** ordene) (med **minst ett** av ordene) (**uten** ordene)	No thesaurus
	Norwegian search string	Samarbeidslæring/"kooperativ læring"/ elevteam effekt biologi matematikk fysikk kjemi geologi elever	(med **alle** ordene) (med **minst ett** av ordene) (**uten** ordene)	No thesaurus
	Danish search string	Samarbejdslæring/"kooperativ læring"/"læring i team" effekt biologi matematik fysik kemi geologi elever	(med **alle** ordene) (med **minst ett** av ordene) (**uten** ordene)	No thesaurus
OpenGrey	No separate components	("cooperative learning" OR "collaborative Learning" OR "team NEAR/2 learning") AND (Outcome* OR result* OR effect OR effects OR skill* OR competenc* OR knowledge OR achievement* OR performanc* OR benefit* OR impact*) AND (stem OR math* OR physic* OR chemi* OR biolog* OR geolog*) AND ("higher education" OR college* OR universit* OR "postsecondary" OR "post secondary" OR graduate* OR undergraduate* OR tertiary OR bachelor*)	No fields	No thesaurus

**Title (TI), Abstract (AB), Keyword (KW), Subject (SU), Key phrase identifier (ID), Topic (TS)
***Descriptors (DE), Main subject, Subject heading
**** No thesaurus and no possibility for combining Boolean operators. Therefore, only free text search and multiple searches. As Google Scholar is inexhaustive, it was decided to only perform searches in Danish, Norwegian and Swedish. English would have yielded too many results for a systematic approach.

Grey literature online source

Higher Education Academy, under HEA-Z (https://www.advance-he.ac.uk/knowledge-hub?field_content_type_target_id=3754):

- With search string: “Cooperative Learning” STEM = 0 results
- With search string: “Collaborative Learning” STEM = 2 results

Supplemental Material 2

Inclusion and Exclusion Criteria

Domain	Included	Excluded
Population	Students in global undergraduate (Bachelor) Mathematics and Science (MS) *discipline* courses. MS discipline courses in this review compromises the following subject disciplines: mathematics, physics, chemistry, biology, and geology.	Students in all other levels and types of global education and courses, including: MS courses for graduate (Master) students MS courses *primarily* for non-major MS students MS courses *primarily* for students belonging to MS subject disciplines *other than* mathematics, physics, chemistry, biology and geology (e.g. computer science). MS courses *primarily* for students in professional studies (such as pre-service teachers, engineers, nurses, doctors, pharmacists, veterinarians, etc.).
Intervention (exposure)	*Highly structured* in-class Co-operative Learning (CL) elements in groups of 3-6 people based on one or more of the guiding CL principles: positive interdependence, individual accountability, heterogeneity, social skills, group reflection and interaction, e.g., Jigsaw, Think-Pair-Share, POGIL Group Formation, Group Roles, STAD, TAI, etc.	Co-operative Learning (CL) in pairs (2 people) or in groups of 7 people or more *only*, Collaborate, Team-Based Learning method(s) and out-of-class group-work based primarily on looser structures and other principles including Peer-Led (Team) Learning, Study Groups, Peer-Assisted Learning, Flipped Classroom, Problem-Based Learning, Online/Digital/Mobile Learning, Clickers and similar Interactive Response Systems, Learning Communities, Project-Based Learning, Collaborative Writing, etc.
Study focus	Empirical examination of the relationship between CL elements and undergraduate MS student outcomes	Non-empirical examination of the relationship between CL elements and undergraduate MS student outcomes or empirical examination of relationships between other variables
Study type (nature)	Primary study	Secondary study (reviews of other studies)
Study type (method)	Both quantitative and qualitative findings	
Study information about intervention	Study includes information about the amount and type of exposure/CL elements used	Study excludes information about the amount and type of exposure/CL elements used

Study information about outcome(s)	Study includes information about the amount and type of outcome(s) the students experience	Study excludes information about the amount and type of outcome(s) the students experience
Date of study	The study is published from 2010 to January 2020	The study is published before 2010 or after January 2020
Language	English, Norwegian, Danish, Swedish	All other languages
Publication type	Peer reviewed and non-peer reviewed articles and papers, books, book chapters, and grey literature	All other types of publication, including abstracts and conference papers

Paper II

Møgelvang, A., & Nyléhn, J. (In press). Interdependence between perceived cooperative learning, sense of belonging, and generic skills in undergraduate STEM education. *Nordic Journal of STEM Education.*

Interdependence between Perceived Cooperative Learning, Sense of Belonging, and Generic Skills in Undergraduate STEM Education

A. Møgelvang[a]* and J. Nyléhn[a]

[a]*bioCEED, Centre for Excellence in Biology Education, Department of Biological Sciences, University of Bergen, Norway*

*Corresponding author:
Anja Møgelvang
BioCEED, Department of Biological Sciences
University of Bergen
Thormøhlens gate 53A
N-5006 Bergen
anja.mogelvang@gmail.com
Mobile: +47 41 16 16 41

ABSTRACT: The development of sense of belonging and generic skills may be considered important to succeed in higher education and in life and may be enhanced through student group work. For group work to succeed, Social Interdependence Theory and Cooperative Learning suggest that group members need to be positively interdependent. In the present study we conducted a cross-sectional survey in a sample of 401 students in undergraduate Science, Technology, Engineering, and Mathematics education in Norway mapping the students' perceptions of cooperative learning, sense of belonging and generic skills. By means of Pearson bivariate correlation analyses and standard multiple regression analyses, we found that 1) Cooperative Learning was positively associated with the development of both sense of belonging and generic skills, 2) Sense of belonging and generic skills were positively interrelated, and 3) Interaction was the cooperative learning principle contributing most to the association with both sense of belonging and generic skills.

1 INTRODUCTION

The importance of preparing students for work and life in a fast-paced world is a recurring topic in international position papers (OECD, 2018; UN, 2015). Among the student factors that can be developed to succeed in higher education specifically and life generally are sense of belonging (SoB) and generic skills (GS). In higher education SoB seems to be linked to student retention (Aurlien et al., 2019; Sæthre, 2014; Thomas, 2012; Tinto, 1975, 1993) and GS to employability (Cornford, 2005; Davey et al., 2018; Male et al., 2011). In life, belonging is regarded as a basic human need (Baumeister & Leary, 1995; Maslow, 1968) and SoB may be defined as our experience of being an integral part of our surrounding systems or environment (Hagerty et al., 1992, p. 173). GS may be regarded tools for lifelong learning (Bourn, 2018) and defined as holistic soft skills which operate across wide ranges of contexts (Taber, 2016, p. 226), and often predict success in life (Heckman & Kautz, 2012, p. 2).

The development of SoB and GS in educational settings seems to be related to student interaction, including group work (Allen et al., 2021; Ballantine & McCourt Larres, 2007; Kember et al., 2007; Virtanen & Tynjälä, 2019). However, groups and group tasks may need to be deliberately designed to ensure that students cooperate to fulfill their tasks. One alternative is to make the students mutually interdependent (Gillies, 2014; 2016). Positive interdependence between students in groups makes up the foundation of Cooperative Learning (CL) by Johnson and Johnson (1989). CL is developed from Social Interdependence Theory (SIT) by Deutsch (2012), which postulates that to facilitate desirable student outcomes, e.g., increased SoB and GS, it is important to structure positive interdependence between students in groups (Deutsch, 2012; Johnson & Johnson, 2005).

In this study, we examine how CL in undergraduate Science, Technology, Engineering, and Mathematics (STEM) education in Norway is related to SoB and GS respectively. The relationships are interpreted with the theoretical frameworks of SIT.

Belonging is a human need (Maslow, 1968). If our need to belong goes unfulfilled we may become lonely which in turn may cause health problems and increase mortality (Baumeister & Leary, 1995; Hayley et al., 2017; Holt-Lunstad et al., 2015; Richardson et al., 2017). Thus, according to Baumeister and Leary (1995), belonging may be regarded just as important for our health and survival as basic physical needs. We fulfil our need to belong by engaging in meaningful interpersonal relationships and social interactions (Baumeister & Leary, 1995). Hence, being a part of groups seems essential to fulfil the need for belonging, in life as in higher education.

Many initiatives may enhance SoB in educational settings. In their review, Allen et al. (2021, p. 91) propose that components such as perceptions of belonging and opportunities to belong reinforce and affect one another continuously in the development of belonging. It seems essential that universities acknowledge that students enter with different perceptions of belonging informed by past experiences. Given this acknowledgement, universities should both strive to add to an existing SoB and create new experiences that may remedy past experiences of alienation. A way to achieve this may be to create learning settings where students are given opportunities to belong on different levels, e.g., with peers, teachers, disciplines, and institution. These learning settings should be structured so that all students are enabled to fulfill the need for belonging through social interactions and meaningful relationships (Allen et al., 2021; Baumeister & Leary, 1995).

Previous research shows that highly structured groups and group work such as CL in undergraduate STEM education may provide the social interactions and meaningful relationships needed to enhance SoB (Møgelvang & Nyléhn, 2022). In a study in undergraduate mathematics, Furuto (2017) reported that implementing CL methods increased SoB among the students. In an undergraduate biology study, Wilton et al. (2019) introduced structured in-class student-student/student-teacher interactions and peer-led discussions and found that the students reported greater SoB than did students in a similar course with traditional teaching. In these studies, both from the US, CL was implemented to enhance belonging among minority groups to strengthen student diversity in higher STEM education. Taken together they show the potential for a positive association between CL and SoB in undergraduate STEM education. In Norway, student loneliness is on the rise (Knapstad et al., 2018; Sivertsen, 2021) and initiatives to enhance belonging are warranted. Thus, in this study we wish to examine if CL may also be positively related to SoB in a Norwegian undergraduate STEM population.

As the definition in the first paragraph implies, we opt for an extended understanding of GS. GS are also known as "life skills", "21st century skills", and "transferable skills" (UN, 2015; UNICEF, 2021), and according to Binkley et al. (2012, pp. 18-19) GS include a) ways of thinking: e.g. creativity, critical thinking, problem-solving, and meta cognition, b) ways of working: e.g. collaboration and communication, c) tools for working: e.g. Information and Communication Literacy (ICT) and d) living in the world: e.g. citizenship. Since GS may promote lifelong learning opportunities across a range of fields and enable students to navigate in and adapt to an unpredictable future, GS are seen as vital to the question of sustainability. Thus, GS play a substantial role in Goal 4 of *Transforming Our World: The 2030 Agenda for Sustainable Development* (UN, 2015; UNESCO, 2016).

In educational settings, it is generally believed that GS are developed through an integration of content knowledge and active learning methods, and especially collaborative learning methods seem to be a strong predicator of the development of GS (Ballantine & McCourt Larres, 2007; Kember et al., 2007; Smith & Bath, 2006; Tynjälä & Gijbels, 2012; Virtanen & Tynjälä, 2019). By integrating both learning content and collaborative learning methods, it is believed that a dual process occurs: when students use theoretical knowledge to discuss and solve practical problems, they also conceptualize their practical experiences using theoretical concepts (Virtanen & Tynjälä, 2019, p. 882). In CL literature, specifically, Millis and Cottell (1998) suggest that the inherent group and task structures of CL may stimulate the development of GS such as problem-solving and critical thinking.

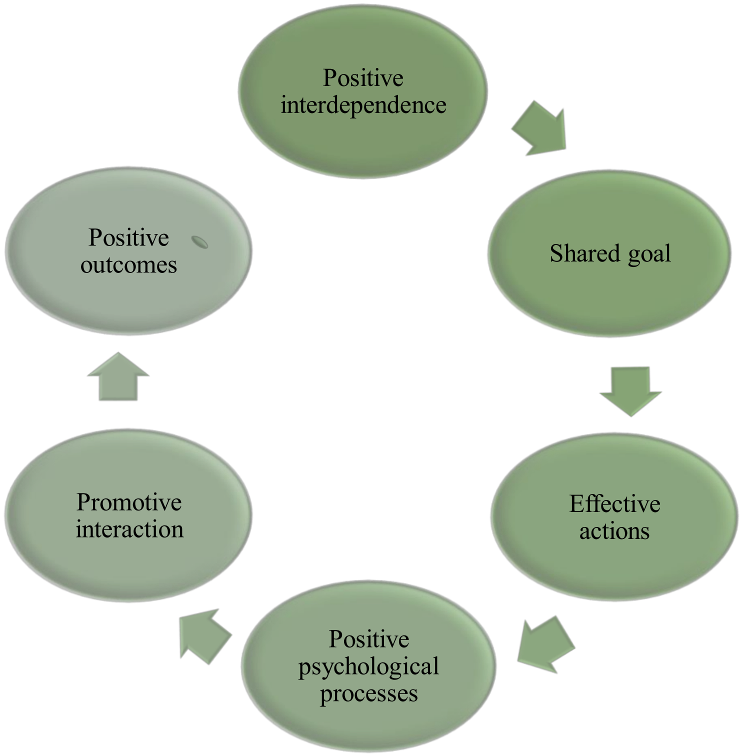
Previous research in undergraduate STEM education indicates that CL may be linked to the development of GS (Møgelvang & Nyléhn, 2022). Thus, in the US, where numerous calls to address gaps in GS have been issued, studies (Canelas et al., 2017; Carson & Glaser, 2010; Cheruvelil et al, 2020; Ott et al., 2018) have examined if CL may help close these gaps. In undergraduate chemistry, Canelas et al. (2017) compared two similar courses, one employing traditional lectures and one employing several CL methods and found that the students in the latter reported higher learning gains in key transferable skills such as problem-solving and collaboration. Similar results, at least regarding collaboration, was found in yet another undergraduate chemistry course employing CL methods (Carson & Glaser, 2010). In undergraduate biology, Cheruvelil et al. (2020) introduced CL elements such as team contracts, teamwork syllabus objective, exercises, and reflection and found that the students' collaboration skills improved significantly. In a mix of STEM disciplines, a study implemented CL roles to enhance collaborative skills. Although the students reported negative perceptions of the roles, they gained valuable collaboration skills (Ott et al., 2018). Taken together, these studies show the potential for a positive association between CL and GS in undergraduate STEM education, at least in the US. The White Paper "Working Life Relevance" (St.Meld.16, 2020-2021) is a clear testimony to similar gaps in Norway – and addresses the responsibility of higher education to address the question of GS in reducing these gaps. Thus, knowledge on teaching and learning strategies related to GS are warranted and, in this study, we examine if CL may be related to the development of GS in a Norwegian undergraduate STEM population.

2 THEORETICAL FRAMEWORK

2.1 Social Interdependence Theory (SIT)

Social Interdependence Theory (SIT) states that we are socially interdependent when our individual outcomes are influenced by other people's actions and was first introduced by Morton Deutsch in the 1940s (Deutsch, 2012). The premise of SIT is that goals, actions, psychological processes, interaction, and subsequently outcomes of individuals are dependent on how social interdependence in groups is structured. There are three ways of structuring social interdependence: positive interdependence, negative interdependence, and no interdependence (Deutsch, 2012). Negative interdependence primarily leads to negative group processes and outcomes, and no interdependence leads to no group processes or outcomes. Positive interdependence, however, is believed to lead to several positive processes and outcomes (Figure 1).

Figure 1. Positive interdependence processes in groups



According to Deutsch, *positive interdependence* arises when individuals in a group think that the only way, they can reach their own goals is if other individuals reach their goals (Deutsch, 1973, p. 20; Johnson & Johnson, 2009, p. 366). Therefore, as shown in Figure 1, when individuals within a group are positively independent and share goals, they engage in *effective actions* (e.g., orientation to task achievement and high productivity) to try to reach their *shared goal*. In the process, they are likely to experience *positive psychological processes* and Deutsch (2012, pp. 5-6) points to three psychological processes: cathexis, substitutability, and inducibility. Cathexis concerns the human innate predisposition to respond positively to stimuli that are beneficial for us and negatively to those that are harmful. Substitutability is a term which is used to describe the degree to which an individual's actions can satisfy another individual's intentions, e.g., division of labor or role specialization. Inducibility refers to the readiness to accept or reject doing what another individual wants us to do. Positive interdependence is likely to affect these three psychological processes in a positive manner and because of that the next step in Figure 1, *promotive interaction* will follow. Promotive interaction is a type of interaction where individuals encourage and ease each other's contributions. Ultimately, the entire process will lead to *positive outcomes* for the individuals in the group. These positive outcomes are characterized by a reciprocal relationship and count high efforts to achieve, positive relationships, and good mental health (Johnson & Johnson, 1989, p. 9). In our study, we relate SoB to positive relationships and good mental health and due to our extended understanding of GS, we relate GS to all three categories: high efforts to achieve, positive relationships, and good mental health.

Deutsch (1973; 2012) stressed that few situations are characterized by purely positive, negative or no interdependence. Thus, to facilitate positive student outcomes, e.g., increased SoB and GS, it may be important to create deliberate structures leading to positive interdependence. Structuring positive interdependence between students in groups is the pillar of the teaching method Cooperative Learning (CL).

2.2 Cooperative Learning (CL)

The premise of SIT has the last forty years been systematically developed into the pedagogy known as Cooperative Learning (CL) by educational psychologists and brothers David and Roger Johnson (Johnson & Johnson, 1989). CL rests on the relationship between SIT, research, and practice - and numerous studies have validated, modified, and extended the theory and the applications of CL (Johnson & Johnson, 2009). Based on SIT, CL underlines positive interdependence in cooperation. In CL, students are not simply assigned to groups and told to work together (Gillies, 2014). Thus, CL tends to be more highly structured than other forms of small-group learning (Millis & Cottell, 1998, p. 10). CL may be defined as: '...a highly structured form of group work' (Millis, 2010, p. 5) and '…the instructional use of small groups so that students work together to maximize their own and each other's learning' (Johnson et al., 1998, p. 14). For true cooperation to occur, both groups and group tasks should be structured according to a set of principles.

Johnson and Johnson (2009) operate with five principles to which CL should adhere. *Positive interdependence* is achieved by structuring the group and the group task in a way that makes group members interdependent and interested in co-working to successfully complete the task (Ballantine & McCourt Larres, 2007, p. 128). *Individual accountability* promotes responsibility and prevents social loafing (Millis & Cottell, 1998). Individual accountability is achieved when the teacher includes a mechanism, e.g., individual tests, for holding group members accountable for learning the material and completing the group task (Ballantine & McCourt Larres, 2007, p. 128). *Promotive interaction* takes place when group members encourage and ease each other's contributions through listening, exchanging ideas, offering explanations, and constructive feedback (Gillies, 2014, p. 131). According to Johnson and Johnson (1990) such reciprocal actions may also lead to group members feeling more accepted and valued. *Appropriate use of social skills* is the explicit training and negotiation of social inclusion, mutual respect, consideration, and assistance within the group (Gillies, 2016). *Group reflection* occurs in two steps: first the group members reflect on which group actions and strategies were useful and which were not and second, they decide which actions and strategies should be maintained and which need altering (Johnson & Johnson, 2009). In addition to the CL principles of Johnson and Johnson (2009), we find other principles or basic elements in CL literature and *Tutoring* is one of these. Tutoring is characterized by the teacher's involvement and support in the group task, process, and product (Atxurra et al., 2015). Teachers who plan, explain, observe, help, and offer feed-back are invaluable to student success (Hattie, 2012).

3 RESEARCH QUESTION

Most of the previous studies on the proposed relationships between CL and SoB and CL and GS in undergraduate STEM education do not provide an in-depth theoretical rationale for this association and have not been conducted in a European setting (Møgelvang & Nyléhn, 2022). To the best of our knowledge, associations between CL on the one hand and the development of SoB and GS on the other, have yet to be examined in Norwegian undergraduate STEM education.

In this study, we contribute to filling these knowledge gaps. In response to higher education challenges (Knapstad et al., 2018; Sivertsen, 2021) and numerous international and national priorities (OECD, 2018; St.Meld.16, 2020-2021; UN, 2015), we consider it important to identify teaching methods that may enhance SoB and GS in a Norwegian higher education context, and to provide possible and thorough theoretical reasons for such associations. Hence, we pose the following research question:

How is cooperative learning related to sense of belonging and generic skills among students in Norwegian undergraduate STEM education?

4 METHODS

4.1 Sample and Procedure

Data used in this study was based on a cross-sectional survey collected during lectures in the fall 2020. For the analyses, we used a sample of undergraduate STEM students at a major Norwegian university. In total 437 students from six different courses were invited to complete a survey and 401 students participated, resulting in a response rate of 92%. The students were studying the following disciplines: chemistry (n=146; 36%), biology (n=126; 31%), geology (n=92; 23%), engineering (n=22; 6%), and physics (n=15; 4%). Only large undergraduate STEM courses implementing variants of student cooperation were invited to participate. Participants consisted of 244 females (61%), 152 males (38%) and 4 students who did not report gender (1%). Age was divided into three intervals: 20 years or younger (52%), 21-24 years of age (41%), and 25 years or older (8%). Regarding the education level of their parents, 14% of the students had no parents with higher education, 35% had one parent with higher education and 51% responded that both parents had higher education.

The data collection procedures followed the regulation of GDPR[1] and the advice of Norwegian Centre for Research Data (NSD), and the study was registered in a data protection portal. The participants were informed of the purpose of the study, that their participation was voluntary, and that no personal, sensitive, nor identifiable data was collected. Each student was allowed 15 minutes to complete the survey. Due to the COVID-19 pandemic the data collection was partly digital, using SurveyXact by Rambøll (Rambøll, 2021), and partly physical, depending on course restrictions. The main researcher or research assistants were available for questions throughout the completion.

4.2 Measures

All measures were validated in a pilot survey administered to 253 STEM students at the same university during the spring 2020. The main purpose of the pilot was to validate the translation and the dimensionality of the *Psychological Sense of School Membership* (PSSM) by Goodenow (1993) and the *Cooperative Learning Application Scale* (CLAS) (Atxurra et al, 2015). Based on student comments and the statistical measures from this pilot some of the items were re-translated while others were removed. The individual measures and their validation steps are explained below.

4.2.1 Cooperative Learning (CL)

Validated scales that measure CL in higher education are rare and therefore, this study made use of a rather new scale called the *Cooperative Learning Application Scale* (CLAS) (Atxurra et al., 2015). This scale consists of 44 items distributed along seven subscales, each responding to a CL principle: Positive interdependence, Interaction, Social skills, Group reflection, Heterogeneity, Assessment, and Tutoring.

Because CLAS was a new scale with little previous validation, we took several steps to validate it in a Norwegian setting before using it in the pilot and main survey. First, we decided to remove two subscales: assessment and heterogeneity. Assessment (six items) was removed as formative assessment has not yet been implemented in all courses at Norwegian universities. Heterogeneity (four items) was removed as our survey was to be administered to large introductory courses where deficient personal knowledge of the students may make the forming of heterogenous groups a difficult task. Lacking measures of student in-coming preparation analogous to ACT or SAT scores, or GPA in previous classes in Norway adds to this difficulty (Cotner et al., 2020).

As recommended by the International Test Commission (ITC) test translation and adaptation guidelines (Hambleton, 2001), the items were translated from English to Norwegian and then back to English again by two different sets of researchers. After agreeing upon the most suitable translation, each item was discussed with a group of five STEM students to ensure that the students' understanding of the items reflected the meaning of the items. Furthermore, the translation and number of items in the scale were subject to change after student feedback and the statistical findings of the aforementioned pilot study (n=253). Ultimately, the validation process resulted in a final scale consisting of five subscales and 23 items. These five subscales each represented the following

[1] General Data Protection Regulation

principles of CL: positive interdependence, promotive interaction, social skills, group reflection and tutoring.

The items in the selected subscales included statements such as "When we work in groups, we can't fulfill a task unless everybody contributes" (Positive interdependence), "In this subject, we have the opportunity to share our opinions with group members" (Interaction) and "The lecturer guides us and helps us with our group task" (Tutoring). The items were measured on a 4-point Likert-type scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*).

4.2.2 Sense of Belonging (SoB)

To measure the students' SoB in their respective courses, we used the *Psychological Sense of School Membership* (PSSM) by Goodenow (1993). The scale has been validated in international higher education and in some but not all previous studies, the PSSM has resulted in three different subscales equivalent to a sense of social belonging (peer-related), a sense of academic belonging (tutor-related) and a general sense of belonging (institution-related). In these studies, the global scale and subscales have demonstrated good internal consistency (Alkan, 2016; Freeman et al., 2007).

The original scale consists of 18 statements, but due to poor fit we decided to remove two of the items. The PSSM underwent the same thorough validation steps as the CLAS. First, it was translated and back translated in accordance with the ITC (Hambleton, 2001) and second, discussed with a group of five STEM students to ensure that the students' understanding of the items reflected the meaning of the items. Further, based on student feedback in the pilot survey (n=253), we reformulated some of the items before including them in the present study. Examples of select items are "Other students here like me the way I am", "The teachers here respect me" and "I feel like a real part of (name of course)". All items were measured on a 5-point Likert scale from *strongly disagree* (1) to *strongly agree* (5).

4.2.3 Generic Skills (GS)

The students' perceptions of their GS in their respective courses were measured using the subscale "Generic skills" in the *Course Experience Questionnaire* (CEQ) (Ramsden, 1991). The subscale comprises six statements such as "The course helped me to develop my ability to work as a team member" and "The course sharpened my analytical skills" which were measured using a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). In previous publications, including in a Norwegian study, the reliability of the CEQ "Generic skills" subscale has been acceptable (Byrne & Flood, 2003; Espeland & Indrehus, 2003; Jansen et al., 2013) and no translation was needed.

4.3 Data Analyses

All preliminary and primary analyses were performed using IBM SPSS 25 (IBM, 2017). First, to assess factor structure, normal distribution, and internal consistency, we conducted exploratory factor analyses, descriptive analyses, and reliability analyses measured with Cronbach's alpha. Second, to explore potential relationships between CL and SoB and GS, we ran Pearson bivariate correlation analyses. Third, to determine how much unique variance each of the CL subscales explain in the prediction of SoB and GS, we conducted standard multiple regression analyses. "Exclude cases pairwise" was the chosen strategy in cases of missing data.

5 RESULTS

5.1 Exploratory Factor Analyses

Due to little previous validation of CLAS and the uncertainties regarding dimensionality of PSSM, we ran Principal component exploratory factor analyses (Tabachnick & Fidell, 2014). Because CLAS and PSSM were reported to consist of correlated factors (Atxurra et al., 2015; You et al., 2011), we used Oblimin rotation (Tabachnick & Fidell, 2014). Our exploratory factor analyses (EFAs) showed acceptable factor loadings. An EFA of CLAS resulted in a four-factor solution with eigenvalues above 1, accounting for 64.38% of the variance. The original subscales "Social skills" and "Group reflection" emerged as one factor which we, based on the items, named "Group work reflection". An EFA of PSSM produced three factors, i.e., general sense of belonging, social sense of belonging, and academic sense of belonging, with eigenvalues above 1 which in total accounted for 56.34% of the variance. An EFA of *Course Experience Questionnaire* (CEQ) (Ramsden, 1991) resulted in a one-factor solution with eigenvalues above 1, accounting for 53.01% of the variance. Full overviews of the EFA solutions and their respective factor loadings are presented in the supplementary materials.

5.2 Descriptive Statistics

As illustrated in Table 1, all variables met the assumptions of normal distribution with skewness (Skw.) and kurtosis (Kurt.) well under the absolute limit of -2 to 2 (Field, 2009). Reliability was measured with Cronbach's alpha (α) and exhibited values of around .80 which is considered good and around .90 which is considered excellent (Cronbach, 1951; Kline, 2016).

Table 1. Descriptive statistics for the study variables

	M	Range	*SD*	Skw.	Kurt.	α
Cooperative learning (CL)	2.87	1-4	0.60	-0.47	-0.05	.93
Positive interdependence	3.07	1-4	0.71	-0.76	0.11	.84
Interaction	3.19	1-4	0.68	-0.84	0.16	.84
Group work reflection*	2.42	1-4	0.80	-0.07	-0.78	.88
Tutoring	2.80	1-4	0.73	-0.42	-0.27	.87
Sense of belonging (SoB)	4.10	1-5	0.57	-0.43	-0.53	.88
General sense of belonging	3.86	1-5	0.80	-0.41	-0.51	.81
Social sense of belonging	4.24	1-5	0.62	-0.79	0.03	.80
Academic sense of belonging	4.18	1-5	0.67	-0.85	0.64	.77
Generic skills (GS)	3.62	1-5	0.71	-0.16	-0.17	.81

Note. *M* = Mean; *SD* = Standard Deviation; Skw = Skewness; Kurt. = Kurtosis;
α = Cronbach's alpha.
*"Group work reflection" is a novel combined subscale

5.3 Correlation Analyses

All the Pearson bivariate correlations between the study variables were significant at $p<.01$ as shown in Table 2. Specifically, and marked in bold, CL correlated strongly with SoB and with GS. Further, SoB was strongly correlated with GS. Lastly, the correlation between the CL subscale Interaction and SoB and GS was stronger than the correlations between the other CL subscales and SoB and GS. All effects in bold were large (r >.50) in magnitude, except for the correlation between Interaction and GS, which was medium (r >.30) in magnitude (Cohen, 2013).

Table 2. Pearson correlation matrix of the study variables

		1	2	3	4	5	6	7	8	9	10
1	Cooperative learning (CL)	-									
2	Positive interdependence	.73	-								
3	Interaction	.83	.57	-							
4	Group work reflection	.80	.34	.53	-						
5	Tutoring	.89	.51	.64	.68	-					
6	Sense of belonging (SoB)	**.56**	.41	**.57**	.39	.47	-				
7	General sense of belonging	.47	.27	.49	.38	.37	.83	-			
8	Social sense of belonging	.41	.35	.43	.22	.32	.83	.51	-		
9	Academic sense of belonging	.52	.38	.47	.37	.47	.80	.49	.54	-	
10	Generic skills (GS)	**.52**	.33	**.46**	.42	.45	**.56**	.55	.35	.46	-

Note. All correlations were significant (p<.01).

5.4 Multiple Regression Analyses

Multiple regression analyses make several assumptions about the data (Tabachnick & Fidell, 2014) and thus, preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity, and homoscedasticity. The overall regression of CL to predict SoB was statistically significant ($R^2 = .35$, $F\ (4, 36) = 48.36$, $p < .001$) and the effect size large in magnitude ($R^2 > .26$) (Cohen, 2013). As marked in bold in Table 3, the only CL subscale to significantly predict SoB was Interaction ($\beta = .40$). Group work reflection, Positive interdependence and Tutoring did not significantly predict SoB. The overall regression of CL to predict GS was statistically significant ($R^2 = .26$, $F\ (3, 36) = 43.01$, $p < .001$) and the effect size large in magnitude ($R^2 > .26$) (Cohen, 2013). The following CL subscales, in descending order, significantly predicted GS: Interaction ($\beta = .24$), Group work reflection ($\beta = 18$), and Tutoring ($\beta = .14$). Positive Interdependence did not significantly predict GS, see Table 3.

Table 3. Standard multiple regression results of the subscales of CLAS in predicting sense of belonging and generic skills

Variable	Unstandardized coefficients *B*	Unstandardized coefficients *SE*	Standardized coefficients *β*	*t*	95% CI *LL*	95% CI *UL*
			Sense of belonging			
Constant	2.43	.13		19.32	2.19	2.68
Group work reflection	.01	.01	.07	1.25	-.01	.03
Positive interdependence	.02	.01	.09	1.76	-.00	.03
Interaction	.07	.01	**.40*** **	6.68	.05	.09
Tutoring	.01	.01	.11	1.63	-.00	.03
			Generic skills			
Constant	11.25	1.02		11.09	9.26	13.25
Group work reflection	.20	.07	**.18** **	2.91	.06	.33
Positive interdependence	.07	.07	.06	.98	-.07	.20
Interaction	.30	.08	**.24*** **	3.63	.14	.46
Tutoring	.14	.07	**.14* **	2.00	.00	.28

Note. CLAS = Cooperative Learning Application Scale; CI = Confidence interval; *LL* = lower limit; *UL* = upper limit.
*$p < .05$. **$p < .01$. ***$p < .001$

6 DISCUSSION

The purpose of this study was to examine how perceived CL relates to the development of perceived SoB and GS among a sample of Norwegian students in undergraduate STEM education and three main findings emerged. First, the results suggested that CL is positively related to both SoB and GS respectively. Second, we found that SoB and GS are positively interrelated. Third, Interaction emerged as the strongest of the significant subscales of CL in the prediction of SoB and GS.

6.1 The Relationship between CL and SoB, and CL and GS

Previous studies from the US show that CL may enhance undergraduate STEM students' SoB and GS (Canelas et al., 2017; Carson & Glaser, 2010; Cheruvelil et al., 2020; Furuto, 2017; Ott et. al, 2018; Wilton et al., 2019). Although not inferring any causal relationships, our study supports a positive relationship between CL and SoB and CL and GS in a context not examined previously, i.e., in a sample of Norwegian STEM undergraduates. Further, by providing an explanation based on social interdependence theory (SIT), the study adds value to the existing literature on the relationship between CL and SoB and GS.

When explaining the findings in light of SIT, it is vital to consider one of the principal conditions for positive interdependence in groups, i.e., shared goals. As shown in Figure 1, SIT states that shared goals between students in groups will lead to positive actions, psychological processes, promotive interaction, and subsequently outcomes, such as SoB and GS in our study (Deutsch, 2012; Johnson & Johnson, 2009, p. 366). Shared goals, and the processes they may cause, could provide the conditions Allen et al. (2021) and Baumeister and Leary (1995) claim are necessary for belonging. Not only might a shared goal provide opportunities to belong, positive psychological processes and promotive interaction (Figure 1) might also facilitate positive student perceptions of belonging – and ultimately satisfy the need to belong and prevent student loneliness (Baumeister & Leary, 1995). Shared goals and the actions, processes, and interactions brought about by positive interdependence (Figure 1) may ultimately also train and strengthen the students' GS such as communication, problem-solving, analytical skills, and collaboration. Such a process would be in line with the hypothesis proposed by Millis and Cottell (1998) that the inherent structures in CL groups and group work may lead to various GS.

6.2 A Relationship between SoB and GS?

A strong positive correlation between SoB and GS was found in the present study (Table 2). Although beyond the original scope of this study, this strong correlation may be relevant for the interpretation of the results and of interest to further studies. When SoB and GS are interrelated, it complicates our understanding of the influence of CL on these variables.

Theoretically, it is likely that SoB may lead to increased GS. According to Baumeister and Leary (1995) we will strive to fulfill our need to belong through interaction, and previous research shows that GS are developed by way of student interaction (Ballantine & McCourt Larres, 2007; Kember et al., 2007; Smith & Bath, 2006; Tynjälä & Gijbels, 2012; Virtanen & Tynjälä, 2019). On the other hand, it is also likely that students who master GS at an early stage are inclined to experience an enhanced SoB. Demonstrating solid GS may lead to recognition from peers and faculty alike and is also likely to be reflected in good grades, which in turn may affect SoB positively. A third explanation may be that SoB and GS reinforce each other. This third explanation would be in accordance with SIT which states that there is a reciprocal relationship among efforts to achieve, positive relationships, and good mental health (Johnson & Johnson, 1989, p. 9). To the best of our knowledge research on the relationship between SoB and GS does not exist, and more research is needed on this topic.

6.3 The Relationship between Interaction and SoB, and Interaction and GS

Our results indicate that of the CL principles measured in this study, Interaction is the most important principle in the development of SoB and GS. Which specific CL principles are most important to student outcomes, e.g., SoB and GS, in higher STEM education has - to the best of our knowledge - not been examined previously. Thus, this study may bring new knowledge to the field. The Interaction subscale measured in our study can be considered to reflect the principle of promotive interaction (Gillies, 2014, p. 131), and in SIT promotive interaction is considered an important step leading to positive student outcomes.

The association between the CL subscale Interaction and SoB might support the theoretical assumption in SIT that promotive interaction leads to positive student relationships and good mental health. The findings may also be in compliance with the CL notion that promotive interaction is considered to bring about a feeling of personal acceptance and value among peers (Johnson & Johnson, 1990). Further, our findings may support the belief that interaction - more than other factors and principles, including in CL practices - is key in the fulfilment of our need to belong (Baumeister & Leary, 1995).

The association between the CL subscale Interaction and GS may also support the theoretical assumption in SIT that promotive interaction increases student efforts to achieve, positive relationships, and good mental health. Johnson and Johnson (2002) suggest that promotive interaction in university populations results in cognitive processes involving oral communication, problem-solving, acquisition of concepts, critical thinking and bridging past and present knowledge. We see many parallels between the cognitive processes pointed out by Johnson and Johnson (2002) and today's sought-after GS. In addition, our findings support the hypothesis by Millis and Cottell (1998), specifically that the inherent structures in CL groups and group work may lead to various GS.

Promotive interaction as one of the key CL principles may underpin many CL structures which in turn may stimulate the development of GS.

6.4 Limitations and Strengths

Our study is a cross-sectional study and thus, no causal relationships could be claimed. Although our results showed clear associations between CL, SoB, and GS it is not possible to infer any direction to these associations. Self-reported instruments and scales not previously validated in a Norwegian setting may also pose a limitation to the study. Finally, there is a possibility that the COVID-19 pandemic and a high degree of digital teaching have affected the students' perceptions of CL, SoB, and GS.

Despite these possible limitations, this study can point to many strengths. The study was conducted in a sample not previously examined and thoroughly pre-validated scales and design through systematic procedures such as student interviews, an extensive pilot study, and review of previous research. The response rate and the magnitude of the relationships may be considered a strength, and the study provides a theoretical rationale for the studied relationships using SIT. Taken together, this study fills knowledge gaps and contributes with new and valuable information about the relationship between CL and SoB and CL and GS among a sample of Norwegian students in undergraduate STEM education.

7 CONCLUSION

We find positive relationships between CL and the development of SoB and GS, but also between SoB and GS among Norwegian students in undergraduate STEM education. The CL principle contributing most to the relationship between CL and SoB and CL and GS is promotive interaction. Except for the positive interrelationship between SoB and GS, our findings seem to be consistent with previous research and all our findings may be understood through the lens of SIT and the benefits of structuring positive interdependence into student group work (Johnson & Johnson, 2009).

7.1 Implications

Our findings suggest that implementing CL in university STEM courses might be a suitable method to strengthen SoB and GS. Our study underlines the importance of promotive interaction and thus we recommend that faculty provide students opportunities to participate in groups where they can help and support each other, exchange ideas, communicate thoughts, and offer explanations and constructive feedback (Gillies, 2014, p. 131; Johnson et al., 2014). However, it must be stressed that we do not recommend merely "group work". We suggest that teachers increase their competency on which elements contribute to successful group work, to ensure that group work is structured to increase positive interdependence among students. Faculty also need to be aware of and avoid negative interdependence.

We recommend that CL structures are applied in one of the large introductory courses. Such a course can be led by an instructor that is skilled in CL, making a foundation for group work in subsequent courses. However, we recommend that all teachers who apply CL have acquired basic competencies in the method, to ensure that students perform group work in a fruitful way.

Studies examining the relationship between CL on the one hand and SoB and GS on the other in higher education are few, conducted outside Europe and do typically not offer thorough theoretical frameworks. More research examining and theoretically explaining this relationship and the relationship between SoB and GS is warranted - in Norway, in other countries and in different types of higher education disciplines.

8 DATA AVAILABILITY

The dataset of the current study is available from the corresponding author on reasonable request.

9 DECLARATION OF CONFLICTING INTERESTS

The authors have no conflicts of interest to disclose.

10 FUNDING

The authors received no financial support for the research, authorship, and/or publication of this article.

11 ACKNOWLEDGEMENTS

The authors thank Anne Grete Danielsen from the Department of Education at the University of Bergen, Norway for valuable help throughout the study process.

REFERENCES

Alkan, N. (2016). Psychological sense of university membership: An adaptation study of the PSSM scale for Turkish university students. *Journal of Psychology, 150*(4), 431-449. https://doi.org/10.1080/00223980.2015.1087373

Allen, K.-A., Kern, M. L., Rozek, C. S., McInerney, D. M., & Slavich, G. M. (2021). Belonging: A review of conceptual issues, an integrative framework, and directions for future research. *Australian Journal of Psychology, 73*(1), 87-102. https://doi.org/10.1080/00049530.2021.1883409

Atxurra, C., Villardón-Gallego, L., & Calvete, E. (2015). Design and validation of the cooperative learning application scale (CLAS). *Revista de Psicodidactica, 20*(2), 339-357. https://doi.org/https://doi.org/10.1387/RevPsicodidact.11917

Aurlien, I., Bjørnsrud, K. M., Haugom, E., & Thrane, C. (2019). Betydningen av målbevissthet og sosial tilhørighet for frafall i høyere utdanning [The importance of goal determination and social belonging for drop out in higher education]. *Uniped, 42*(2), 125-138. https://doi.org/10.18261/issn.1893-8981-2019-02-02

Ballantine, J., & McCourt Larres, P. (2007). Cooperative learning: A pedagogy to improve students' generic skills? *Education & Training, 49*(2), 126-137. https://doi.org/10.1108/00400910710739487

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497-529. https://doi.org/10.1037/0033-2909.117.3.497

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 https://doi.org/10.1007/978-94-007-2324-5_2

Bourn, D. (2018). *Understanding global skills for 21st century professions*. Springer International Publishing. https://doi.org/10.1007/978-3-319-97655-6_3

Byrne, M., & Flood, B. (2003). Assessing the teaching quality of accounting programmes: An evaluation of the Course Experience Questionnaire. *Assessment and Evaluation in Higher Education, 28*(2), 135-145. https://doi.org/10.1080/02602930301668

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

Cheruvelil, 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, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic Press.

Cornford, I. (2005). Challenging current policies and policy makers' thinking on generic skills. *Journal of Vocational Education & Training, 57*(1), 25-45. https://doi.org/10.1080/13636820500200273

Cotner, S., Jeno, L. M., Walker, J. D., Jørgensen, C., & Vandvik, V. (2020). Gender gaps in the performance of Norwegian biology students: The roles of test anxiety and science confidence. *International Journal of STEM Education, 7*(1), 1-10. https://doi.org/10.1186/s40594-020-00252-1

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. https://doi.org/10.1007/BF02310555

Davey, T., Meerman, A., Galán-Muros, V., Orazbayeva, B., & Baaken, T. (2018). *The state of university-business cooperation in Europe (Final report).* https://www.ub-cooperation.eu/

Deutsch, M. (1973). *The resolution of conflict*. Yale University Press.

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* (Vol. 2, pp. 275). SAGE Publications Ltd. https://doi.org/10.4135/9781446249222.n40

Espeland, V., & Indrehus, O. (2003). Evaluation of students' satisfaction with nursing education in Norway. *Journal of Advanced Nursing, 42*(3), 226-236. https://doi.org/10.1046/j.1365-2648.2003.02611.x

Field, A. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)* (3rd ed.). SAGE.

Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education, 75*(3), 203-220. https://doi.org/10.3200/JEXE.75.3.203-220

Furuto, M. A. (2017). Mathematics pedagogical strategies to create a positive college classroom community. *Malaysian Journal of Mathematical Sciences, 11* (1), 9-21.

Gillies, R. M. (2014). Developments in cooperative learning: review of research. *Anales de Psicología 30*(3), 792-801. https://doi.org/10.6018/analesps.30.3.201191

Gillies, R. M. (2016). Cooperative learning: review of research and practice. *Australian Journal of Teacher Education, 41*(3), 39-54.

Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in Schools, 30*(1), 79-90.

Hagerty, B. M. K., Lynch-Sauer, J., Patusky, K. L., Bouwsema, M., & Collier, P. (1992). Sense of belonging: A vital mental health concept. *Archives of Psychiatric Nursing, 6*(3), 172-177. https://doi.org/10.1016/0883-9417(92)90028-H

Hambleton, R. K. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment, 17*(3), 164-172. https://doi.org/10.1027//1015-5759.17.3.164

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Hayley, A. C., Downey, L. A., Stough, C., Sivertsen, B., Knapstad, M., & Øverland, S. (2017). Social and emotional loneliness and self-reported difficulty initiating and maintaining sleep (DIMS) in a sample of Norwegian university students. *Scandinavian Journal of Psychology, 58*(1), 91-99. https://doi.org/10.1111/sjop.12343

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics, 19*(4), 451-464. https://doi.org/10.1016/j.labeco.2012.05.014

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: A meta-analytic review. *Perspectives on Psychological Science, 10*(2), 227-237. https://doi.org/10.1177/1745691614568352

IBM. (2017). *IBM SPSS statistics for Macintosh, version 25.0.* In IBM Corp.

Jansen, E., van der Meer, J., & Fokkens-Bruinsma, M. (2013). Validation and use of the CEQ in The Netherlands. *Quality Assurance in Education, 21*(4), 330-343. https://doi.org/10.1108/QAE-11-2012-0041

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction Book Company.

Johnson, D. W., & Johnson, R. T. (1990). Cooperative learning and achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 23-37). Praeger.

Johnson, D. W., & Johnson, R. T. (2002). Social interdependence theory and university instruction - theory into practice. *Swiss Journal of Psychology, 61*(3), 119-129. https://doi.org/10.1024//1421-0185.61.3.119

Johnson, D. W., & Johnson, R. T. (2005). New developments in social interdependence theory. *Genetic Social and General Psychology Monographs, 131*(4), 285-358. https://doi.org/10.3200/MONO.131.4.285-358

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365-379. https://doi.org/10.3102/0013189X09339057

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). *Active learning: Cooperation in the college 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.

Kember, D., Leung, D. Y. P., & Ma, R. S. F. (2007). Characterizing learning environments capable of nurturing generic capabilities in higher education. *Research in Higher Education, 48*(5), 609-632. https://doi.org/10.1007/s11162-006-9037-0

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guildford.

Knapstad, M., Heradstveit, O., & Sivertsen, B. (2018). *Students' Health and Wellbeing Study 2018*. SiO. https://www.studenthelse.no/tema/sosiale_forhold

Male, S. A., Bush, M. B., & Chapman, E. S. (2011). An Australian study of generic competencies required by engineers. *European Journal of Engineering Education, 36*(2), 151-163. https://doi.org/10.1080/03043797.2011.569703

Maslow, A. H. (1968). *Toward a psychology of being* (2nd ed.). D. Van Nostrand.

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.

Møgelvang, A., Nyléhn, J. (2022). Co-operative learning in undergraduate mathematics and science education: A scoping review. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-022-10331-0

OECD. (2018). *The Future of Education and Skills 2030.* OECD. https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf

Ott, L. E., Kephart, K., Stolle-McAllister, K. & LaCourse, W. R. (2018). 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

Rambøll. (2021). *SurveyXact*. Retrieved 19.10.2021 from https://www.surveyxact.no

Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education, 16*(2), 129-150. https://doi.org/10.1080/03075079112331382944

Richardson, T., Elliott, P., & Roberts, R. (2017). Relationship between loneliness and mental health in students. *Journal of public mental health, 16*(2), 48-54. https://doi.org/10.1108/JPMH-03-2016-0013

Sivertsen, B. (2021). *Students' Health and Wellbeing Study 2021*. SiO. https://sioshotstorage.blob.core.windows.net/shot2018/SHOT2021.pdf

Smith, C., & Bath, D. (2006). The role of the learning community in the development of discipline knowledge and generic graduate outcomes. *Higher Education, 51*(2), 259-286. https://doi.org/10.1007/s10734-004-6389-2

St.Meld.16. (2020-2021). *Utdanning for omstilling - Økt arbeidslivsrelevans i høyere utdanning [Education for change - Increased work relevance in higher education]*. https://www.regjeringen.no/no/dokumenter/meld.-st.-16-20202021/id2838171/?ch=1

Sæthre, H. Å. (2014). *Å tilrettelegge for at studentene skal lykkes: en verktøykasse for utdanningsledere, studieadministrasjon, studenttillitsvalgte [Facilitating student succes: A toolbox for educational leaders, study administration, student representatives]*. Haas1.

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th, new international ed.). Pearson.

Taber, K. S. (2016). Learning generic skills through chemistry education. *Chemistry Education Research and Practice Journal, 17*, 225-228.

Thomas, L. (2012). *Building student engagement and belonging in higher education at a time of change.* P. H. Foundation. https://www.heacademy.ac.uk/system/files/what_works_final_report_0.pdf

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research ,45*(1), 89-125. https://doi.org/10.3102/00346543045001089

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). The University of Chicago Press.

Tynjälä, P., & Gijbels, D. (2012). Changing world: Changing pedagogy. In P. Tynjälä, M.-L. Stenström, & M. Saarnivaara (Eds.), *Transitions and transformations in learning and education* (pp. 205-222). Springer Netherlands. https://doi.org/10.1007/978-94-007-2312-2_13

UN. (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. Retrieved 20.04.22 from https://sdg4education2030.org/the-goal

UNESCO. (2016). *Unpacking Sustainable Development Goal 4: Education 2030; guide*. Retrieved 15.10.2022 from https://unesdoc.unesco.org/ark:/48223/pf0000246300

UNICEF. (2021). *Adolescent education and skills*. Retrieved 21.04.22 from https://www.unicef.org/education/skills-development

Virtanen, A., & Tynjälä, P. (2019). Factors explaining the learning of generic skills: A study of university students' experiences. *Teaching in Higher Education, 24*(7), 880-894. https://doi.org/10.1080/13562517.2018.1515195

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-ar53. https://doi.org/10.1187/cbe.18-08-0155

You, S., Ritchey, K. M., Furlong, M. J., Shochet, I., & Boman, P. (2011). Examination of the latent structure of the Psychological Sense of School Membership scale. *Journal of Psychoeducational Assessment, 29*(3), 225-237. https://doi.org/10.1177/0734282910379968

Paper II, Supplementary Materials

Table 1. Exploratory factor analysis (Principal Component, Oblimin Rotation) of CLAS

Item	Group work reflection	Positive interdependence	Interaction	Tutoring
The lecturer provides us with guidelines to solve conflicts that might arise in the group	.90			
The lecturer helps us identify and define the difficulties in group work	.89			
The lecturer teaches us to properly manage in-group situations	.88			
During the lesson we have time to reflect on our ways of working in the group and how to improve	.73			
The lecturer provides us with tools for us to reflect on how we are working in the group	.52			
When we work in groups, we can't fulfill a task unless everybody contributes		.86		
When we work in groups, we need the ideas of all of us in order to achieve success		.81		
When we work in groups, each member has a task to contribute to		.75		
When we work in groups, we have to share materials or information in order to complete the task		.70		
In this subject, each group member has to make an effort in order to help the group achieve their results		.66		
This subject allows me to interact with my fellow group members			-.83	
In this subject, we have the opportunity to share our opinions with group members			-.80	
This subject creates opportunities for us to interact with others			-.79	
This subject encourages us to freely express our points of view			-.50	
In this subject we exercise our social skills			-.45	
The lecturer monitors the tasks fulfilled by the group and each of the group members, and helps us improve				-.88
The lecturer oversees group work as we carry it out				-.77
The lecturer guides us and helps us with our group task				-.76
The lecturer motivates us to make progress in our group task				-.61
The lecturer intervenes when we need it in order to ensure that we make progress in our group task				-.55
The lecturer helps us determine the level of efficiency at which the group has performed				-.52

Note. CLAS = Cooperative Learning Application Scale.
"Group work reflection" is a novel combined subscale made up of items from the original subscales "Social skills" and "Group reflection".

Table 2. Exploratory factor analysis (Principal Component, Oblimin Rotation) of PSSM

Item	Social sense of belonging	General sense of belonging	Academic sense of belonging
I can really be myself *(at this school)	.80		
Other students here like me the way I am	.79		
It is hard for people like me to be accepted here**	.79		
People *(at this school) are friendly to me	.75		
Other students *(in this school) take my opinions seriously	.59		
I feel very different from most other students here**	.35		
I feel like a real part of *(name of school)		.86	
Sometimes I feel as if I don't belong here**		.82	
I wish I were *(in a different school) **		.79	
I feel proud of belonging to *(name of school)		.57	
I am included in lots of activities at *(name of school)		.40	.31
Teachers here are not interested in people like me**			.82
Most teachers at *(name of school) are interested in me			.78
There's at least one teacher or other adult *(in this school) I can talk to if I have a problem			.68
The teachers here respect me			.68
I am treated with as much respect as other students			.62

Note. PSSM = Psychological Sense of School Membership Scale.
*() translated to *in this course* to fit the university setting.
**reversed.

Table 3. Exploratory factor analysis (Principal Component) of CEQ

Item	Generic skills
As a result of the course, I feel confident about tackling unfamiliar problems	.82
The course sharpened my analytical skills	.79
The course developed my problem-solving skills	.79
The course helped me to develop my ability to plan my own work	.69
The course improved my skills in written communication	.66
The course helped me to develop my ability to work as a team member	.60

Note. CEQ = Course Experience Questionnaire.

Paper III

Møgelvang, A., Vandvik, V., Ellingsen, S., Strømme, C. B., & Cotner, S. (2023). Cooperative learning goes online: Teaching and learning intervention in a digital environment impacts psychosocial outcomes in biology students. *International Journal of Educational Research, 117*, 102114. https://doi.org/10.1016/j.ijer.2022.102114

International Journal of Educational Research 117 (2023) 102114

Contents lists available at ScienceDirect

International Journal of Educational Research

journal homepage: www.elsevier.com/locate/ijedures

Cooperative learning goes online: teaching and learning intervention in a digital environment impacts psychosocial outcomes in biology students

Anja Møgelvang [a,*], Vigdis Vandvik [b], Ståle Ellingsen [b], Christian Bianchi Strømme [a], Sehoya Cotner [a]

[a] bioCEED, - Centre for Excellence in Biology Education, Department of Biological Sciences, University of Bergen, 5006 Bergen, Norway
[b] Department of Biological Sciences, University of Bergen, 5006 Bergen, Norway,

A R T I C L E I N F O

Keywords:
Online cooperative learning
Sense of belonging
Science confidence
Perceived generic skills
Loneliness
Intervention study

A B S T R A C T

Identifying evidence-based teaching and learning strategies that can ease teacher challenges and mitigate student concerns in digital settings becomes increasingly important. In this intervention study we compared the effect of digital cooperative learning (CL) and digital lectures on a range of psychosocial outcomes, specifically students' sense of belonging, science confidence, perceived generic skills, and loneliness, among a Norwegian sample of undergraduate biology students ($n = 71$). Employing a one-group pretest/posttest quasi-experimental design with a double pretest and follow-up, we found that students' scores on psychosocial outcomes improved significantly following digital CL compared to digital lectures. Further, the effect sizes suggest that the effect of CL on psychosocial outcomes in digital settings is at least as substantial as in physical settings.

1. Introduction

The need for digital teaching and learning strategies that can lead to desirable student outcomes in higher education has received considerable attention, and the COVID-19 pandemic has only spurred this need (Damşa et al., 2015; European Commision, 2021; Henrie et al., 2015; Lillejord et al., 2018; Lashley et al., 2020a). Following imposed restrictions on social contact, teachers all over the world moved their courses from physical classrooms to online platforms almost overnight (Crawford et al., 2020). For many of the teachers this transition led to both frustration and difficulties in facilitating a digital learning environment (Houlden & Veletsianos, 2020; Mazur et al., 2021a; Watermeyer et al., 2021). The transition to digital learning was also challenging for students, who reported lack of social contact, increased loneliness, and difficulties in studying from home (Børve et al., 2021; Deng et al., 2021; Lederer et al., 2021; Phillips et al., 2022; Sivertsen, 2021; Tice et al., 2021; Werner et al., 2021). Even though it is difficult to determine whether these student concerns were mainly due to digital learning, or the restrictions of the pandemic in general, they deserve closer attention. Digital teaching may affect students differently than traditional in-person teaching. For example, the inherent distance embedded in digital teaching and the reported student concerns indicate that psychosocial outcomes are particularly vulnerable to a digital setting.

Despite these teacher and student challenges, initiatives to implement digital learning in higher education have only gained strength in the wake of the pandemic (European Commision, 2021). While specialized online courses have been offered for years,

* Corresponding author.
E-mail address: anja.mogelvang@gmail.com (A. Møgelvang).

https://doi.org/10.1016/j.ijer.2022.102114
Received 14 September 2022; Received in revised form 25 November 2022; Accepted 30 November 2022

remote instruction is likely to remain part of the standard curricula in higher education (European Commision, 2021; Hodges et al., 2020). Thus, in this study we sought an evidence-based and student-centered teaching and learning strategy that could be implemented in a digital setting, effectively facilitate learning, and mitigate student psychosocial concerns, specifically students' sense of belonging, science confidence, perceived generic skills, and loneliness.

Previous studies from non-digital higher education settings show that active and cooperative learning (CL) methods enhance student learning and wellbeing more so than lectures (Deslauriers et al., 2019; Freeman et al., 2014; Loh & Ang, 2020). While these methods have traditionally been applied in physical classrooms, it is hypothesized that these findings may be transferable to a synchronous online context (Davidson, 2021). To test this hypothesis, we conducted an intervention study in a sample of undergraduate biology students (n = 71) in Norway, employing a one-group pretest/posttest quasi-experimental design with a double pretest and follow-up (Shadish et al., 2002). During the intervention, we implemented several evidence-based CL methods in a digital setting and explored if such methods were associated with any change in the students' psychosocial outcomes, i.e., students' expressed sense of belonging, science confidence, perceived generic skills, and loneliness compared to traditional digital lectures.

2. Cooperative learning (CL)

Cooperative learning (CL) was developed by Johnson and Johnson (1989), among others, and may be defined as: '...a highly structured form of group work' (Millis, 2010, p. 5) and '…the instructional use of small groups so that students work together to maximize their own and each other's learning' (Johnson et al., 1998, p. 14). According to Millis and Cottell (1998), key principles of CL in higher education include (i) positive interdependence, i.e., that the group members need to cooperate to complete the tasks, and have mutual gain and shared goals, and (ii) individual accountability, i.e., that each group member is accountable, thus preventing social loafing.

The principles of positive interdependence and individual accountability should underpin both CL group features and structures (Gillies, 2016; Millis & Cottell, 1998) and to achieve this, CL literature suggests, in compliance with the definition, that groups need to be highly structured (Gillies, 2003). Thus, groups should be relatively small, of three to five students, to avoid free-riding behaviour, allow less forthright students to participate, and more generally enhance learning (Kagan, 2021; Millis & Cottell, 1998). Also, groups should be formed by the teacher based on student characteristics such as academic ability, background, age, and gender, with the aim of creating diverse groups (Johnson et al., 1994; Kagan, 2021; Millis & Cottell, 1998). This will enforce mutual dependence, perspective taking, and case argumentation (Piaget, 1985), the knowledge and autonomy level of both low-ability and higher-ability students (Jacobs et al., 2006; Lou et al., 1996), and group discussion quality and group performance (Curşeu et al., 2018). Further, the use of CL structures helps the students to complete the assignment in alignment with the CL principles (Millis & Cottell, 1998). Because CL structures are content-free strategies (Kagan, 2021), they can be employed in any subject and on any educational level, including higher education. Key CL structures used in the present study are presented in Box 1 and 2.

2.1. CL outcomes

CL outcomes are traditionally divided into three categories (Johnson & Johnson, 2018): 1) efforts to achieve (e.g., academic motivation, persistence, productivity, and performance, 2) relationships (e.g., social skills, promoting each other's success, and forming academic and personal relationships), and 3) mental health (e.g., personal ego-strength, self-confidence, and autonomy), and there is a reciprocal relationship between these three categorical outcomes (Johnson & Johnson, 1989, p. 9). In higher education most CL studies have been related to the first of these three categories and in particular to higher student academic performance (Loh & Ang, 2020). The relationship between CL and academic performance is thoroughly established and the greatest impacts on students during the digital transformation induced by the pandemic have been psychosocial (Børve et al., 2021; Deng et al., 2021; Händel et al., 2020; Lederer et al., 2021; Phillips et al., 2022; Sivertsen, 2021). Therefore, this study has focused on student *affect* in relation to digital CL methods. Previous studies on CL in Science, Technology, Engineering, and Mathematics (STEM) higher education indicate that CL in physical learning environments may be positively associated with a range of psychosocial outcomes (Canelas et al., 2017; Espinosa et al., 2019; Furuto, 2013, 2017; Kocak, 2008; Pilcher et al., 2015; Rattanatumma & Puncreobutr, 2016; Rivera, 2013; Wilton et al., 2019; Yapici, 2016). However, to the best of our knowledge these studies are few, have been conducted in educational settings very different from ours, or have not explored CL methods in a fully digital setting. Thus, in this intervention study we implemented several CL methods digitally and examined the following psychosocial outcomes: students' expressed sense of belonging, science confidence, perceived generic skills, and loneliness.

To belong is regarded as a basic human need (Baumeister & Leary, 1995; Maslow, 1968) and **sense of belonging** is defined as our experience of being an integral part of our surrounding systems or environment (Hagerty et al., 1992, p. 173). One such system may be higher education; for example, previous research shows that sense of belonging seems to be particularly important for student retention in STEM higher education (Appleton et al., 2006; Rainey et al., 2018, 2019; Smith et al., 2013; Trujillo & Tanner, 2014). We fulfil our need to belong by engaging in meaningful interpersonal relationships and social interactions (Baumeister & Leary, 1995) and as a result, groups and group work facilitating student interaction in higher education may lead to increased sense of belonging. Thus, sense of belonging has been examined as one of many student outcomes of CL, including in undergraduate STEM education (Furuto, 2017; Wilton et al., 2019; Yapici, 2016), but never in a fully digital setting. Further, perceptions of belonging and opportunities to belong, among others, reinforce each other continuously in the development of belonging (Allen et al., 2021). Hence, universities should implement intentional, consistent, and systematic practices to become *places of belonging* (Murdock-Perriera et al., 2019), especially now as students' sense of belonging has decreased during the pandemic (Lederer et al., 2021; Sivertsen, 2021; Tice et al., 2021). In an attempt to increase and examine the students' sense of belonging in a digital setting, we include sense of belonging in the present study.

Science confidence is a malleable trait (Burns et al., 2016). In this study, we understand science confidence as a student's perception of their own abilities *to learn and complete* tasks specific to science, such as articulating a testable hypothesis, designing an experiment, or explaining scientific concepts to peers (Cotner et al., 2020b; Seymour et al., 2004). Thus, our understanding aligns with literature in the discipline-based educational research communities stating that science confidence appears closely related, but not identical to, *(science) self-efficacy* (Ainscough et al., 2016; Ballen et al., 2017a; Cotner et al., 2020b; Nissen & Shemwell, 2016; Rittmayer & Beier, 2008). Self-efficacy is the belief in one's capacity to organize and execute the courses of action required to manage prospective situations (Bandura, 1997) and consequently *scientific self-efficacy* may capture students' belief that they are capable enough to master scientific tasks and obtain successful results. Like science confidence, self-efficacy may also be considered malleable. In fact, Camfield et al., (2021) recently demonstrated how STEM students' efficacy beliefs were positively influenced by student-centered interventions during COVID-19 remote instruction. In previous undergraduate STEM education studies on CL interventions specifically, scientific self-efficacy has also improved (Espinosa et al., 2019; Furuto, 2013, 2017; Rivera, 2013). Thus, we consider it likely that task-specific science confidence may be a positive student outcome of CL. Generally, confidence and self-efficacy are positively associated with desirable student outcomes such as better performance and retention in a discipline (Ballen et al., 2017a; Rittmayer & Beier, 2008). These associations may be understood through the lens of social cognitive career theory (Bandura, 1986; Lent et al., 1994) - whereby a perceived reduced capacity, and possibly a consequent *lack of belonging*, in a discipline informs an individual's self-evaluation and sense of a future in that discipline (Stake & Mares, 2005; Wonch Hill et al., 2017). As student sense of belonging has decreased during the pandemic (Lederer et al., 2021; Sivertsen, 2021; Tice et al., 2021), we hypothesize that science confidence may be affected, too. Further, science confidence has, to the best of our knowledge, not been examined as one of the typical student outcomes of CL in science higher education - and certainly not in Norwegian higher education. Thus, the present study examines the students' expressed science confidence.

Generic skills are also known as "21st century skills", "life skills", and "transferable skills", (UN, 2015; UNICEF, 2021) and may be understood as holistic skills which operate across wide ranges of contexts (Taber, 2016, p. 226), and often predict success in life (Heckman & Kautz, 2012, p. 2). Such skills are developed through interaction with our surroundings and are important tools for lifelong learning (Bourn, 2018), highly desired by employers (Davey et al., 2018; Male et al., 2011), and believed essential to navigate in an unpredictable future. This broad understanding of generic or 21st century skills together with the focus on and measurement of the students' *perceived* generic skills, lay the foundation for categorizing generic skills as a psychosocial outcome in the present study. Generic skills (or 21st century skills) include a) ways of thinking: e.g., creativity, critical thinking, problem solving, and meta cognition, b) ways of working: e.g., collaboration and communication, c) tools for working: e.g., information and communication literacy, and d) living in the world: e.g., citizenship and life and career (Binkley et al., 2012, pp. 18-19). Generic skills seem to be developed through an integration of content knowledge and active learning methods, particularly group work (Ballantine & McCourt Larres, 2007; Kember et al., 2007; Smith & Bath, 2006; Tynjälä & Gijbels, 2012; Virtanen & Tynjälä, 2019; 2022) and as a result, improved generic skills may be one of the more prominent student outcomes of CL (Johnson & Johnson, 2016; Millis & Cottell, 1998). In previous studies in STEM higher education, the positive effect of CL on generic skills and/or perceived generic skills has been examined in both physical settings (Canelas et al., 2017; Pilcher et al., 2015; Rattanatumma & Puncreobutr, 2016) and digital settings (Lee et al., 2016; Parsazadeh et al., 2018). However, these studies are few and have never been conducted in a fully digital setting in STEM higher education similar to ours. Thus, we include perceived generic skills as one of our outcome measures.

Loneliness is a subjective feeling of distress due to deficiencies in an individual's social relationships (Peplau & Perlman, 1982, p. 3). Moving away from the safety of home and suddenly having to manage on their own, university students may be particularly vulnerable to loneliness (Cutrona, 1982; Stewart-Brown et al., 2000). Student loneliness is on the rise (Knapstad et al., 2018) and has increased substantially during the COVID-19 pandemic (Phillips et al., 2022; Sivertsen, 2021; Werner et al., 2021). Furthermore, loneliness among university students may cause health problems (Hayley et al., 2017; Richardson et al., 2017). To prevent and counteract loneliness, universities may facilitate initiatives (both within and beyond the classroom, physically and digitally) to decrease loneliness (Adriansen & Madsen, 2012), and increasing opportunities for social interaction is particularly promising (Hawkley & Cacioppo, 2010). CL is an example of an initiative that may offer such opportunities and previous research indicates that CL may lead to psychological health benefits (Johnson et al., 2014), including reducing feelings of loneliness in university populations (Kocak, 2008). In this study, we test if CL may also reduce students' self-reported loneliness in a digital setting.

3. Hypothesis and prediction

Increased digital components in standard university courses, and post-pandemic student concerns in higher education in Norway and elsewhere indicate a need to identify digital learning strategies which strengthen student outcomes. Cooperative learning (CL) offers a coherent and evidence-based framework for such methods but has not previously been systematically implemented or tested in a digital undergraduate STEM setting similar to ours. We capitalize on a rare opportunity offered by an undergraduate course in biology during the COVID-19 pandemic, where we implemented an experimental digital CL period into a course otherwise consisting of remote lectures and labs. Specifically, we consider the course in three parts, a lecture period, a cooperative learning experimental period, and a second lecture period, to address the following hypothesis and specific prediction:

Hypothesis 1. (H1): *Digital cooperative learning methods lead to beneficial changes in biology students' self-reported psychosocial outcomes compared to traditional digital lectures.*

Specifically, we predict that a digital cooperative learning intervention will increase the students' self-reported sense of belonging, science confidence, and perceived generic skills, and lead to a decrease in loneliness scores.

4. Methods

4.1. Participants

The participants in this intervention study all attended an undergraduate introductory biology course and thus made up a convenient sample. The course counted 86 students, but we only included the 71 students who participated in all four data collections throughout the semester, resulting in a response rate of 83%. This rate is high, especially given the amount of measurement time points (Yu et al., 2017). Participants included 39 females (55%) and 32 males (45%). Age was divided into three intervals: born in year 2000 or later (17%), born in year 1998 or 1999 (53%) and born in year 1997 or earlier (30%). 10% of the students stated that neither of their parents had higher education, 24% had one parent with higher education and 66% responded that both parents had higher education.

4.2. Procedure: Data collection

The intervention was conducted during the spring semester 2021 and all data used in this study were collected by means of four digital surveys conducted during that period. The content of the four surveys was identical but the framing differed from the first to the remaining surveys. As the first survey was administered in the first lecture of the semester, the students were asked to give their answers drawing on experience from similar large mandatory undergraduate biology courses in the *preceding semester* and for most students, in the same program. The three remaining surveys were administered at different time points and equally distributed throughout the semester; here, the students were asked to give their answers based on their experiences in the *preceding month* (Fig. 1). Each student was allowed 15 minutes to complete each of the four surveys. The data collection was digital, using SurveyXact by Rambøll (Rambøll, 2021) and the main researcher was available for questions throughout survey completion.

The data collection procedures followed the regulation of the General Data Protection Regulation (European Commision, 2022) and was approved by the Norwegian Centre for Research Data (Norwegian Centre for Research Data (NSD), 2021). Due to the psychological variable "loneliness" the Regional Committee for Medical and Health Research Ethics (Regional Committees for Medical and Health Research Ethics (REK), 2020) was also consulted to ensure that the data collection adhered to the health research ethics regulations. The participants were informed of the purpose of the study, that their participation was voluntary and that they could withdraw from the study at any time. Furthermore, they were informed that any personal data, specifically the student ID-number which we collected to trace the responses of each student, would be de-identified, treated confidentially, and deleted after the completion of the study.

4.3. Procedure: Study design

The intervention study was carried out in a large mandatory undergraduate biology course at a Norwegian university. It lasted one semester and included the following remote elements: five weeks of lectures, five weeks of cooperative learning (CL), a week of lab work, and an additional five weeks of lectures. The lectures took place on Zoom, with the CL sessions in Zoom break-out rooms. To carry out the intervention and measure possible changes in the students' psychosocial outcomes, we adopted a quasi-experimental design (Shadish et al., 2002). Quasi-experimental studies are particularly useful when randomization is not feasible due to natural criteria, e.g., class enrollment (Crano et al., 2015; Shadish et al., 2002). Given the fact that we lacked an equivalent control group, we used a one-group pretest/posttest design with a double pretest and follow-up. Using a double pretest and follow-up reduces the risk of errors due to student maturation and regression (Shadish et al., 2002) and a minimum of five weeks between every measurement reduces the risk of survey fatigue (Creswell, 2014). To illustrate the flow of the study, we refer to Figure 1.

This study – in addition to the double pre-test and follow-up measurements - adopted a range of design controls (Shadish et al., 2002) such as: a) Two long, uninterrupted, and consistently structured periods of lectures in order to establish a baseline and follow-up score, b) An equally long, uninterrupted, and consistently structured CL intervention period to allow comparison across the periods and their relative impacts, c) Uniform instruction and training of the teaching assistants (TAs) to control for between-group differences (Cox, 2015), d) Mandatory student attendance to avoid missing by design (Jeno et al., 2017), e) Allocating the digital lectures and the digital CL group seminars to a fixed time of day to remove environmental differences, f) Incorporating the same kinds of topics throughout both the lecture and the intervention period to avoid any between-interest effects (Jeno et al., 2017) and g) Not revealing our hypotheses to the teachers or the TAs to avoid an impact on their interactions with the students.

Fig. 1. Overview of measurement time points (blue) and digital teaching and learning (orange = lectures, green = cooperative learning (CL) intervention, grey = lab exercises) during the semester.

4.4. Procedure: The cooperative learning intervention

Cooperative learning (CL) methods implemented in the intervention period followed CL principles (Johnson et al., 1998), in particular positive interdependence and individual accountability (Millis & Cottell, 1998), and previous research on CL interventions in STEM higher education (Møgelvang & Nyléhn, 2022a). First, we included cooperation as a learning objective in the course description to be able to motivate the intervention and its assessment (Cheruvelil et al., 2020). Second, we formed 20 heterogeneous "home groups" of four students and two home groups of three students based on gender, age, and study program (Millis & Cottell, 1998). During the five-week CL intervention, the student home groups met twice a week: once in group Zooms initiated and conducted by the groups themselves and once in course-scheduled group breakout rooms. Each meeting lasted one to two hours. Third, we asked the groups to draw up group contracts as such contracts are associated with more positive student attitudes and higher student contribution within the groups (Aakre & Mørkve, 2021; Oakley et al., 2004). To ensure both an overview of the planned group work and ownership of the contract, the contracts consisted of two parts: (i) A table displaying scheduled individual and group tasks and (ii) A list of points for the group to consider ensuring mutual responsibilities and expectations. For the sake of reproducibility, we have included an excerpt of the group contract (Box 1). Fourth, we implemented "jigsaw" as the key CL structure throughout the intervention period. In previous research, jigsaw in undergraduate STEM education is linked to academic success, (perceived) generic skills, and positive student attitudes (Daniel, 2016; Pilcher et al., 2015; Yimer & Feza, 2020). One of the jigsaws in our study was related to the learning of academic writing (introduction, methods, results, and discussion) and the other jigsaw was related to the learning of life cycles in different species (mouse, frog, bird, and sea urchin), see example in Box 2. In both jigsaws, the groups divided the four expert responsibilities (puzzle pieces) between them, met in expert groups (in pre-set breakout rooms) before teaching it to their home group (in other pre-set breakout rooms). In line with the focus on structure, all the activities in the course-scheduled group discussions were subject to time and task management by the instructor and TAs using different call and chat functions in Zoom breakout rooms. For each jigsaw, each home group made a short presentation summarizing their findings and the learning outcomes were synthesized in academic reports. To promote individual accountability, each group member handed in their own report to be assessed individually. Finally, the student groups were allocated their own group site on the university's digital learning platform where they could keep in touch and share resources.

4.5. Measures

To measure the students' self-reported sense of belonging, science confidence, perceived generic skills, and loneliness, the present study employed validated scales capturing each of the latent variables. All scales have previously been employed both in international and Norwegian studies, and we used available translations and documentation as detailed below.

4.5.1. Sense of belonging

To measure the students' sense of belonging in the course, we used the *Psychological Sense of School Membership* (PSSM) by Goodenow (1993). In recent higher education studies, the PSSM seems to be multidimensional resulting in three different factors equivalent to a sense of social belonging (peer-related), a sense of academic belonging (tutor-related), and a general sense of belonging with reliability estimates ranging from low .70s to .90 (Alkan, 2016; Freeman et al., 2007; Møgelvang & Nyléhn, 2022b). The original scale consists of 18 statements such as "Other students in this university/school/course take my opinions seriously", "The teachers here respect me" and "I can really be myself at this university/school/course" which are measured on a 5-point Likert scale from *strongly disagree* (1) to *strongly agree* (5). However, due to poor fit in a previous Norwegian study (Møgelvang & Nyléhn, 2022b), we removed two of the items.

4.5.2. Science Confidence

For measuring the students' science confidence, we used a scale of items adapted from previous studies investigating students' science confidence (Lopatto, 2004; Seymour et al., 2004). The scale used in the present study has previously been employed and tested for reliability for US (Walker et al., 2008) and Norwegian biology students (Cotner et al., 2020b). The scale has 13 items which are measured on a five-point Likert-scale ranging from *strongly disagree* (1) to *strongly agree* (5) as response alternatives to items such as: I am confident I can… "Discuss scientific concepts with my friends or family" and "Interpret tables and graphs".

4.5.3. Perceived generic skills

The students' perceptions of their generic skills were measured using the subscale "Generic skills" in the *Course Experience Questionnaire* (CEQ) (Ramsden, 1991). The subscale comprises six statements such as "The course sharpened my analytical skills" which are measured using a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). In previous international publications, the reliability of the CEQ "Generic skills" subscale has been acceptable, ranging from the high .60s to high .70s (Byrne & Flood, 2003; Jansen et al., 2013) and from .77 to .81 in Norwegian studies (Espeland & Indrehus, 2003; Møgelvang & Nyléhn, 2022b).

4.5.4. Course loneliness

We measured loneliness using the short "Three-Item Loneliness Scale" (TILS) developed by Hughes et al., and Cacioppo (2004). This instrument consists of three items: "How often do you feel that you lack companionship?", "How often do you feel left out?", and "How often do you feel isolated from others?" These three items are measured on a 5-point Likert scale spanning from *never* (1) to *very often* (5). In previous research among adults, the TILS has shown acceptable reliability ($\alpha = .72$ in Hughes et al. (2004) and $\alpha = .81$ in

Matthews-Ewald and Zullig (2013)). The TILS has also been employed in a large Norwegian study (Knapstad et al., 2018).

4.6. Statistical analyses

Only participants who completed all four surveys were included in our analyses. To address the research hypothesis and prediction, we conducted a range of *confirmatory* analyses using IBM SPSS Statistics 27 (IBM, 2021). First, we assessed factor structure, normal distribution, and internal consistency by running factor analyses, descriptive analyses, and reliability analyses measured with Cronbach's alpha (Cronbach, 1951). Second, to explore whether the students' scores on the outcome variables changed throughout the semester, we employed one-way repeated measures ANOVA, used Wilks' Lambda to assess significance, and adjusted for multiple comparisons through Bonferroni correction (Field, 2018). Effect sizes were measured in Partial Eta squared, which may be considered small (η_p^2 >.01 to .05), moderate ($\eta_p^2 >$ .06 to .13), or strong (η_p^2 >.14) in magnitude (Cohen, 2013). As previous research shows that psychosocial outcomes can vary meaningfully by gender and generation in college (Abdul Karim et al., 2012; Ballen et al., 2017b; Hoffman et al., 2002; Master et al., 2016; Sæthre, 2014; Sahil et al., 2020; Salehi et al., 2019; Salehi et al., 2021b; Sivertsen, 2021; Thomas, 2012; Tinto, 1975, 1993) we also conducted *exploratory* analyses. Specifically, we ran Independent-samples t-tests and One-way between-groups ANOVA with Tukey post-hoc tests to detect whether our outcomes varied by student gender and generation in college.

5. Results

All variables met assumptions of normal distribution (Field, 2009) and acceptable reliability (Cronbach, 1951) at all four measurement time points (Table 1).

Through one-way repeated measures ANOVA, we found that coming into the course, the students scored relatively high (around 4.0 on the Likert scale) on expected sense of belonging, scientific confidence, and perceived generic skills, and relatively low on loneliness (2.5) (Fig. 2, time point 1). After five weeks of digital lectures, student's self-reported sense of belonging (-.17 [95% CI, -.30 to -.04] $p< .05$), science confidence (-.46 [95% CI, -.67 to -.25] $p< .05$), and especially perceived generic skills (-.63 [95% CI, -.92 to -.35] $p< .05$) had decreased significantly (values given are mean [CI]) (Fig. 2, time point 2). Loneliness did not change. After the next 5 weeks, during which students were engaged in the cooperative learning (CL) module, these trends were reversed, as sense of belonging (.27 [95% CI, .14 to .41] $p< .05$), science confidence (.30 [95% CI, .08 to .51] $p< .05$), and particularly perceived generic skills (.41 [95% CI, .14 to .67] $p< .05$), were significantly increased, and loneliness (-.23 [95% CI, -.42 to -.03] $p< .05$) significantly decreased (Fig. 2, time point 3). After the last 5 weeks of another series of digital lectures only sense of belonging (-.14 [95% CI, -.26 to -.02] $p< .05$) was significantly decreased. Science confidence and perceived generic skills did not change significantly and neither did loneliness (Fig. 2, time point 4). Further, students reported higher scores on sense of belonging and lower scores on loneliness after the intervention (time point 3) than they did coming into the course (time point 1). Science confidence and perceived generic skills, on the other hand, did not reach the levels from time point 1. None of the differences in scores between time point 1 and time point 3 were statistically significant. Taken together, all outcome variables changed over time, with large effect sizes for sense of belonging ($\eta_p^2 =$ .34), science confidence ($\eta_p^2 = .34$), and perceived generic skills ($\eta_p^2 = .44$), and a more moderate effect for loneliness ($\eta_p^2 = .13$) (Fig. 2, Appendix A).

Except for a significant difference between females and males on perceived generic skills coming into the course (time point 1), we found no statistically significant differences in mean scores when running exploratory analyses, based on either gender or generation in college, on any of the measured outcomes (Appendix B).

Table 1
Descriptive statistics for the study variables (range 1-5) at all four time points (T1-T4).

	M	SD	Skw.	Kurt.	α
Sense of belonging, T1	3.99	0.63	0.04	-1.03	.90
Sense of belonging, T2	3.82	0.72	0.03	-1.15	.93
Sense of belonging, T3	4.09	0.65	-0.29	-1.03	.91
Sense of belonging, T4	3.95	0.72	-0.24	-1.01	.93
Science confidence, T1	4.03	0.54	0.04	-0.39	.92
Science confidence, T2	3.56	0.80	-0.99	1.80	.96
Science confidence, T3	3.86	0.62	-0.80	1.27	.94
Science confidence, T4	3.75	0.79	-0.95	1.50	.97
Perceived generic skills, T1	4.13	0.63	-0.18	-0.86	.90
Perceived generic skills, T2	3.50	0.89	-0.87	1.24	.93
Perceived generic skills, T3	3.91	0.90	-0.89	0.76	.93
Perceived generic skills, T4	3.80	0.86	-0.47	-0.08	.93
Loneliness, T1	2.51	0.84	0.17	-0.31	.76
Loneliness, T2	2.56	0.94	0.15	-0.62	.80
Loneliness, T3	2.34	0.93	0.26	-0.75	.86
Loneliness, T4	2.39	1.02	0.55	-0.40	.86

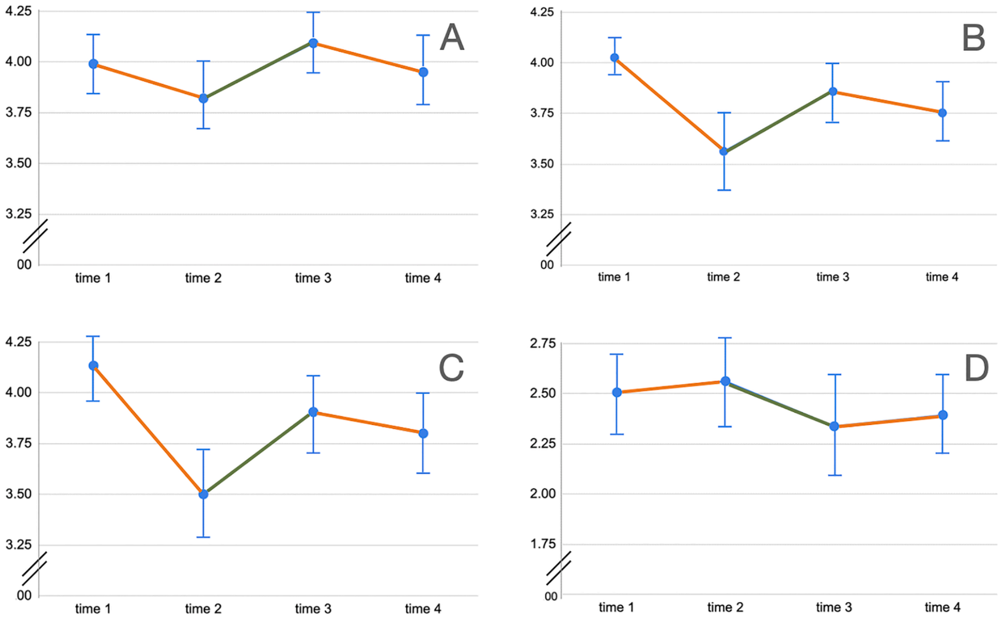


Fig. 2. Close-up four-panel figure illustrating the change in students' expressed sense of belonging (A), science confidence (B), perceived generic skills (C), and loneliness (D) including confidence intervals measured at four time points using Likert scales from 1 (strongly disagree/never) to 5 (strongly agree/very often).

6. Discussion

6.1. Main findings

The purpose of this quasi-experimental study was to implement digital cooperative learning (CL) methods and explore if such methods were associated with any change in the students' psychosocial outcomes compared to digital lectures. The students' scores on all psychosocial outcomes changed positively and significantly after five weeks of digital CL compared to the five weeks of digital lectures preceding the CL intervention. In short, the CL intervention led to the students reporting increased sense of belonging, science confidence, and perceived generic skills, and decreased loneliness in comparison to traditional digital lectures. These positive changes cannot be explained by mere development over time. The four time points of our study design reveal that the downward trend following the first period of digital lectures was reversed during the CL intervention and then stopped or even reversed again when the CL methods were replaced by the second period of digital lectures. Consequently, the positive changes in the students' reported psychosocial outcomes seem to result from the implemented CL methods. In other words, if students' psychosocial outcomes in digital STEM higher education matter, the type of digital strategies we implement in our teaching should also matter. Our study indicates that digital lectures are not only unsuited for a positive development of students' sense of belonging, science confidence, perceived generic skills, and loneliness – rather, digital lectures seem to impair these psychosocial outcomes. Taken together, these psychosocial outcomes are nourished by collaborative and student-centered teaching strategies such as CL, not lectures, and our study demonstrates this, for the first time, in a digital learning environment. To the best of our knowledge all previous research on the effect of CL on psychosocial outcomes in STEM higher education, except for a few studies on the effect of online CL on select generic skills (Lee et al., 2016; Parsazadeh et al., 2018), has been conducted in a physical setting (Canelas et al., 2017; Espinosa et al., 2019; Furuto, 2013, 2017; Kocak, 2008; Pilcher et al., 2015; Rattanatumma & Puncreobutr, 2016; Rivera, 2013; Wilton et al., 2019; Yapici, 2016). Our results indicate that the general findings of these studies could be transferable to a digital setting.

6.2. Findings in light of previous research

When comparing our findings on the relationship between CL and psychosocial outcomes in a digital setting to previous research on the equivalent relationships in physical settings, we identify several correspondences, but also novelties, particularly regarding effect sizes. Like our study, previous CL intervention studies targeting sense of belonging showed a significant increase in sense of belonging following enhanced structure (Wilton et al., 2019) and various CL methods (Yapici, 2016). Both the influence (r = .30 to .49) of enhanced structure and the effect (d = 0.5 to 0.8) (own calculations) of CL methods on sense of belonging were moderate in magnitude (Cohen, 2013). In comparison, our study identified a very large effect size (η_p^2 = .34). Large effect sizes measured in eta squared or partial eta squared equal ($\eta^2_{(p)} > .14$). A study on the effect of CL on loneliness concluded that CL heterogenous groups and CL structures led to a significant decrease in loneliness compared to traditional teaching (Kocak, 2008), with an effect of ($\eta^2 = .06$). In our study, the effect was ($\eta^2_p = .13$). According to Cohen (2013), both may be considered moderate ($\eta^2_{(p)} = .06$ to $.13$), but they figure in each end of the range. Our study is the first from STEM higher education in which CL has been specifically associated with science confidence. Previous research exists, however, on the relationship between CL and different types of science self-efficacy (Furuto, 2017; Rivera, 2013), which may resemble science confidence. These studies found significant increases in student self-efficacy following CL interventions employing heterogenous grouping, roles, and jigsaws compared to traditional teaching. Our calculations of the data in these studies show that the effect sizes varied from moderate ($d = 0.5$ to 0.8) (Rivera, 2013) to very large ($d > 0.8$) (Furuto, 2017) while in our study scientific confidence changed over time with a very large effect size ($\eta^2_p = .34$). In the case of perceived generic skills, it is hypothesized that the inherent group and task structures of CL may stimulate the development of a range of generic skills (Millis & Cottell, 1998). A previous study from a physical setting did, like our study, find significant increases and very large effect sizes ($d > 0.8$) (Cohen, 2013) in a range of perceived generic skills following a CL intervention versus traditional lectures (Canelas et al., 2017). Likewise, a qualitative study replacing recipe-based lab teaching by jigsaw methods found a substantial increase in the students' (perceived) metacognitive skills (Pilcher et al., 2015). Contrary to relationships between CL and sense of belonging, loneliness, and self-efficacy, we did identify a few studies on the effects of online CL on generic skills in higher STEM education. These were conducted in settings very different from ours, but our study from a digital setting in a Norwegian higher biology sample supports their findings. One of these studies showed that the effect sizes of CL on information evaluation skills (Parsazadeh et al., 2018) were very large ($d > 0.8$) (Cohen, 2013) and the other that CL leads to significantly increased critical thinking skills development (Lee et al., 2016).

We contribute to the understanding of CL impacts by supporting the findings of previous research on the effect of CL on several psychosocial outcomes compared to lectures, albeit in a *digital learning environment*. Our effect size for loneliness was in the upper range of moderate and the rest all very large in magnitude – which are equal to and in some instances even stronger than the effect sizes of previous research on CL in physical settings. Thus, our study suggests that the effect of CL on psychosocial outcomes in digital settings is *at least as* substantial as it is in physical settings. Taken together with the drastic decrease in sense of belonging, science confidence, perceived generic skills, and the increase in loneliness after the first period of digital lectures, we argue that psychosocial outcomes may in fact be particularly vulnerable to traditional teacher-centered instruction in a digital setting. If that is the case, it is an additional indication that teachers need to carefully consider how to teach digitally, and, encouragingly, that in a digital setting the added value of student-centered learning methods, exemplified by CL, may exceed those in physical settings.

6.3. Findings in light of theory

The drastic and positive change in the students' self-reported psychosocial outcomes during the digital cooperative learning (CL) intervention may indicate that digital CL provides many of the conditions theorized to improve student affect. In the case of science confidence, it is plausible that the CL group zooms provided the necessary experiences, e.g., interaction, working with specific scientific tasks, and immediate feedback to inform the students' positive self-evaluation and sense of a future in biology (Bandura, 1986; Lent et al., 1994). During the CL intervention the students worked cooperatively to solve biology tasks by means of jigsaw structures, thus integrating the two conditions hypothesized as needed to develop generic skills: content knowledge and active group work (Ballantine & McCourt Larres, 2007; Kember et al., 2007; Smith & Bath, 2006; Tynjälä & Gijbels, 2012; Virtanen & Tynjälä, 2019, 2022). Further, initiatives to increase students' inclusion and wellbeing (Adriansen & Madsen, 2012) through interaction and meaningful relationships (Baumeister & Leary, 1995) may be effective means to reducing loneliness and increasing sense of belonging. In our study, the students were given many opportunities to interact and experience meaningful relationships through highly structured, small, heterogeneous, and fixed CL groups with shared goals during the CL intervention (Gillies, 2003; Johnson et al., 1994; Kagan, 2021; Millis & Cottell, 1998). In fact, it should be noted that sense of belonging and loneliness, as the only of the measured variables, improved not only in comparison to the first period of lectures (time point 2), but also to time point 1 (Fig. 2). Time point 1 was when the students stated their expected sense of belonging and loneliness throughout the course based on experiences from similar courses the preceding semester. Hypothetically, the improvement in sense of belonging and loneliness measured after the CL intervention compared to both time point 2 (highly significant) and time point 1 (not significant), could be influenced by the types of student groups. Belonging to highly structured, small, heterogeneous, and fixed groups (Gillies, 2003; Johnson et al., 1994; Kagan, 2021; Millis & Cottell, 1998) may to a larger degree, than other types of groups, build up the necessary personal safety to interact and promote inclusion (Adriansen & Madsen, 2012; Baumeister & Leary, 1995), particularly in a digital setting where the threshold to initiate dialog seems higher than in a physical setting.

Theory may also help us understand why the positive developments in all psychosocial outcomes stopped or even reversed after the second period of lectures, and particularly why sense of belonging decreased significantly anew. Perceptions of belonging and opportunities to belong reinforce and affect one another continuously in the development of belonging (Allen et al., 2021), making consistent and systematic inclusion practices in higher education both important and required to secure the continuous sense of belonging among the students (Murdock-Perriera et al., 2019). The systematic practice to include all the students through the CL intervention disappeared when the second period of lectures began. Thus, the interaction needed to fulfil the need to belong also disappeared, and the students no longer experienced the same opportunities to belong. In a digital setting it is not just a matter of catching up with your group members in the break. The natural meeting place is simply not present, meaning that the importance of intended and systematic university practices to include all students become even more warranted in digital settings. The other psychosocial outcomes were more consistent when compared to sense of belonging. This may be related to science confidence being a malleable trait (Burns et al., 2016), while generic skills have been explained in terms of life-long learning (Bourn, 2018). Also, interventions offering targeted opportunities for social interaction seem to be successful in reducing loneliness over time (Hawkley & Cacioppo, 2010). All in all, the fact that none of the measured variables, except for sense of belonging, changed significantly in the second period of digital lectures may mean that there can be a lot to be gained from implementing periods of CL methods in digital settings.

7. Limitations and strengths

There are some limitations to our study. The design is quasi-experimental and lacks the advantages provided by randomization, and a control group could have strengthened the causal claims (Shadish et al., 2002). Self-reported measures, with their inherent biases, might also pose a limitation to the study (Czaja & Blair, 2005). Further, our conflation of the terms "science confidence" and "(science) self-efficacy" may be considered a limitation since the terms may not be as closely linked as the discipline-based educational research community seem to suggest. To counter any confusion and to promote transparency on this potential issue, we have included an overview (Appendix C) of how we measure science confidence (and all the other constructs) in the present study. The intervention was conducted in the spring 2021, i.e., in the middle of the COVID-19 pandemic when higher education in Norway was subjected to severe restrictions and the Norwegian population showed increasing signs of de-motivation and frustration with the seemingly endless situation. This backdrop may have influenced the answers of the students. Perhaps they would not have felt that discouraged from digital lectures nor that positive towards cooperative learning (CL) methods under different circumstances. Even so, we employed many design controls to strengthen causal claims (Shadish et al., 2002). These design controls and in particular the use of a double pre-test and a follow-up indicated a real and substantial difference in the students' psychosocial outcomes following the implemented digital CL methods compared to digital lectures. Further, a repeated measures design enables us to detect within-person change over time and has high statistical power (Guo et al., 2013, p. 100) and the effect sizes of the intervention were all very large, except for loneliness which was in the upper range of moderate. Thus, it is reasonable to infer that the changes in student scores in our sample are linked to the CL intervention rather than to confounding variables (Shadish et al., 2002). Also, the study adds new and valuable knowledge on a sample, i.e., Norwegian undergraduate biology students, and topic, i.e., digital CL methods, not previously researched. With digital teaching and learning on the rise, the knowledge gained from the digital nature of the study is especially valuable.

Our ability to extrapolate these findings to a larger higher-education community is limited. Our target population - Norwegian biology students - is relatively homogeneous, being composed primarily of continuing-generation, White, Norwegian-educated students, and may not be representative of populations elsewhere. Here we analyzed two aspects of diversity, i.e., gender and generation in college, and we did not detect any variability in response to the intervention. We cannot say for sure how students from, e.g., underrepresented racial or ethnic minorities might have responded to the intervention, but we hypothesize, based on similar studies (Binning et al., 2020; Hammarlund et al., 2022c) that the positive effects might have been even *greater* for these students. For example, a US-based study in introductory chemistry, (Hammarlund, 2022c) shows that a sense-of-belonging intervention positively impacted underrepresented minority students, but it did not differentially affect the performance of women or first-generation college students. We note further that the field of psychosocial classroom interventions is beset by a "replication crisis," driven largely by studies conducted in one classroom, in one subject, in one geographic area - often the United States (Camerer et al., 2018; Open Science Collaboration, 2015). Few studies have investigated similar interventions in Norway, and none, to our knowledge, have explored the impacts of CL in a digital learning environment. Thus, our study population makes this work a valuable contribution to the growing field.

Finally, the addition of qualitative approaches would have added layers of understanding. For instance, interviews with students could have allowed us to better understand *which* features of CL best activated the positive outcomes and perhaps brought us closer to a more elaborate understanding of the benefits of CL. Thus, future studies will be strengthened by the incorporation of qualitative elements in the study design.

8. Conclusion

Taken together, our study adds new and valuable knowledge about digital teaching and learning in higher education. First, it underlines the differences between two digital teaching strategies in relation to psychosocial outcomes among students in higher education – and concludes that the positive effect of digital cooperative learning (CL) on psychosocial outcomes compared to digital lectures is substantial. Second, the findings suggest that psychosocial outcomes may be particularly vulnerable in a digital setting, making it even more important for instructors to pursue teaching strategies favoring such outcomes. Third, the effect sizes of the digital CL methods implemented in the study on psychosocial outcomes among the students were at least as strong as the effect sizes in previous studies on physical CL interventions (Canelas et al., 2017; Furuto, 2017; Kocak, 2008; Pilcher et al., 2015; Rattanatumma & Puncreobutr, 2016; Wilton et al., 2019; Yapici, 2016). This suggests not only that CL methods from physical settings may be transferable to digital settings, but also that the effect of CL in digital settings may be even greater than in physical settings. Fourth, digital settings do not offer spontaneous and informal meeting places for the students, making it even more important that higher education institutions implement systematic digital inclusion practices in all courses for all students. Our findings are in alignment with theory in the field of CL, sense of belonging, scientific confidence, generic skills, and loneliness.

In sum, teachers developing digital courses should take care to include student-centered approaches in their courses. Our findings suggest the positive impacts of this engagement can be meaningful and far-reaching, and we urge our colleagues in higher education to seek and implement digital strategies —such as digital CL—that promote student well-being.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

None.

Appendix A

Table A.1

Appendix B

Table B.1 and Table B.2

Appendix C

Scales and items used in the study "Cooperative Learning Goes Online: Teaching and Learning Intervention in a Digital Environment Impacts Psychosocial Outcomes in Biology Students"

Science confidence scale (Lopatto, 2004; Seymour et al., 2004)

I am confident I can...

1. Discuss scientific concepts with my friends or family
2. Think critically about scientific findings I read about in the media
3. Read the scientific literature (current papers, written by scientists, in scientific journals)
4. Determine what is - and is not - valid scientific evidence
5. Make an argument using scientific evidence
6. Interpret tables and graphs
7. Pose questions that can be addressed by collecting and evaluating scientific evidence
8. Collect research-related data
9. Analyze data
10. Present scientific results in writing or orally
11. Understand scientific processes behind important scientific issues in the media
12. Understand the science content of this course
13. Use scientific thinking to solve problems outside of this course

Likert-scale ranging from 1 (strongly disagree), 2 (partly disagree), 3 (neither agree nor disagree), 4 (partly agree), to 5 (strongly agree)

Table A.1
Means, standard deviations, and one-way repeated measures ANOVA statistics for the study variables.

Variable	*M*	*SD*	$F(3, 68)$	*p*	η_p^2
Sense of belonging			11.87	<.001	.34
Time 1	3.99	0.63			
Time 2	3.82	0.72			
Time 3	4.09	0.65			
Time 4	3.95	0.72			
Scientific confidence			11.55	<.001	.34
Time 1	4.03	0.54			
Time 2	3.56	0.80			
Time 3	3.86	0.62			
Time 4	3.75	0.79			
Generic skills			17.99	<.001	.44
Time 1	4.13	0.63			
Time 2	3.50	0.89			
Time 3	3.91	0.90			
Time 4	3.80	0.86			
Loneliness			3.24	.028	.13
Time 1	2.51	0.84			
Time 2	2.56	0.94			
Time 3	2.34	0.93			
Time 4	2.39	1.02			

Table B.1
Means, standard deviations, and independent-samples T-tests for the study variables

Variables	Females		Males				
	M	*SD*	*M*	*SD*	*t*(69)	*p*	*d*
Sense of belonging							
Time 1	4.12	0.66	3.83	0.56	-1.95	.056	.46
Time 2	3.92	0.73	3.70	0.68	-1.32	.192	.31
Time 3	4.17	0.68	3.99	0.60	-1.17	.244	.28
Time 4	4.09	0.70	3.79	0.72	-1.76	.082	.42
Science confidence							
Time 1	4.11	0.46	3.93	0.61	-1.40	.165	.34
Time 2	3.54	0.85	3.60	0.74	0.31	.757	.07
Time 3	3.86	0.67	3.86	0.56	0.02	.986	.00
Time 4	3.79	0.77	3.71	0.83	-0.42	.676	.10
Generic skills							
Time 1	4.34	0.62	3.88	0.55	-3.27	.002	.78
Time 2	3.49	0.89	3.51	0.89	0.09	.929	.02
Time 3	3.96	0.91	3.84	0.89	-0.57	.569	.14
Time 4	3.94	0.86	3.64	0.86	-1.47	.146	.35
Loneliness							
Time 1	2.62	0.73	2.38	0.96	-1.20	.235	.29
Time 2	2.62	0.84	2.49	1.05	-0.60	.552	.14
Time 3	2.50	0.83	2.15	1.02	-1.59	.115	.38
Time 4	2.39	1.01	2.40	1.04	0.01	.991	.00

Note. Females (n=39); Males (n=32).

Table B.2
Means, standard deviations, and one-way between groups ANOVA statistics for the study variables.

Variable	NPHE		OPHE		BPHE				
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*(2, 68)	*p*	η^2
Sense of belonging									
Time 1	4.13	0.60	3.89	0.73	4.01	0.60	0.42	.657	.01
Time 2	4.10	0.77	3.73	0.83	3.81	0.67	0.68	.511	.02
Time 3	4.01	0.80	4.20	0.63	4.07	0.64	0.35	.709	.01
Time 4	3.98	0.74	4.09	0.70	3.90	0.74	0.46	.636	.01
Science confidence									
Time 1	3.85	0.28	4.03	0.61	4.05	0.54	0.44	.649	.01
Time 2	3.79	0.47	3.44	0.86	3.57	0.82	0.49	.615	.01
Time 3	3.77	0.67	3.89	0.59	3.86	0.63	0.09	.914	.00
Time 4	3.92	0.73	3.87	0.65	3.69	0.85	0.51	.606	.01
Generic skills									
Time 1	4.31	0.71	4.24	0.70	4.07	0.60	0.72	.492	.02
Time 2	3.55	0.83	3.45	0.90	3.51	0.91	0.04	.962	.00
Time 3	3.74	1.32	4.03	0.93	3.89	0.83	0.29	.751	.01
Time 4	3.81	1.17	4.08	0.92	3.70	0.79	1.22	.302	.03
Loneliness									
Time 1	2.71	0.83	2.37	0.91	2.52	0.83	0.43	.651	.01
Time 2	2.90	1.07	2.33	0.90	2.60	0.93	1.01	.371	.03
Time 3	2.81	1.15	2.02	0.88	2.38	0.90	2.01	.142	.06
Time 4	2.95	1.41	2.08	0.91	2.43	0.97	1.94	.151	.05

Note. NPHE = No parent has higher education (n=7); OPHE = One parent has higher education (n=17); BPHE = Both parents have higher education (n=47)

(Perceived) generic skills (Ramsden, 1991)

1 The course developed my problem-solving skills
2 The course sharpened my analytical skills
3 The course helped me to develop my ability to work as a team member
4 As a result of my course, I feel confident about tackling unfamiliar problems
5 The course improved my skills in written communication
6 The course helped me to develop my ability to plan my own work

Likert-scale ranging from 1 (strongly disagree), 2 (partly disagree), 3 (neither agree nor disagree), 4 (partly agree), to 5 (strongly agree)

Sense of belonging (Goodenow, 1993**)***

Box 1
Implemented CL structure: Group contract.

A group contract provides guidelines for group work and group tasks with the purpose of establishing common expectations and tools to secure workflow, develop constructive communication, and manage potential conflicts (Oakley et al., 2004). Here, we share specific guidelines from our intervention.

Group contract

Name of group:____

A group contract may ensure a positive work flow, harmony, and autonomy in group work. This group contract consists of two parts:

Part 1: An overview of fixed contractual group tasks
Part 2: A list of points to consider in your group work

Part 1:

Week	Individual preparations	Group tasks	Responsibility
9	For the third group zoom: Read text, study figures and answer the question about life cycle in general.	In the third group zoom: - Discuss life cycle in general - Agree on and distribute responsibility (expertise knowledge) for the life cycle of the following species: frog, bird, sea urchin and mouse - E-mail course responsible with an overview of the distribution of species (who becomes expert in what). Deadline: - Agree on time and convene for the fourth group zoom	Name: Name:
	For the fourth scheduled group zoom: Become an expert in "your" species and note down key content	In the fourth scheduled group zoom: - In expert groups: Compare and discuss your expert notes and decide on a presentation strategy - In home groups: Present «your» species and compare and discuss the life cycle of the four species. - Make a short joint presentation and upload it to the group site	Name:

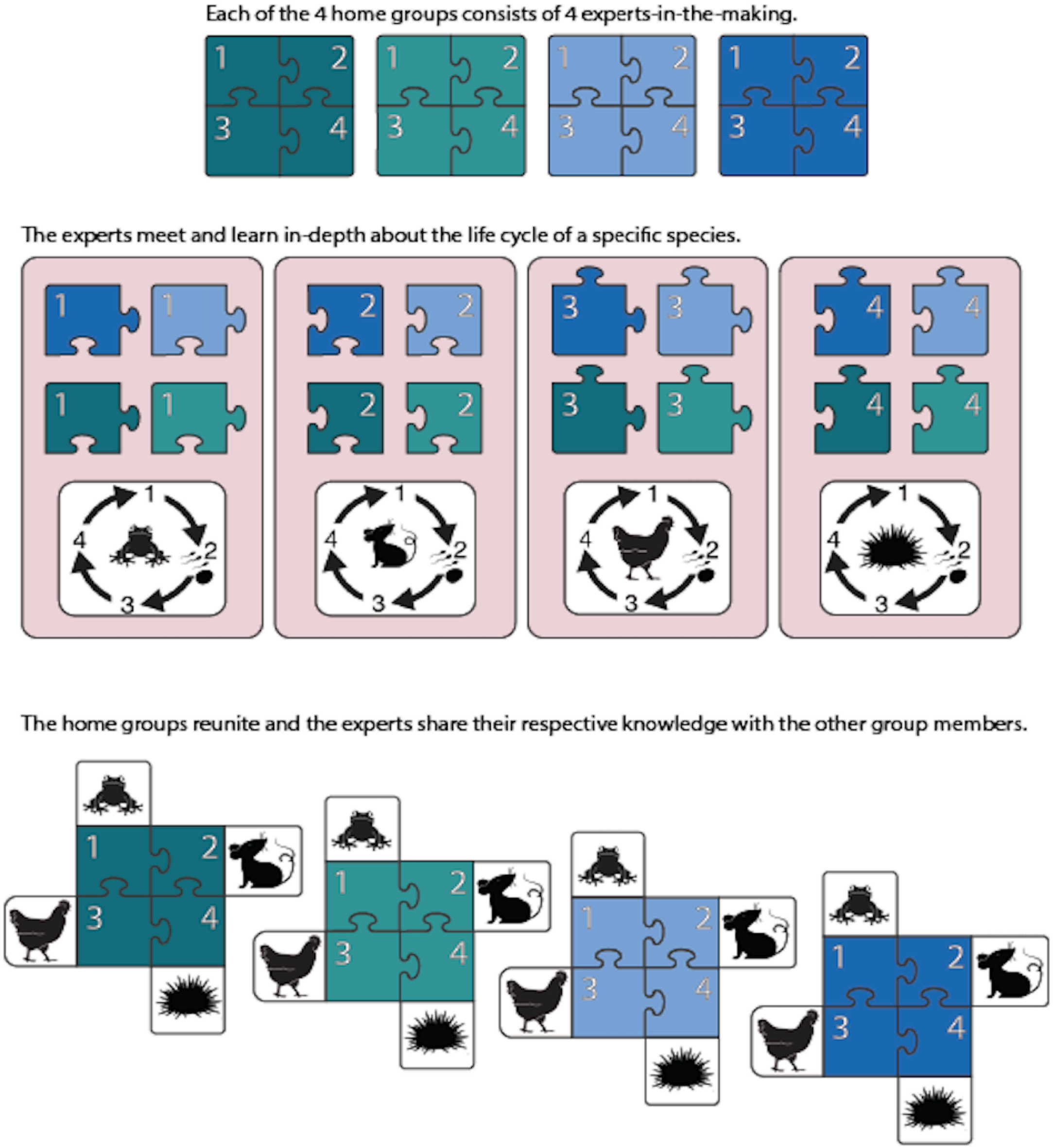
Part 2:
Discuss and agree on the points you want to guide your group work. Write down and sign.

- Decisions. How do we make good and fair decisions in our group?
- Disagreement. How do we deal with disagreement?
- Contact. Where and when can we reach each other?
- Participation. How do we make sure that everyone participates on equal terms?
- ETC.

1 I feel like a real part of this course
2 It is hard for people like me to be accepted here
3 Other students in this course take my opinions seriously
4 Most teachers in this course are interested in me
5 Sometimes I feel as if I don't belong here
6 There's at least one teacher or other adult in this course I can talk to if I have a problem
7 People in this course are friendly to me
8 Teachers here are not interested in people like me
9 I am included in lots of activities in this course

Box 2
Implemented CL structure: Jigsaw.

Jigsaw: each group member takes responsibility for learning a specific part (piece) of a complex whole (jigsaw) and teaching it to the rest of the group. This way the group cooperates to put all the pieces of the jigsaw together (Millis & Cottell, 1998). Here, we share a specific jigsaw example from our intervention.



10 I am treated with as much respect as other students
11 I feel very different from most other students here
12 I can really be myself in this course
13 The teachers here respect me
14 I wish I were in a different course
15 I feel proud of belonging to this course
16 Other students here like me the way I am

Likert-scale ranging from 1 (strongly disagree), 2 (partly disagree), 3 (neither agree nor disagree), 4 (partly agree), to 5 (strongly agree)

*Items adapted to the studied course (not school) and two items removed

Loneliness short item scale (Hughes et al., 2004)

1 How often do you feel that you lack companionship?
2 How often do you feel left out?
3 How often do you feel isolated from others?

Likert-scale ranging from 1 (never), 2 (rarely), 3 (sometimes), 4 (often), to 5 (very often)

References

Aakre, I. G., & Mørkve, M. S. (2021). *Det er lurt å skrive en avtale! Ferske studenters opplevelse av gruppearbeid og gruppeavtaler.* MNT conference, University of Agder.

Abdul Karim, A. M., Abdullah, N., Abdul Rahman, A. M., Noah, S. M., Wan Jaafar, W. M., Othman, J., Borhan, L., Badushah, J., & Said, H. (2012). A nationwide comparative study between private and public university students' soft skills. *Asia Pacific Education Review, 13*(3), 541–548. https://doi.org/10.1007/s12564-012-9205-1

Adriansen, H. K., & Madsen, L. M. (2012). Facilitation: A novel way to improve students' well-being. *Innovative Higher Education, 38*(4), 295–308. https://doi.org/10.1007/s10755-012-9241-0

Ainscough, L., Foulis, E., Colthorpe, K., Zimbardi, K., Robertson-Dean, M., Chunduri, P., & Lluka, L. (2016). Changes in biology self-efficacy during a first-year university course. *CBE Life Science Education, 15*(2), ar19. https://doi.org/10.1187/cbe.15-04-0092

Alkan, N. (2016). Psychological sense of university membership: An adaptation study of the PSSM scale for turkish university students. *The Journal of Psychology, 150*(4), 431–449. https://doi.org/10.1080/00223980.2015.1087373

Allen, K.-A., Kern, M. L., Rozek, C. S., McInerney, D. M., & Slavich, G. M. (2021). Belonging: A review of conceptual issues, an integrative framework, and directions for future research. *Australian Journal of Psychology, 73*(1), 87–102. https://doi.org/10.1080/00049530.2021.1883409

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology, 44*(5), 427–445. https://doi.org/10.1016/j.jsp.2006.04.002

Ballantine, J., & McCourt Larres, P. (2007). Cooperative learning: A pedagogy to improve students' generic skills? *Education & Training, 49*(2), 126–137. https://doi.org/10.1108/00400910710739487

Ballen, C. J., Salehi, S., & Cotner, S. (2017b). Exams disadvantage women in introductory biology. *Plos One, 12*(10), Article e0186419-e0186419. https://doi.org/10.1371/journal.pone.0186419

Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K. R. (2017a). Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning. *CBE—Life Sciences Education,, 16*(4), ar56.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology, 4*(3), 359–373. https://doi.org/10.1521/jscp.1986.4.3.359

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

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). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-2324-5_2.

Binning, K. R., Kaufmann, N., McGreevy, E. M., Fotuhi, O., Chen, S., Marshman, E., … Singh, C. (2020). Changing social contexts to foster equity in college science courses: An ecological-belonging intervention. *Psychological Science, 31*(9), 1059–1070. https://doi.org/10.1177/0956797620929984

Bourn, D. (2018). *Understanding global skills for 21st century professions.* Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-97655-6_3

Burns, K. M., Burns, N. R., & Ward, L. (2016). Confidence - more a personality or ability trait? It depends on how it is measured: A comparison of young and older adults. *Frontiers in Psychology, 7*, 518. https://doi.org/10.3389/fpsyg.2016.00518

Byrne, M., & Flood, B. (2003). Assessing the teaching quality of accounting programmes: An evaluation of the Course Experience Questionnaire. *Assessment and Evaluation in Higher Education, 28*(2), 135–145. https://doi.org/10.1080/02602930301668

Børve, K., Straume, K., Smith, R., Surén, P., Samdal, O., & Hågensen, E. (2021). Tiltak for oppfølging av studenter ved fagskoler, høyskoler og universiteter under koronoapandemien [Initiatives to follow up students in higher education during the Corona pandemic]. https://www.regjeringen.no/contentassets/459df2260de74e57a9f1317bd0bb492f/tiltak-for-oppfolging-av-studenter-ved-fagskoler-hoyskoler-og-universiteter-under-pandemien.pdf.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., & Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour, 2*(9), 637–644. https://doi.org/10.1038/s41562-018-0399-z

Camfield, E. K., Schiller, N. R., & Land, K. M. (2021). Nipped in the bud: COVID-19 reveals the malleability of STEM student self-sfficacy. *CBE Life Science Education, 20*(2), ar25. https://doi.org/10.1187/cbe.20-09-0206

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

Cheruvelil, 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. https://doi.org/10.1525/abt.2020.82.1.18

Cotner, S., Jeno, L. M., Walker, J. D., Jørgensen, C., & Vandvik, V. (2020b). Gender gaps in the performance of Norwegian biology students: The roles of test anxiety and science confidence. *International Journal of STEM Education, 7*(1), 1–10. https://doi.org/10.1186/s40594-020-00252-1

Cox, C. (2015). Incorporating more individual accountability in group activities in general chemistry. *Journal of College Science Teaching, 44*(3), 30–36. https://doi.org/10.2505/4/jcst15_044_03_30

Crano, W. D., Brewer, M. B., & Lac, A. (2015). *Principles and methods of social research*. New York: Routledge, 3rd ed. ed.

Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., Magni, P. A., & Lam, S. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses. *Journal of Applied Learning & Teaching, 3*(1), 1–20. https://doi.org/10.37074/jalt.2020.3.1.7

Creswell, J. W. (2014). *Research design : Qualitative, quantitative, and mixed methods approaches*. Los Angeles: SAGE, 4th ed.; International student ed.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

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.

Cutrona, C. E. (1982). Transition to college: Loneliness and the process of social adjustment. In L. A. Peplau, & D. Perlman (Eds.), *Loneliness : A sourcebook of current theory, research and therapy*. New York: Wiley, Article 291309.

Czaja, R., & Blair, J. (2005). Designing surveys: A guide to decisions and procedures (2nd ed.). Thousand Oaks: Pine Forge.

Damşa, C., de Lange, T., Elken, M., Esterhazy, R., Fossland, T., Frølich, N., Hovdhaugen, E., Maassen, P., Nerland, M. B., Nordkvelle, Y. T., Stensaker, B., Tømte, C., Vabø, A., Wiers-Jenssen, J., Aamodt, P. O., & Stensaker, B. (2015). *Quality in Norwegian higher education: A review of research on aspects affecting student learning*. Nordic Institute for Studies in Innovation Research and Education (NIFU). https://nifu.brage.unit.no/nifu-xmlui/bitstream/handle/11250/2360199/NIFUreport2015-24.pdf?sequence=4&isAllowed=y.

Daniel, K. L. (2016). Impacts of active learning on student outcomes in large-lecture biology courses. *The American Biology Teacher, 78*(8), 651–655.

Davey, T., Meerman, A., Galán-Muros, V., Orazbayeva, B., & Baaken, T. (2018). The state of university-business cooperation in Europe (Final report). https://www.ub-cooperation.eu/.

Davidson, N. (2021). Synthesis of CL approaches and a multifaceted rationale for CL - past, present, and future. In N. Davidson (Ed.), *Pioneering perspectives in cooperative learning* (pp. 234–254). New York: Routledge.

Deng, J., Zhou, F., Hou, W., Silver, Z., Wong, C. Y., Chang, O., Drakos, A., Zuo, Q. K., & Huang, E. (2021). The prevalence of depressive symptoms, anxiety symptoms and sleep disturbance in higher education students during the COVID-19 pandemic: A systematic review and meta-analysis. *Psychiatry Research, 301*, Article 113863. https://doi.org/10.1016/j.psychres.2021.113863

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Procedings of the National Academy of Sciences USA, 116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Espeland, V., & Indrehus, O. (2003). Evaluation of students' satisfaction with nursing education in Norway. *Journal of Advanced Nursing, 42*(3), 226–236. https://doi.org/10.1046/j.1365-2648.2003.02611.x

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).

European Commision. (2021). Digital education action plan (2021-2027): Resetting education and training for the digital age [Online]. Retrieved 16.06.22 from https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan_en.

European Commision. (2022). Complete guide to GDPR compliance. Retrieved 20.04.22 from https://gdpr.eu/.

Field, A. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll* (3rd ed.). Los Angeles: SAGE.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Los Angeles: SAGE.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Procedings of the National Academy of Science USA, 111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education, 75*(3), 203–220. https://doi.org/10.3200/JEXE.75.3.203-220

Furuto, M. (2013). *A case study on a diverse college algebra classroom: Analyzing pedagogical strategies to enhance students' mathematics self-efficacy*. [Oregon State University]. ProQuest Disertation Publishing.

Furuto, M. A. (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

Gillies, R. M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education, 41*(3), 39–54.

Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in Schools, 30*(1), 79–90.

Guo, Y., Logan, H. L., Glueck, D. H., & Muller, K. E. (2013). Selecting a sample size for studies with repeated measures. *BMC Medical Research Methodology, 13*(1), 100–108. https://doi.org/10.1186/1471-2288-13-100

Hagerty, B. M. K., Lynch-Sauer, J., Patusky, K. L., Bouwsema, M., & Collier, P. (1992). Sense of belonging: A vital mental health concept. *Archives of Psychiatric Nursing, 6*(3), 172–177. https://doi.org/10.1016/0883-9417(92)90028-H

Hawkley, L. C., & Cacioppo, J. T. (2010). Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of Behavioral Medicine, 40*(2), 218–227. https://doi.org/10.1007/s12160-010-9210-8

Hayley, A. C., Downey, L. A., Stough, C., Sivertsen, B., Knapstad, M., & Øverland, S. (2017). Social and emotional loneliness and self-reported difficulty initiating and maintaining sleep (DIMS) in a sample of Norwegian university students. *Scandinavian Journal of Psychology, 58*(1), 91–99. https://doi.org/10.1111/sjop.12343

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics, 19*(4), 451–464. https://doi.org/10.1016/j.labeco.2012.05.014

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers and education, 90*, 36–53. https://doi.org/10.1016/j.compedu.2015.09.005

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, M. (2020). The difference between emergency remote teaching and online learning. *Educause Review*. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning.

Hoffman, M., Richmond, J., Morrow, J., & Salomone, K. (2002). Investigating "sense of belonging" in first-year college students. *Journal of College Student Retention: Research, Theory & Practice, 4*(3), 227–256.

Houlden, S., & Veletsianos, G. (2020). *Coronavirus pushes universities to switch to online classes – but are they ready?*, 12.03. The Conversation https://theconversation.com/coronavirus-pushes-universities-to-switch-to-online-classes-but-are-they-ready-132728.

Hughes, M. E., Waite, L. J., Hawkley, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys: Results from two population-based studies. *Research on Aging, 26*(6), 655–672. https://doi.org/10.1177/0164027504268574

Hammarlund, S. P., Scott, C., Binning, K. R., & Cotner, S. (2022c). Context matters: How an ecological-belonging intervention can reduce inequities in STEM. *BioScience, 72*(4), 387–396.

Händel, M., Stephan, M., Gläser-Zikuda, M., Kopp, B., Bedenlier, S., & Ziegler, A. (2020). Digital readiness and its effects on higher education students' socio-emotional perceptions in the context of the COVID-19 pandemic. *Journal of Research on Technology in Education*, 1–13. https://doi.org/10.1080/15391523.2020.1846147

IBM. (2021). *IBM SPSS Statistics for Macintosh, Version 27.0*. IBM Corp.

Jacobs, G. M., McCafferty, S. G., & Iddings, A. C. D. (2006). Roots of cooperative learning in general education. In G. M. Jacobs, S. G. McCafferty, & A. C. D. Iddings (Eds.), *Cooperative learning and second language teaching* (pp. 9–17). Cambridge University Press.

Jansen, E., van der Meer, J., & Fokkens-Bruinsma, M. (2013). Validation and use of the CEQ in The Netherlands. *Quality Assurance in Education, 21*(4), 330–343. https://doi.org/10.1108/QAE-11-2012-0041

Jeno, L. M., Raaheim, A., Kristensen, S. M., Kristensen, K. D. B., Hole, T. N., Haugland, M. J., & Mæland, S. (2017). The relative effect of team-based learning on motivation and learning: A self-determination theory perspective. *CBE Life Science Education, 16*(ar 59), 1–12. https://doi.org/10.1187/cbe.17-03-0055

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina: Interaction Book Company.

Johnson, D. W., & Johnson, R. T. (2016). Cooperative learning and teaching citizenship in democracies. *International Journal of Educational Research, 76*, 162–177. https://doi.org/10.1016/j.ijer.2015.11.009
Johnson, D. W., & Johnson, R. T. (2018). Cooperative learning: The foundation for active learning. In S. M. Brito (Ed.), *Active learning: Beyond the future*. IntechOpen. https://doi.org/10.5772/intechopen.73460.
Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1994). *Cooperative learning in the classroom*. Association for Supervision and Curriculum Development.
Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). *Active learning: Cooperation in the college classroom* (2nd ed.). Edina: 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.
Kagan, S. (2021). The structural approach and Kagan structures. In N. Davidson (Ed.), *Pioneering perspectives in cooperative learning* (pp. 78–127). New York: Routledge.
Kember, D., Leung, D. Y. P., & Ma, R. S. F. (2007). Characterizing learning environments capable of nurturing generic capabilities in higher education. *Research in Higher Education, 48*(5), 609–632. https://doi.org/10.1007/s11162-006-9037-0
Knapstad, M., Heradstveit, O., & Sivertsen, B. (2018). Students' Health and Wellbeing Study 2018. SiO. https://www.studenthelse.no/tema/sosiale_forhold.
Kocak, R. (2008). The effects of cooperative learning on psychological and social traits among undergraduate students. *Social Behavior and Personality, 36*(6), 771–782. https://doi.org/10.2224/sbp.2008.36.6.771
Lashley, M. A., Acevedo, M., Cotner, S., & Lortie, C. J. (2020a). How the ecology and evolution of the COVID-19 pandemic changed learning. *Ecology and evolution, 10* (22), 12412–12417. https://doi.org/10.1002/ece3.6937
Lederer, A. M., Hoban, M. T., Lipson, S. K., Zhou, S., & Eisenberg, D. (2021). More than inconvenienced: The unique needs of U.S. college students during the COVID-19 pandemic. *Health Educuation and Behavior, 48*(1), 14–19. https://doi.org/10.1177/1090198120969372
Lee, H., Parsons, D., Kwon, G., Kim, J., Petrova, K., Jeong, E., & Ryu, H. (2016). Cooperation begins: Encouraging critical thinking skills through cooperative reciprocity using a mobile learning game. *Computers & Education, 97*, 97–115. https://doi.org/10.1016/j.compedu.2016.03.006
Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45*(1), 79–122. https://doi.org/10.1006/jvbe.1994.1027
Lillejord, S., Børte, K., Nesje, K., & Ruud, E. (2018). *Learning and teaching with technology in higher education - a systematic review*. Knowledge Center for Education. https://www.uis.no/sites/default/files/inline-images/slbTVKbYroGY36QPKNy3PzogniywJ7zrgc6Jatku2uQCMD7Ndb.pdf.
Loh, R. C.-Y., & Ang, C.-S. (2020). Unravelling cooperative learning in higher education. *Research in Social Sciences and Technology, 5*(2), 22–39. https://doi.org/10.46303/ressat.05.02.2
Lopatto, D. (2004). Survey of undergraduate research experiences (SURE): First findings. *Cell Biology Education, 3*(4), 270–277.
Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66*(4), 423–458.
Male, S. A., Bush, M. B., & Chapman, E. S. (2011). An Australian study of generic competencies required by engineers. *European Journal of Engineering Education, 36*(2), 151–163. https://doi.org/10.1080/03043797.2011.569703
Maslow, A. H. (1968). *Toward a psychology of being* (2nd ed.). New York: D. Van Nostrand.
Master, A., Cheryan, S., & Meltzoff, A. N. (2016). Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of Educational Psychology, 108*(3), 424–437. https://doi.org/10.1037/edu0000061
Matthews-Ewald, M. R., & Zullig, K. J. (2013). Evaluating the performance of a short loneliness scale among college students. *Journal of College Student Development, 54*(1), 105–109. https://doi.org/10.1353/csd.2013.0003
Mazur, C., Creech, C., Just, J., Rolle, C., Cotner, S., & Hewlett, J. (2021a). Teaching during COVID-19 times: A community college perspective. *Journal of Microbiology & Biology Education, 22*(1). https://doi.org/10.1128/jmbe.v22i1.2459
Millis, B. J. (2010). Why faculty should adopt cooperative learning approaches. In B. J. Millis (Ed.), *Cooperative learning in higher education: Across the disciplines, across the academy* (1st ed.). Sterling: Stylus.
Millis, B. J., & Cottell, P. G. (1998). *Cooperative learning for higher education faculty*. Westport: American Council on Education/Oryx.
Møgelvang, A., & Nyléhn, J. (2022a). Co-operative learning in undergraduate mathematics and science education: A scoping review. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-022-10331-0
Møgelvang, A. & Nyléhn, J. (2022b). Interdependence between perceived cooperative learning, sense of belonging, and generic skills in undergraduate STEM education. *Nordic Journal of STEM Education*. (Re-submitted).
Murdock-Perriera, L. A., Boucher, K. L., Carter, E. R., & Murphy, M. C. (2019). Places of belonging: Person- and place-focused interventions to support belonging in college. In M. B. Paulsen, & L. W. Pern (Eds.), *Higher education: Handbook of theory and research* (pp. 291–323). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-03457-3_7.
Nissen, J. M., & Shemwell, J. T. (2016). Gender, experience, and self-efficacy in introductory physics. Physical review. *Physics Education Research, 12*(2), Article 020105. https://doi.org/10.1103/PhysRevPhysEducRes.12.020105
Norwegian Centre for Research Data (NSD). (2021). *Data protection services*. Retrieved 20.04.22 from https://www.nsd.no/en/data-protection-services.
Oakley, B., Felder, R. M., Brent, R., & Elhajj, I. (2004). Turning student groups into effective teams. *Journal of Student Centered Learning, 2*(1), 9–34.
Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), 1–8. https://doi.org/10.1126/science.aac4716
Parsazadeh, N., Ali, R., & Rezaei, M. (2018). A framework for cooperative and interactive mobile learning to improve online information evaluation skills. *Computers & Education, 120*, 75–89. https://doi.org/10.1016/j.compedu.2018.01.010
Peplau, L. A., & Perlman, D. (1982). Perspectives on loneliness. In L. A. Pepla, & D. Perlman (Eds.), *Loneliness : A sourcebook of current theory, research and therapy* (pp. 1–18). New York: Wiley.
Phillips, R., Seaborne, K., Goldsmith, A., Curtis, N., Davies, A., Haynes, W., McEnroe, R., Murphy, N., O'Neill, L., Pacey, C., Walker, E., & Wordley, E. (2022). Student loneliness through the pandemic: How, why and where? *The Geographical Journal, 188*(2), 277–293. https://doi.org/10.1111/geoj.12438
Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. University of Chicago Press.
Pilcher, L. A., Darren, L. R., Kgadi, C. M., & 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.
Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *International Journal of STEM Education, 5*(1), 10–14. https://doi.org/10.1186/s40594-018-0115-6
Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2019). A descriptive study of race and gender differences in how instructional style and perceived professor care influence decisions to major in STEM. *International Journal of STEM Education*, 6(1), 1–13. 10.1186/s40594-019-0159-2.
Rambøll (2021). SurveyXact. Retrieved 19.10.2021 from https://www.surveyxact.no.
Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education, 16*(2), 129–150. https://doi.org/10.1080/03075079112331382944
Rattanatumma, 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.
Regional Committees for Medical and Health Research Ethics (REK). (2020). Guidance and Support. Retrieved 20.04.22 from https://www.norcrin.no/en/regional-committees-for-medical-and-health-research-ethics-rek/.
Richardson, T., Elliott, P., & Roberts, R. (2017). Relationship between loneliness and mental health in students. *Journal of Public Mental Health, 16*(2), 48–54. https://doi.org/10.1108/JPMH-03-2016-0013
Rittmayer, A. D., & Beier, M. E. (2008). Overview: Self-efficacy in STEM. *SWE-AWE CASEE Overviews, 1*(3), 12.
Rivera, N. (2013). *Cooperative learning in a community college setting: Developmental coursework in Mathematics*. [Arizona State University]. ProQuest Disertation Publishing.

Sahil, S. A. S., Hassan, H., & Ismail, W. R. W. (2020). Measuring student learning approaches and generic skills: A multi-group analysis of gender differences. *International Journal of Humanities and Social Sciences, 12*(5), 1–12.

Salehi, S., Berk, S. A., Brunelli, R., Cotner, S., Creech, C., Drake, A. G., & Ballen, C. J. (2021b). Context matters: Social psychological factors that underlie academic performance across seven institutions. *CBE Life Science Education, 20*(4), Article ar68-ar68. https://doi.org/10.1187/cbe.21-01-0012

Salehi, S., Cotner, S., Azarin, S. M., Carlson, E. E., Driessen, M., Ferry, V. E., & Ballen, C. J. (2019). Gender performance gaps across different assessment methods and the underlying mechanisms: The case of incoming preparation and test anxiety. *Frontiers in Education (Lausanne), 4*. https://doi.org/10.3389/feduc.2019.00107

Seymour, E., Hunter, A.-B., Laursen, S. L., & Deantoni, T. (2004). Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. *Science Education, 88*, 493–534.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont: Wadsworth, Cengage Learning.

Sivertsen, B. (2021). Students' Health and Wellbeing Study 2021. SiO. https://sioshotstorage.blob.core.windows.net/shot2018/SHOT2021.pdf.

Smith, C., & Bath, D. (2006). The role of the learning community in the development of discipline knowledge and generic graduate outcomes. *Higher Education, 51*(2), 259–286. https://doi.org/10.1007/s10734-004-6389-2

Smith, J. L., Lewis, K. L., Hawthorne, L., & Hodges, S. D. (2013). When trying hard isn't natural: Women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns. *Personality and Social Psychology Bulletin, 39*(2), 131–143. https://doi.org/10.1177/0146167212468332

Stake, J. E., & Mares, K. R. (2005). Evaluating the impact of science-enrichment programs on adolescents' science motivation and confidence: The splashdown effect. *Journal of Research in Science Teaching, 42*(4), 359–375. https://doi.org/10.1002/tea.20052

Stewart-Brown, S., Evans, J., Patterson, J., Petersen, S., Doll, H., Balding, J., & Regis, D. (2000). The health of students in institutes of higher education: An important and neglected public health problem? *Journal of Public Health, 22*(4), 492–499. https://doi.org/10.1093/pubmed/22.4.492

Sæthre, H. Å. (2014). Å tilrettelegge for at studentene skal lykkes: En verktøykasse for utdanningsledere, studieadministrasjon, studenttillitsvalgte [Facilitating student succes: A toolbox for educational leaders, study administration, student representatives]. Haas1.

Taber, K. S. (2016). Learning generic skills through chemistry education. *Chemistry Education Research and Practice Journal, 17*, 225–228.

Thomas, L. (2012). *Building student engagement and belonging in higher education at a time of change*. P. H. Foundation. https://www.heacademy.ac.uk/system/files/what_works_final_report_0.pdf.

Tice, D., Baumeister, R., Crawford, J., Allen, K.-A., & Percy, A. (2021). Student belongingness in higher education: Lessons for professors from the COVID-19 pandemic. *Journal of University Teaching & Learning Practice, 18*(4), 2.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125. https://doi.org/10.3102/00346543045001089

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: The University of Chicago Press.

Trujillo, G., & Tanner, K. D. (2014). Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity. *CBE Life Science Education, 13*(1), 6–15. https://doi.org/10.1187/cbe.13-12-0241

Tynjälä, P., & Gijbels, D. (2012). Changing world: Changing pedagogy. P. Tynjälä, M.-L. Stenström, & M. Saarnivaara. In P. Tynjälä, M. L. Stenström, & M. Saarnivaara (Eds.), *Transitions and transformations in learning and education* (pp. 205–222). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-2312-2_13.

UN. (2015). The 2030 agenda for sustainable development. Retrieved 20.04.22 from https://sdg4education2030.org/the-goal.

UNICEF. (2021). Adolescent education and skills. Retrieved 21.04.22 from https://www.unicef.org/education/skills-development.

Virtanen, A., & Tynjälä, P. (2019). Factors explaining the learning of generic skills: A study of university students' experiences. *Teaching in Higher Education, 24*(7), 880–894. https://doi.org/10.1080/13562517.2018.1515195

Virtanen, A., & Tynjälä, P. (2022). Pedagogical practices predicting perceived learning of social skills among university students. *International Journal of Educational Research, 111*, Article 101895. https://doi.org/10.1016/j.ijer.2021.101895

Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A delicate balance: Integrating active learning into a large lecture course. *CBE Life Science Education, 7*(4), 361–367. https://doi.org/10.1187/cbe.08-02-0004

Watermeyer, R., Crick, T., Knight, C., & Goodall, J. (2021). COVID-19 and digital disruption in UK universities: Afflictions and affordances of emergency online migration. *Higher Education, 81*(3), 623–641. https://doi.org/10.1007/s10734-020-00561-y

Werner, A. M., Tibubos, A. N., Mülder, L. M., Reichel, J. L., Schäfer, M., Heller, S., … Beutel, M. E. (2021). The impact of lockdown stress and loneliness during the COVID-19 pandemic on mental health among university students in Germany. *Scientific Reports, 11*(1), 22637. https://doi.org/10.1038/s41598-021-02024-5

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

Wonch Hill, P., McQuillan, J., Talbert, E., Spiegel, A., Gauthier, G., & Diamond, J. (2017). Science possible selves and the desire to be a scientist: Mindsets, gender bias, and confidence during early adolescence. *Social Sciences, 6*(2), 55. https://doi.org/10.3390/socsci6020055

Yapici, I. U. (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 ontegrated with GeoGebra. *International Electronic Journal of Mathematics Education, 15*(1).

Yu, S., Alper, H. E., Nguyen, A.-M., Brackbill, R. M., Turner, L., Walker, D. J., Maslow, C. B., & Zweig, K. C. (2017). The effectiveness of a monetary incentive offer on survey response rates and response completeness in a longitudinal study. *BMC Medical Research Methodology, 17*(1), 77. https://doi.org/10.1186/s12874-017-0353-1

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230856574 (print)
9788230864388 (PDF)