Journal of Educational Change
https://doi.org/10.1007/s10833-022-09472-4

Shared responsibility between teachers predicts student achievement: A mixed methods study in Norwegian co-taught literacy classes

Aslaug Fodstad Gourvennec[1] · Oddny Judith Solheim[1] · Njål Foldnes[1] · Per Henning Uppstad[1] · Erin M. McTigue[1]

Accepted: 13 October 2022

Abstract

Having two teachers work collaboratively in the same class has been suggested as a possible solution to several instructional challenges, including the inclusion of students with special needs in mainstream classrooms and as part of school-wide prevention models to increase student achievement. In this, shared responsibility between teachers is regarded as a prerequisite to successful co-teaching. However, few studies have investigated whether shared responsibility between teachers actually leads to improved student achievement. This mixed methods study investigates shared responsibility in a sample of 148 classrooms where two general educators worked collaboratively in literacy instruction through first and second grade. First, we analysed whether the degree of shared responsibility between the two teachers for planning, enacting, and evaluating literacy instruction predicted student reading when controlling for pre-reading skills at baseline. Second, we carried out in-depth individual interviews with six collaborative teacher dyads purposefully selected from high- and low-performing classrooms to investigate what characterized their sharing of responsibility. The results show that shared responsibility significantly predicts students' reading achievement. Further, the interviews reveal a surface level collaboration between coteachers in low-performing classes, yet a more profound level of collaboration with influence on key teaching decisions in high-performing classes.

Keywords Co-teaching · Early literacy instruction · Mixed methods · Teacher collaboration · Student–teacher ratio

Aslaug Fodstad Gourvennec and Oddny Judith Solheim have contributed equally to this work.

✉ Aslaug Fodstad Gourvennec
aslaug.f.gourvennec@uis.no

Extended author information available on the last page of the article

Published online: 08 December 2022

Rationale

Any one teacher must feel safe and empowered to perform optimally in instruction and facilitate students' academic growth. With experience, most individual teachers find strategies to navigate these needs. However, when two teachers are set to collaborate about instruction in the *same* class[1] (i.e., co-teaching), established classroom practices can be challenged and full engagement may falter. In our work, co-teaching is defined broadly and describes cooperation among two or more professionals with similar or different formal competencies, which allows for shared responsibility (Conderman, 2011; Krammer et al., 2018). This collaborative situation demands new decisions regarding instructional choices, and it affects the conditions under which each individual teachers' core needs of autonomy, relatedness, and competence are satisfied (Alexander, 1997; Deci & Ryan, 2008; Pesonen et al., 2020); therefore, it is a challenging situation for teachers to enter and worthy of nuanced investigation.

Co-teaching has been suggested as a possible solution to several instructional challenges – in particular the inclusion of students with special needs in general education. Having two teachers work collaboratively in the same class has also been suggested as part of school-wide prevention models to increase student achievement, particularly in reading and mathematics. Such models allow for increased individualized attention to students, additional teacher time, and a reduction in the student-teacher ratio (STR) in particular subjects (Andersen et al., 2020; Solheim & Opheim, 2018; Friend, 2021). Despite much optimism and promise, however, the results are often inconclusive, yielding few certainties from which to make policy. As co-teaching continues to hold enthusiastic support in many educational communities, regardless of the recognized cautions within the research base (Jones & Winters, 2022), it is imperative that we better understand this educational intervention. Rather than framing research in binary terms – for example, "Does co-teaching benefit students?" – we need to acknowledge that we already know that sometimes it works, and sometimes it does not. Hence, we need to reconceptualize the approach taken and pose more nuanced questions, such as: "Under what conditions does co-teaching most benefit students?" and "What characterizes the most successful co-teaching relationships?".

A key assumption in discussions about the relationship between co-teaching and student outcomes is that it enables instruction more conducive to student learning (e.g., more varied and differentiated instruction) (Krammer et al., 2018; Solheim et al., 2017). Having two teachers work collaboratively in the same class offers multiple opportunities for more interaction and flexible ways of organizing instruction, for example by enabling teachers to switch between whole-group instruction, work in groups or pairs, and one-to-one teaching. Further, the collaborative nature of the arrangement allows for shared knowledge construction (Rytivaara & Kershner, 2012; Rytivaara et al., 2021) and enables peer modeling and mentoring where two professionals can learn from each other (Johnston & Tsai, 2018).

[1] "Class" refers to a group of students taught together.

Still, the benefits of collaborative teaching, like those of class-size reduction, do not always materialize (Weiss & Brigham, 2000), and previous research into co-teaching has highlighted the importance of shared responsibility or parity. Whereas these concepts are rarely defined in the literature, several aspects of what they imply are presented. First, Alexander (1997) found that in order to establish effective partnerships, both educators need to be actively engaged in the planning and teaching of each lesson (Alexander, 1997). In line with this, but adding to the picture, Friend (2021) notes that to achieve parity in the collaboration between educators with different formal competence, co-teachers need to share responsibility for participation and decision making. Friend also emphasizes the importance of their mutual recognition of bringing "different but equally important knowledge and skills to their shared classroom" (Friend et al., 2015, p. 84). Unfortunately, research validates what many co-teachers actually experience-a lack of parity (Karten & Murawski, 2020).

Several challenges for establishing parity have been identified. At the level of the individual, according to Alexander (1997), teachers' notions of territory, ownership, and autonomy may undermine collaboration. At the level of the partnership, collaborating teachers often have difficulty defining and maintaining classroom roles in which they both contribute meaningfully to the instruction (Pratt, 2014; Karten & Murawski, 2020). Further, differences in personality, communication, conflict styles, gender, etc., can create tensions, as teachers in co-teaching arrangements tend to believe that similar learning philosophies and classroom-management styles are necessary for an effective co-teaching relationship (c.f., Pratt, 2014). Furthermore, external factors at the school level, such as a lack of shared planning time, physical distance of working spaces, and incompatible schedules, present practical challenges to the most compatible of teaching partners and require administrative support (Karten & Murawski, 2020; Main, 2012). Finally, the customs and patterns of the school culture (e.g., teaching in isolation) can present an obstacle for teams' effectiveness (Main, 2012).

Unfortunately, it must be acknowledged that the importance of parity/shared responsibility in co-teaching is more supported by theory and anecdote rather than systematic research, particularly research using mixed methods and large samples. While often assumed to be true, the question of whether parity and shared responsibility impact the actual goal of increasing student achievement is not yet empirically established. Several research overviews have pointed to a lack of efficacy research and characterized the body of research as limited (e.g., Friend et al., 1993; Jones & Winters, 2022; Murawski & Swanson, 2001; Weiss, 2004; Weiss & Brigham, 2000), or even anecdotal (Murawski & Swanson, 2001).

Contribution of this work

In this mixed methods study, we address a central gap in the research base by investigating how the degree of shared responsibility is associated with student achievement in a large sample of classes where two general educators (a "dyad") co-taught literacy instruction through first and second grade. We define the concept of shared responsibility as collectively taking responsibility for the whole class throughout all

phases of the instruction, including the planning, enacting, evaluating, reflecting, and adjusting instruction. This responsibility includes an active engagement in the collaborative project. As a non-example, when a co-teacher simply assists in an ongoing activity, without making instructional decisions about the activity, they are contributing to the productivity of the classroom but are not sharing the responsibility. However, as long as daily activities and decisions are connected to long term co-constructed goals, shared responsibility may be well maintained even when two teachers divide the class in two, or one of them teaches individual students.

The study was carried out as part of a large randomized controlled trial (RCT), *Two teachers*, which investigated the effect of STR on student outcomes (Haaland et al, 2022; Solheim et al., 2017). The RCT yielded quantitative results for student achievement (i.e., reading) and teachers' self-reported collaboration for a large number of teacher dyads. To allow for a more nuanced understanding of the quantitative findings, we selected six extreme dyads based on student achievement post intervention. Specifically, we studied dyads who had taught classes demonstrating either very strong or weak progress in reading – referred to below as low- and high-performing classes.

As is consistent with mixed methods research, which combines multiple paradigms, different theories informed our work at each level of analysis. Because our focus was primarily on the teachers' experience of their new professional situation (co-teaching), we drew upon adult-learning and motivational theories, specifically *Self-Determination Theory* (SDT) (Deci & Ryan, 2008). SDT considers how social and cultural factors promote or undermine a person's motivation, directly influencing self-concept and performance. It posits that people can act with volition only if they experience *autonomy*, *competence*, and *relatedness*. When those needs are met, people are internally motivated, act with engagement, and exhibit higher levels of performance.

For characterizing teachers' collaboration, we drew upon Pratt's (2014) *Achieving Symbiosis Theory*. Based on interviews with and observation of secondary-school co-teaching teams characterized as *effective*, this theory describes how collaborative teachers create an effective partnership through the stages of (1) *initiation*, (2) *symbiosis spin*, and (3) *fulfillment*. At the fulfillment stage, both teachers are "fulfilled professionally and personally in their collaborative teaching relationship and actively involved in instruction" (p. 10). This stage is characterized by (i) valuing of the relationship, (ii) smooth handling of challenges, (iii) seamless instruction, (iv) presence of all necessary dimensions (external and internal), (v) reflection, and (vi) compatibility. Based on the above, this study seeks to contribute to the research base on co-teaching by investigating how the teacher dyads' degree of shared responsibility is associated with student achievement. We examine this by addressing the following research questions:

1. To what extent does the degree of shared responsibility for planning, enacting, and evaluating literacy instruction predict student reading in Grade 2 when controlling for pre-reading skills at school entry?
2. How do the two teachers in a dyad, in high- or low-performing classes, characterize their sharing of responsibility?

Methods

Contextual frames for the study

The Norwegian educational system is founded on the principle of a unified school system that provides equal and adapted instruction in an inclusive environment based on a single national curriculum (Ministry of Education & Research, 2020). Within the framework of statutes and national curricula, teachers are free to choose teaching methods (i.e., the curricula formulate aims for student learning but there are no expectations for specific teaching methods). Relating to STR, at the time of the Two teachers intervention, The Education Act (1998) only stated that municipalities had to ensure particularly low STR in Grades 1–4 in Norwegian (mother tongue) and Mathematics. Schools were free to organize the student group aligned to their judgement, for example deploying more than one teacher in the class in one or more subjects.

The Two teachers study represented a major collaboration with the Norwegian Ministry of Education where participating schools received funding for an additional teacher (i.e., with an educational degree) in Norwegian lessons for 8×45 min a week (about 32% of the weekly instruction) through first and second grade. The obligation of an additional teacher was met by integrating the eight teacher hours into the schedule of a teacher already working at the respective school. School leaders were asked about their main emphasis when selecting the additional teacher, and 35.1% reported experience in literacy instruction, 22.3% reported having an education in literacy instruction, and 19.6% reported experience in teaching Grades 1–4.

Upon entering the intervention (the Two teachers project), all participating teachers were assigned the title of either *homeroom teacher* or *co-teacher*. These titles, or surface-level roles, come endowed with predefined tasks and responsibilities as well as undefined spaces to be occupied. Specifically, in a Norwegian context, homeroom teachers are assigned overall responsibility for their students, for collaboration across the teacher team at the grade level in question, and for communication with parents and the school administration. In Grades 1 and 2, the homeroom teacher usually spends most of the school day with their class, being responsible for the instruction in all, or most of, the subjects. These responsibilities are assigned by tradition, by the individual school's established guidelines, and – partially – by Norwegian law. The co-teacher's role, by contrast, lacks both guidelines and tradition. As participating teachers did not receive any professional development related to co-teaching in general, or shared responsibility specifically, the teachers in the 148 dyads were tasked with both interpreting and occupying this role, as well as establishing how the roles related to each other.

Schools were free to choose how the two teachers would work together. In a teacher questionnaire administered towards the end of the intervention teachers self-reported on their main organizational approaches. While a variety of approaches were reported (e.g., divide the group in two and teach one group each or make one teacher teach the class while the other helps where needed), one organizational form predominated, i.e., one teacher taught the class while the other worked with either

a small group or with individual students. Specifically, about half of all homeroom teachers reported that one teacher taught the class while the other worked with a smaller group or individual students for around half of their co-taught lessons.

Design

In the present study we employed an explanatory sequential design (Creswell & Clark, 2017) (see Fig. 1 for summary and integration points) with two phases beginning with quantitative analysis and shifting to qualitative analysis. The quantitative phase served two purposes. First, we investigated the predictive value of shared responsibility on students' achievement. Second, teaching partners who were, or were not, particularly effective were identified. These partnerships were then explored in the next phase through qualitative methodologies (interviews and thematic coding) to explore potential explanations as to co-teachers' effectiveness in facilitating student achievement.

More specifically, we first collected data on students' reading before and after the intervention. Second, we distributed surveys to the homeroom teachers, who self-reported on the extent to which the two teachers shared responsibility for planning, enacting, and evaluating instruction. We then analyzed whether the degree of shared responsibility predicted student reading in Grade 2 when controlling for pre-reading skills at school entry. Further, we selected three high-performing classes and three low-performing classes. To understand the experience of collaborative teaching, we conducted in-depth individual interviews with both teachers in each of those six classes. The interview data were analyzed thematically. Finally, the results of the quantitative and qualitative analyses were integrated into a final conclusion.

Sample

A total of 150 schools in 53 municipalities across 9 counties in southern Norway were enrolled in Two teachers in 2016. Two incoming first-grade classes in each school (N = 300) were randomly assigned to either a treatment or control condition. 100 of the 150 participating schools were also randomized to receive teacher professional development (TPD) in literacy instruction (i.e., the TPD did not explicitly focus on co-teaching).[2] The present study only includes co-taught (i.e., treated) classes. As two schools decided to withdraw from the study after Grade 1, our final sample consists of 148 classrooms. Class size varies between 13 and 30 students (median = 21). A total of 2685 students are included (student participation rate = 95.8%). Average age at school entry was 6.2 years (47.1% girls).

The teacher sample in the quantitative analysis includes the 148 homeroom teachers, of whom 98% were female; 11% were 29 years old or younger, 24% were 30–39, 35% were 40–49, 22% were 50–59, and 6% were over 60 years old. Most (89%) held

[2] We collapse the different co-taught conditions into one sample as initial analyses showed no differences in how teachers self-reported on the survey.

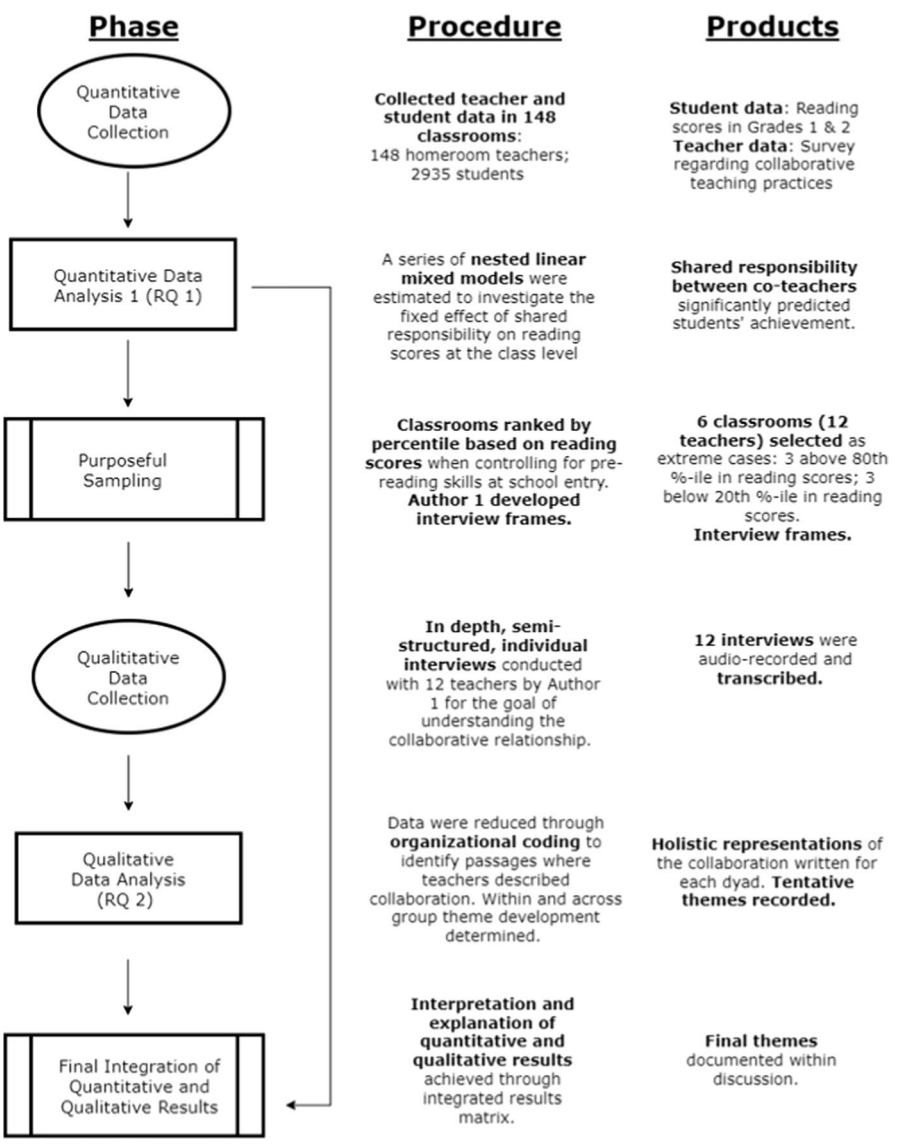

Fig. 1 Methodological flow chart documenting data collection and integration points

a bachelor's degree; 3% held a master's degree. Teaching experience ranged from 1 to 40 years, with a mean of 13 (Standard deviation (SD) = 8). Experience of providing early literacy instruction ranged from 1 to 20 years, with a mean of 4 (SD = 3).

The teacher sample in the qualitative analysis includes twelve teachers composing six co-teaching dyads (see Table 1). Three dyads taught in classes with reading scores above the 80th percentile when controlling for pre-reading skills at school

Table 1 Overview of participants in the interviews

	Teacher dyad	Teacher names[a]	Age[b]	Years of experience	Class size
Low- performing classes	B1	Bridget and Arthur	40–49 30–39	16 6	20
	B2	Brenda and Anna	30–39 25–29	10 6	23
	B3	Beth and August	40–49 25–29	14 3	21
High- performing classes	T1	Tania and Sara	30–39 40–49	14 17	21
	T2	Tina and Sofia	40–49 40–49	20 15	25
	T3	Theresa and Stella	40–49 40–49	12 20	25

[a]All teacher names are pseudonyms

[b]At the end of the Two teachers intervention

entry.[3] The other three dyads taught in classes with such scores below the 20th percentile. Hence the teachers can be seen as representing extreme cases. Extreme cases are often dense in information and may thus offer insights into both especially beneficial and especially problematic aspects of teachers' collaboration (Flyvbjerg, 2006, p. 13).

To select dyads for interviews, we first identified classes above or below the cut-off points, respectively. Then contact was made with teachers regarding willingness to participate, and a final selection was made, comprising one teacher dyad above and one below the cutoff points for each of the three different co-taught conditions. When possible, an attempt was made to avoid geographical concentration. The participating teachers are from six schools in five municipalities – both urban and rural – in four different counties. All six homeroom teachers (Bridget, Brenda, Beth, Tania, Tina, and Theresa) are female and relatively experienced (13–21 years' working experience). Among the co-teachers, there are two males (Arthur and August) and four females (Anna, Sara, Sofia, and Stella), representing a large span of teaching experience (3–20 years). Two co-teachers (Arthur and Anna) were only part of the dyad during the second year of the intervention, while the other four dyads worked together for two years.

Measures

Our mixed methods study relies on data from three sources: (1) assessment of students' reading before and after the intervention, (2) a survey answered by 148

[3] From the fitted model M_1 we extracted the conditional means of reading scores (the conditional modes of the random intercepts c). Classes whose conditional class mean were below the 20th percentile, or above the 80th percentile, were considered.

homeroom teachers toward the end of second grade, and (3) individual interviews with the two teachers in six co-taught classes conducted in January and February 2019.

Student level measures

Students were assessed (i) at baseline, i.e., school entry in first grade (August/September 2016), and (ii) immediately post-intervention, i.e., at the end of second grade (May/June 2018). On both occasions trained and certified research assistants administrated the assessment (see Table 2 for description of measures). Students were assessed individually at their respective schools. The baseline measures were administrated on tablet computers.

Teacher level measures

Homeroom teachers were asked about the extent to which they shared responsibility for planning, enacting, and evaluating literacy instruction with their co-teacher. Specifically, we asked teachers to choose one of three options to characterize their collaboration. The options reflected a progressively greater sharing of responsibility for planning and teaching: (1) "I have the leading role when it comes to planning, teaching, and evaluation. My co-teacher helps where needed" ("Assisting"); (2) "I have the leading role when it comes to planning, teaching, and evaluation. I give my co-teacher distinct tasks for each lesson" ("Delegating"); and (3) "We make plans, teach, and evaluate together, and take equal responsibility for all students in the class. We share the management role and the attendant obligations'' ("Co-operating"). The online survey had a response rate of 100%.

Statistical analysis

To explain variation in student achievement at the end of the second grade, a series of nested two-level linear mixed models were estimated, with the dependent variable being a standardized (zero mean, unit variance) combination of the word-reading fluency and reading comprehension measures. All models contained a random intercept and fixed slopes. To further ease interpretation and comparison among predictors, the baseline student measures (Letter knowledge (LK), Phoneme isolation (PI), Phoneme blending (PB), Word reading (WR), Rapid Automatized Naming (RAN), Vocabulary (VOK), Digit span (DS), and Numeracy (NUM) were also standardized. The classroom-level variable was degree of shared responsibility. In our modeling, at the student level, we included the baseline reading measures in a vector X_i for student *i.* At the classroom level, in classroom *c,* shared responsibility was encoded in vector of dummy variables S_c. The following sequence of increasingly complex models was estimated for $Y_{i,c}$, the standardized outcome in reading (word reading and reading comprehension at the end of Grade 2) for child *i* in classroom *c*:

Table 2 Overview of student level measures

Measure	Task	Procedure	Number of items	Alpha[a]
Baseline measures				
Letter knowledge	Match a letter sound to a written letter	Students heard a pre-recorded letter sound and responded by pressing one of four letters appearing on the screen	24	.90
Phoneme isolation	Isolate and pronounce the first sound of a spoken word[b]	The tester pointed to a picture, named the object and asked the student to give the first sound of the word, e.g. 'This is a ball. What is the first sound in *ball*?')	10	.93
Phoneme blending	Blend a sequence of phonemes (2–7) into a word[b]	Students were asked to identify words that were presented phoneme by phoneme (one per second), e.g. "Here you see pictures of /ri/, /rips/, /ris/, and /ring/ Listen carefully and press the picture that goes with: /r/ /i/ /s/". Students responded by pressing one of the four pictures	10	.91
Word reading	Read words orally[b]	Students were asked to read orally words that were presented on the screen, one by one	8	.93
Rapid automatized naming	Timed naming of familiar objects (Solheim et al., 2018)	Twenty stimuli were presented in a 4 X 5 matrix. Students were asked to name each stimulus as quickly and accurately as possible	2 trials	
Vocabulary	Norwegian Vocabulary Test (NVT) (Størksen et al., 2013)	Students were presented with different pictures on a tablet screen and had to tell the name of the object depicted	20	.82
Short term memory	Digit Span Forward Test from the Wechsler Intelligence Scales for Children, 3rd ed. (Wechsler, 1991)	The tester orally presented a sequence of digits and students were asked to repeat the sequence of digits in the same order	16	
Early numeracy skills	Four items from "Ani Banani Math Test" (ten Braak & Størksen, 2021) Three items from "First Step Study" (Zhang et al., 2014)	Students were asked to count forward and backwards, e.g. "Can you count backwards from 12?" Students responded orally	7	.71
Post-intervention measures				

Table 2 (continued)

Measure	Task	Procedure	Number of items	Alpha[a]
Word reading fluency	Test of Word Reading Efficiency (TOWRE) (Torgersen et al., 1999)	Students were given a list of printed words and told to read aloud as many of them as possible in 45 s	1 trial	
Reading comprehension	Neale Analysis of Reading Ability (NARA) (Neale, 1997).[b]	Students read short passages and answered comprehension questions orally	36	.88

[a]Alpha = Cronbach's internal consistency measure

[b]Items/texts were ordered by difficulty and the test was discontinued after set criteria

$$Y_{i,c} = \alpha_c + \varepsilon_{i,c} \quad (M_0)$$

$$Y_{i,c} = \beta_X X_i + \alpha_c + \varepsilon_{i,c} \quad (M_1)$$

$$Y_{i,c} = \beta_X X_i + \gamma_S S_c + \alpha_c + \varepsilon_{i,c} \quad (M_2)$$

In the baseline model $M_{0,}$ only the variance of the random intercept is estimated, resulting in an estimate of the intraclass correlation coefficient (ICC). Model M_1 introduces the baseline individual student scores. The final model M_2 is obtained by further including a class-level fixed effect of shared responsibility, γ_S. To compare the explanatory power of the models, we inspected information criteria such as AIC and BIC, as well as conducting nested mixed modeling anova. All analyses were performed in the *R* computing environment (R Core Team, 2020) using package lme4 (Bates et al., 2015; Kuznestova et al., 2017).

Interviews

Individual, in-depth, semi-structured interviews with the twelve teachers (six co-teaching dyads) were conducted by author 1 in January and February 2019, 7–8 months after the end of the intervention. The interviews took place in a quiet room at the teachers' schools. The interview guide (cf. appendix) addressed the themes of (i) literacy instruction and teachers' and students' roles, (ii) co-teaching/the Two teachers project and (iii) teachers' understandings of students' literacy development. Each theme was introduced by at least one open-ended question. In the present study, we mainly address the second interview theme. However, given the semi-structured design of the interviews, allowing the interviewees' narrations to unfold, each theme is informed by teacher utterances from across the initiated themes. The interviews were audio-taped and transcribed in accordance with a simplified version of Jefferson's transcription key (cf. Atkinson & Heritage, 1999).

The transcription of the interviews represents a first data reduction. The data were further reduced through organizational coding (Maxwell, 2009) to identify passages where the teachers described their collaboration, explicitly or implicitly. Finally, holistic representations of the collaboration between the two teachers in each of the six dyads were written through a hermeneutically-driven close-reading of both previous reductive steps (transcripts and coded sections in individual interviews). This final step was conducted separately by author 1 and 4 for one of the dyads in order to check for consistency. The findings reported in the article draw upon the holistic representations.

Table 3 Correlations between student level measures

		LK	PI	PB	WR	RAN	VOC	DS	NUM
Baseline	LK	–							
	PI	.58	–						
	PB	.51	.64	–					
	WR	.62	.69	.8	–				
	RAN	-.29	-.27	-.25	-.29	–			
	VOC	.36	.38	.33	.36	-.27	–		
	DS	.28	.3	.34	.31	-.16	.27	–	
	NUM	.53	.49	.45	.52	-.39	.41	.32	–
Grade 2	Reading	.43	.39	.38	.45	.39	.34	.28	.49

Results

Research question 1

First, we investigated whether the degree of shared responsibility for planning, enacting, and evaluating literacy instruction predicted student reading in Grade 2 when controlling for pre-reading skills at baseline.

Descriptive statistics

At the student level, we standardized all baseline measures as well as the dependent variable, i.e., the reading score at the end of the second grade. The correlations among these numeric variables are given in Table 3. In the total sample of 2890 students, 59 missed scores at baseline and 205 missed reading scores at the end of second grade (7.1%). We have no reason to believe that the mechanism for missingness is related to any of the class level predictors employed in the present study. For instance, a chi-square test of the association between missingness in reading scores and shared responsibility did not lend support for an association (chi-square = 0.91, df = 2, p-value = 0.63).

Classroom variables were teacher-reported. About half of the homeroom teachers reported equal collaboration while half reported more hierarchical arrangements: Of the 148 teachers, 52% (n = 77) said that they shared responsibility for all students in the class equally with their co-teacher, including planning, teaching, and evaluating together (Co-operating); 26% (n = 39) answered that they had the leading role and delegated specific tasks for each lesson (Delegating); and 22% (n = 32) claimed to have the leading role, with their co-teacher helping as needed (Assisting).

Table 4 Anova table for nested sequence of models

	npar	AIC	BIC	logLik	Chisq	Df	p-value
M_0	3	7316.34	7333.95	-3655.17	4		
M_1	11	6194.56	6259.10	-3086.28	1137.79	8	0.00
M_2	13	6192.14	6268.42	-3083.07	6.41	2	0.04

npar Number of parameters. *AIC* Akaike information criterion. *BIC* Bayesian information criterion. *logLik* Log-likelihood. *Chisq* Chi-square of nested model testing. *Df* Degrees of freedom for nested model testing. *p*-value = *p*-value of nested model test

Table 5 Regression coefficients and standard errors for linear mixed model M_2 explaining variation in second-grade reading scores

	Model 2
(Intercept)	−0.11*
	(0.05)
LK	0.11***
	(0.02)
PI	0.00
	(0.02)
PB	0.00
	(0.03)
WR	0.15***
	(0.03)
RAN	−0.18***
	(0.02)
VOC	0.08***
	(0.02)
DS	0.08***
	(0.02)
NUM	0.23***
	(0.02)
Delegating	0.07
	(0.07)
Co-operating	0.15*
	(0.06)

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. Standard errors in parentheses

Regression models

The baseline M_0 yielded an estimate of the random intercept variance equal to 0.073. That is, the intraclass correlation coefficient of student reading scores with respect to class was 0.073, so that 7.3% of the Grade 2 student reading score variation resulted

from between-class variation, and the remaining 92.7% from within-class variation. Anova outputs of models M_0-M_2 are given in Table 4.[4]

It was seen that M_1 was preferable to M_0 in terms of model fit as assessed by AIC and BIC. Therefore, the inclusion of baseline individual pre-reading scores as predictors in M_1 was a highly significant improvement with respect to the baseline model. When also adding shared responsibility to M_1, the resulting model M_2 did further improve model fit, as reflected both by the information criteria, and the formal test of model fit. The regression results for M_2 are presented in Table 5. As expected, most baseline measures are significant predictors of second grade reading ability. Looking at shared responsibility, we see that both Delegating and Co-operating were associated with an increase in student reading at the end of second grade, relative to the reference category Assisting, while controlling for the individual baseline assessments. However, it was only the gap between Co-operating and Assisting that was statistically significant. The effect size associated with Co-operating versus Assisting was 0.15, when taking baseline measures into account.

Research question 2

As hypothesized, shared responsibility between the teachers was found to significantly predict students' achievement. In the qualitative analysis we explored teachers' collaboration in depth by asking the homeroom teacher and co-teacher in six extreme dyads to characterize their sharing of responsibility.

In the following sections, we first examine how the dyads in the low- and high-performing classes, respectively, characterized their shared responsibility, before broadening our field of vision to present their overall perception of their collaboration.

Distribution of responsibility

Throughout the interviews, the distribution of responsibility between the two collaborating teachers is characterized at two different levels: (1) at a surface level, related to the formal distribution of specific tasks or areas, and (2) at a more profound level, related to influence on key teaching decisions, which may not have been formalized.

[4] In the initial analysis we included class size and teaching experience (number of years) as predictors. As these variables neither came out as significant, nor improved model fit they were excluded for reasons of parsimony.

Low-performing classes

According to the teachers' descriptions of their collaboration, all three dyads in low-performing classes had an unequal distribution of responsibility. At the surface level, related to concrete tasks, this is quite evident: the homeroom teacher has "a more overall perspective" (Beth, B3, p. 46) and the main responsibility for "Norwegian lessons" (Arthur, B1, p. 6)[5] and the plenary sessions and/or the largest student group during instruction. The co-teacher, on the other hand, takes the main responsibility for the follow-up of struggling readers, individually or in small groups. For Bridget's (B1) and Brenda's (B2) dyads, instruction is mainly divided in advance, entailing a division of responsibility. Beth's (B3) dyad differs on this point: the two teachers give the impression that Beth evaluates the situation during instruction on an ongoing basis, deciding whether it allows August (B3) to take some of the struggling readers out of the classroom. August refers to this approach as "a bit random" (pp. 53–54).

August is also given responsibility to carry out reading conferences, a requirement under the overall RCT. Interestingly, whereas he found these conferences informative regarding the students' reading, this information was not discussed between the two teachers. Beth (B3) explains her lack of interest in the reading conferences by saying that she already knew the students' reading skills very well and did not believe she would learn anything more from the conferences: "maybe August did […]. But I believe I know ((laughs))" (Beth, p. 38). Hence August seems to have been responsible for a required task to which the homeroom teacher ascribed no pedagogical significance.

The unequal distribution of responsibility in the low-performing classes is also reflected in the teachers' characterization of their instructional planning. In all three dyads, the homeroom teacher has the final say in planning. Interestingly, both Beth's (B3) and Brenda's (B2) dyads initially describe planning in positive terms. However, it gradually emerges from these two dyads' descriptions that equality in planning refers to an equal willingness to contribute, to time spent, and to administrative tasks performed, not to actual influence on instructional content and methods. The reasons given for the homeroom teacher's dominance over planning differ between these two dyads. Beth implicitly justifies her domination by the difference in age and experience: "I was so much older and was in possession of so many things" (Beth, p. 50). By contrast, Anna (B2) ascribes significance to the physical arrangement of workplaces, explaining her lack of influence on planning with the fact that Brenda (B2) and she did not share offices. Brenda had her office next to that of the homeroom teacher of another class in the same grade, and they had a very close collaboration. Anna mentions that she repeatedly suggested collaborative planning, only to find that the two homeroom teachers had already made plans, feeling as though "they did the main work and that [she] just did some polishing" (p. 30).

[5] All quotations have been translated from Norwegian by author 1. The labels B1–B3 and T1–T3 refer to the dyad the teacher in question belongs to (cf. Table 1). The page numbers given refer to the individual interview transcripts, which comprise the following numbers of pages: Bridget 46, Arthur 35, Brenda 26, Anna 53, Beth 72, August 62, Tania 38, Sara 29, Tina 54, Sofia 44, Theresa 43, and Stella 56.

In the third low-performing class (B1), planning was not presented as a collaboration between the two teachers. Instead, Bridget and Arthur report that planning took place either with all teachers at the same grade level or individually for the groups of students for which they were responsible (i.e., the whole class versus individual struggling readers). This division of planning tasks is presented as a choice made in order to avoid too close a collaboration, as problematic aspects of the relationship between Bridget and Arthur escalated. One point of disagreement was whether Arthur was allowed to communicate with parents without Bridget's consent. Arthur considered it very problematic that Bridget deprived him of this opportunity, whereas she could "speak freely" (Arthur, p. 19) with them. For her part, Bridget refers to the school's guidelines, according to which the homeroom teacher is responsible for the class and thus in charge of communicating with parents.

High-performing classes

At the surface level, the dyads in the high-performing classes report a distribution of responsibility that resembles that of the low-performing classes: the co-teachers have particular responsibility for struggling readers or more "immature students" (Sofia, T2, p. 19), whereas the homeroom teachers tend to lead plenary sessions or larger student groups. During instruction, however, this division is not as rigid. For instance, although Sara (T1) has the main responsibility for letter instruction and for individual tutoring while Tania (T1) takes the main responsibility for the plenary sessions and the "big picture" (Tania, p. 22), they share responsibility during the group's "morning ritual," both working individually with students and taking on identical tasks. Hence the impression is that of seamless integration of instruction during lessons.

Indeed, a clear division of responsibility by domain seems to exist only at the surface level in the high-performing classes – higher-level decision-making during planning was shared for all three dyads. All six teachers emphasize that they "discussed all the time" (Sara, T1, p. 17) and that an ongoing shared reflection – about how to best exploit the enhanced teacher resource when it came to instructional organization, teaching plans, students' skills, and social development in order to meet every student's needs – was at the core of their collaboration.

This emphasis on shared reflection implies that the co-teacher was as engaged as the homeroom teacher during both planning and instruction – the co-teacher never acted as a "radiator heater" (i.e., just sitting on the radiator) (Sara, T1, p. 27) or as an "assistant teacher" (Stella, T3, p. 31). Even at the surface level of the division of responsibility, both Sofia (T2) and Stella were given typical homeroom-teacher tasks (e.g., communicating with parents, attending social events, managing class-council meetings). Theresa (T3) says she believes the students perceived Stella as their teacher, meaning that the students did not perceive the homeroom teacher as more important than the co-teacher. From the teachers' perspective, shared detailed knowledge about the students seems to have enabled them to share responsibility at a more profound level; Stella emphasizes the great benefit of "not being alone with this serious responsibility: teaching children how to read and write" (p. 30).

Perception of the collaboration

We now broaden our field of vision, looking at the dyads' overarching characterizations of their collaboration. Throughout their descriptions of their collaboration, more or less implicit evaluations of it draw attention to aspects that they deemed to promote or hinder shared responsibility.

Low-performing classes

All three dyads in low-performing classes had experienced negative challenges regarding their collaboration. However, the aspects considered challenging varied, as did the intensity of those challenges.

The most explicit description of a problematic collaboration comes from Bridget and Arthur (B1). Their accounts of a collaboration, developing from difficult to very problematic, are consistent with each other; they eventually gave up any real attempts to collaborate. Although this resignation is explicitly addressed by both teachers, they explain it differently. According to Arthur, differences regarding pedagogical questions – such as how to manage externalizing children – were at the core of the conflict. From his perspective, there was no room for common reflection on such issues as their collaboration problems escalated; Bridget seemed to avoid him whenever he asked for a discussion about something that he found problematic during instruction. Arthur describes his struggle to "advocate [my] own perspective and ways of doing things" (p. 16) as "so stressful" that after several months of struggling "I pulled back a bit more and then I thought, OK, it's probably just for another six months, so I'll just do my bit and that's it" (Arthur, p. 16). Bridget, on the other hand, makes it clear that she perceives this resignation as a kind of personal defeat. Somewhat ashamed in retrospect, she characterizes her inability to disregard her initial impression of Arthur, as a know-it-all despite his limited teaching experience (compared with hers), as "almost childish" (Bridget, p. 32). The ensuing lack of real collaboration seems to be emphasized by Arthur's discursive distancing when talking about what took place during their year of working together. For instance, when reporting on instructional organization, he says that "Bridget was of the opinion … Well, she was the one telling me how it should be" (p. 24). Further, he emphasizes their different opinions on the limits to the co-teacher's opportunities for communication with students' families.

Beth and August (B3) have a very different starting point, explicitly referring to their dyad as involving "one hundred percent super" (Beth, p. 52) chemistry between the two of them, including "a lot of humor" (August, p. 42). However, this very positive picture becomes more complex as both teachers describe their struggle trying to find a good way to collaborate "that closely" (Beth, p. 41), and August admits that it "has actually been very difficult to figure out […] for two years" (p. 25) what role to take on within the collaborative dyad. Throughout both interviews, it is clear that Beth, as the experienced teacher, took the lead, whereas August, starting his teaching career, followed her lead and asked to help where needed in any given situation. These stances seemed to perpetuate a cycle leading to greater disparity: although Beth emphasizes her willingness and efforts to ask for August's opinions and ideas

during planning, August emphasizes both his uncertainty and his reluctance to insist on his own opinions and ideas. Further, he reports how his uncertainty increased because of his perception of Beth's attitude – communicating to him that she always knew best. Ultimately, they both question whether the collaboration was worthwhile: August goes so far as to state that he "would prefer not to be two teachers again" (p. 28), whereas Beth characterizes their collaboration as rather inefficient, noting that dividing the tasks more clearly might have been better than doing "a lot of things in parallel" (Beth, p. 51).

Brenda and Anna (B2) mainly describe their collaboration in positive terms, and they clearly appreciate each other as persons, both at work and outside. Nevertheless, they both reflect on the potential benefits of closer collaboration within the dyad. Brenda notes that it was probably difficult for Anna "to be number two" (p. 24) and that, if engaging in similar co-teaching arrangements in the future, she would strive for greater equality between the two teachers. Anna gives a glimpse of what closer collaboration could have looked like when referring to occasional instances when they took the time to evaluate the instruction: "I remember I sometimes thought that,'Oh, yeah! This is really nice'" (Anna, p. 27). She clearly appreciated these occasions, which she perceived as "a kind of peer mentoring" (p. 27).

While Brenda and Anna clearly value their collaborative relationship, they would have liked to be more equal. Both identify an external factor – the distance between their workspaces – as hindering closer collaboration. The situation is somewhat different for the other two dyads. Beth and August (B3) value each other as persons but ascribe limited value to their professional collaboration. Finally, Bridget and Arthur (B1) are clear that their relationship was destructive, emphasizing pedagogical differences, communicational challenges, and personal traits as aspects hindering collaboration.

High-performing classes

All three dyads in high-performing classes clearly value the teacher collaboration, describing the development of "a very close collaboration" (Tina, T2, p. 22) as well as very good chemistry. Theresa (T3) refers to the chemistry with her colleague as a "match made in heaven" (p. 30), whereas Tina underlines that the good chemistry "makes it a lot more fun to work together" (p. 33).

At the same time, the teachers in all three dyads explicitly emphasize that they differ strongly from the other teacher in personality. In other words, their close collaboration does not depend on similarity. Examples of differences mentioned include Sara's (T1) reference to different preferences in terms of classroom rules and Tina's (T1) reference to Sofia's creativity and her own tendency to take time to ponder new ideas as well as valuing the ritual aspects of well-known activities. However, rather than presenting their differences as problematic, these dyads refer to open dialogue about disparate preferences and potential disagreements. What is important to recognize here is that the differences are ascribed to personality traits, teaching styles, or general preferences, whereas more fundamental pedagogical beliefs seem to be shared within the dyad – they "essentially agree about […] what enhances learning" (Theresa, T3, p. 29).

Another characteristic of these three dyads is how they frame their differences in experience, knowledge, or competence as valuable for their professional development. Tania (T1) stresses that collaborating with Sara (T1) was very valuable, not so much because of her own lack of elementary-school teaching experience as because of Sara's experience and knowledge about reading instruction at this level. Thus, their collaboration gave her an opportunity to seek advice and obtain feedback about her own instruction (Tania, pp. 15). Significance is similarly ascribed to the other teacher's competence in Theresa's (T3) account of learning a great deal from Stella (T3) because she was "extremely subject-oriented" (p. 30) and in Tina's (T2) characterization of Sofia (T2) as the "L1 lady", referring to her disciplinary knowledge and competence, and abundant creativity.

A perception of the dyad as a collaborative unit is prevalent in these interviews. Sofia (T2) notes that she felt as "kind of a full member of the grade level" (p. 26), since she mainly taught at that level (unlike a rotating teacher moving between grades). In line with this, Stella (T3) contrasts the work within the dyad, and the close collaboration with Theresa (T3) about periodical teaching plans, with her experience from the following year, when she was back to normal and "more on the sideline" (p. 15). Further, it is notable that Theresa uses a discursive "we" about the dyad throughout the interview, giving the impression of a very close collaboration through her emphasis on shared responsibility, ownership, and credit for the instruction and the students' learning outcomes.

Discussion: The importance of shared responsibility

Before embarking on our discussion, we first return to the purpose of this study – examining how shared responsibility in co-taught literacy classes relates to student achievement. Our quantitative analysis revealed that a profound level of shared responsibility between teachers predicted student reading scores postintervention when controlling for reading scores preintervention. This finding extends previous research on co-teaching by empirically linking shared responsibility between co-teachers for planning, enacting, and evaluating instruction to student outcome. A possible interpretation of this finding is that shared responsibility in all phases of the instructional process may contribute to parity (where both teachers engage in meaningful instruction (cf., Pratt, 2014). However, no significant difference in student reading was found between the Delegating group, on the one hand, and those in the Assisting or Co-operative group, on the other.

The quantitative analysis further allowed us to identify six teacher dyads representing extreme cases of particularly low- or high-performing classes and then explore their collaborative teaching experience through in-depth, individual interviews. Through the qualitative analyses, we identified strong differences in shared responsibility between the groups of high- and low-performing classes. Despite the differences, we see an interesting pattern in the six homeroom teachers' responses to the item characterizing the degree of shared responsibility: Three of the homeroom teachers (Bridget (B1), Tina (T2), and Theresa (T3), reported as expected on the survey item (i.e., low-performing classes in the Assisting group and high-performing

classes in the Co-operating group). However, homeroom teachers of both low- (Brenda (B2), Beth (B3) and high-performing (Tania (T1) classes, report that they delegate tasks to the co-teacher. Based on this pattern, we may assume that the response option for Delegating, i.e., having the lead role and delegating distinct tasks, might have attracted teachers of both a more assisting and co-operating nature. This may partly explain why there was no significant difference in student achievement between the Delegating group and the two other groups, an assumption that was explored in the qualitative interviews. We will return to this throughout the discussion.

The mixed methods design thus helped us extend our understanding of how responsibility was shared to different degrees and at more or less profound levels. For example, having the lead may refer simply to having the final say after shared reflection, or to making all key decisions about instruction. Similarly, the tasks themselves, whether delegated or co-determined, vary in value and hence in shared responsibility. Some delegated tasks come with very detailed instructions (and therefore require little pedagogical knowledge) or are not considered to be at the core of instruction. In those cases, sharing is limited. By contrast, other tasks may involve entrusting the co-teacher with responsibility for core aspects of instruction. This difference may be reflected in the different roles that the two teachers ascribe to each other. Hence it is of crucial importance for the degree of shared responsibility.

It is important to acknowledge, particularly when considering how the concept of shared responsibility can be operationalized, that our study's model with two general educators co-teaching, provided the partners with less predefined roles than in other models of co-teaching, e.g., where a special education teacher is paired with a general educator teacher (Friend, 2008), or another specialist is paired with a generalist, such as a speech language pathologist, reading specialist (Krammer et al., 2018) or gifted education teacher (Mofield, 2020) with a classroom teacher. In generalist/specialist dyads, the unique preparation of the specialist can implicitly direct (and therefore constrain) the responsibility of specialists, who often focus on the subset of students with a specific designation. However, while such an arrangement may be efficient, these divisions of labor also constrain co-teachers in achieving authentic shared responsibility because with a mindset characterized by silos of responsibility, both teachers may not feel empowered to plan, enact instruction and evaluate all of the students in the class. Therefore, our results, aligned with Friend's (2021) advocacy for parity, indicate that shared responsibility does not determine "who does what" but is about making instructional decisions together, such as reviewing assessment results together and planning how to respond to all students' needs.

This concept of shared responsibility is also reflected in research on effective instructional teams in schools, and one can reconceptualize co-teaching as a very small team. In fact, numerous parallels can be found between our work and such teams. For example, teams producing innovation in schools (in our study innovation can be defined as co-teachers trying out new instructional practices) is highly influenced with working time, in the form of frequency and regularity of meetings, because these protected spaces allowed for discussions and deeper processes (Somech & Drach-Zahavy, 2007). Both our work and previous research (e.g., Pratt, 2014) also puts high value on frequent conversations between co-teachers. Such

innovation has specifically been associated with four interactive processes: sharing information, learning from each other, motivating and negotiating (Somech & Drach-Zahavy, 2007). We argue that such processes will be more likely to occur if co-teachers view all students as their responsibility and reflect together.

The possibilities and challenges that characterized the six extreme dyads' collaboration, promoting or hindering shared responsibility, thus seem to be well aligned with previous findings in the broader co-teaching and team instruction literature. Hence, when moving to discuss two key features of effective co-teaching in our study, we assume that they are also of high relevance for other co-teaching arrangements. The two key features are role determination and framing of professional differences.

Taking on and ascribing roles

The process of defining and clarifying roles between co-teachers has been identified as critical for achieving effective collaboration (Weiss & Brigham, 2000). However, our results showed that there are multiple levels of roles that must be considered and revisited. As mentioned, the roles of homeroom teacher and co-teacher were to very different extent endowed with predefined tasks and responsibilities – both due to the Norwegian school context in general and the Two teachers project not specifying the roles at any detailed level. Taking into account that the homeroom teacher's role is defined more clearly through both formalities and tradition, it is critical to consider how the co-teacher role is to be formed and defined at a profound level, relative to that of the homeroom teacher. In other words, regardless of titles, it must be considered in what ways and to what extent the co-teacher exerts influence on instruction and planning.

As indicated by Gourvennec (2021), a homeroom teacher holds a gatekeeper function in co-taught classes. This implies that s/he may invite or hinder, to different degrees, the co-teacher's professional engagement. The most striking example here of a co-teacher's perception of restricted access is Arthur's (B1) experience of not being allowed to contact students' parents, not being able to discuss disagreements with Bridget, and not having a say about the handling of externalizing students. However, both Anna (B2) and August (B3) perceive some obstacles relating to the homeroom teacher, i.e., limited opportunities to influence planning (Anna) and the experience of not being listened to when presenting suggestions for instruction or reporting from the reading conferences (August). In stark contrast to this, Sara (T1), Sofia (T2), and Stella (T3) all describe a situation where the homeroom teacher gives them access to professional engagement in all parts of co-teaching.

Another aspect that stands out as important across all six dyads is the co-teacher's response to the role that the homeroom teacher invites him or her to take on. When provoked by what she perceives as Arthur's (B1) know-it-all behavior, Bridget is less willing to give him access to all parts of co-teaching. When Beth (B3) feels that August is insecure and prefers her to take the leading role, she seems to take control over *all* parts of their collaboration. By contrast, the co-teachers in the

high-performing classes are perceived by the homeroom teachers as responding to their invitation by engaging profoundly in all parts of the co-teaching.

Given the imbalance in the predefined expectations of the two roles, the role of a co-teacher often depends on an invitation – by the homeroom teacher – regarding the co-teacher's opportunities for professional engagement. After an invitation is extended, the enactment of this role also depends on how the co-teacher responds to it. Although these more profound divisions of labor may be more difficult to observe, they influence how teachers' professional autonomy and integrity are ensured and so affect motivation and job satisfaction.

Our findings regarding the contribution of roles are supported by previous empirical research. Specifically, difficulties navigating and establishing equal classroom roles when co-teaching, which were described by Scruggs et al. (2007), are the most salient in the low-performing classes. However, these role-related difficulties do not seem to be limited to the classroom activities, but rather appear to be a challenge outside of it as well; the reasons for these difficulties seem manifold. The highest-conflict situation – that of Bridget and Arthur (B1) – strikes us as strongly linked to differing pedagogical beliefs, which supports previous research by Brownell et al. (2006). However, other reasons for difficulties seem to be linked to what Pratt (2014) refers to as external factors – in particular that Anna (B2) would have liked to have her workspace next to that of her homeroom teacher so that she would be invited to an ongoing shared reflection during planning and evaluation.

In terms of theoretical underpinnings, Self Determination Theory (SDT) (Ryan & Deci, 2000) provides a framework for understanding the experience and motivation of an individual teacher nested within a collaborative teaching situation as well as the importance of professional roles. Most notably, co-teachers in low-performing classes often experienced limited *autonomy* in the domains of planning, instruction, or communication with families. Such limits to autonomy may hinder a teacher's ability to experience *competence* in teaching, which often comes from successful handling of challenges. For example, the two male teachers in low-performing classes – both relatively inexperienced compared with the homeroom teacher – were frequently given tasks arguably below their true competence level. Such interaction patterns in the low-performing classes may have helped create weak or even hostile relationships, which would not provide opportunities for *relatedness* for either the co-teacher or the homeroom teacher.

By contrast, the teachers in high-performing classes provide much evidence of strong relationships and trust – both through their examples of collaboration and through their language describing their work (e.g., the use of the discursive "we"). According to SDT, the collaboration may even have enhanced their performance by helping to fulfill the need for relatedness (Ryan & Deci, 2000). Teachers working in their own classrooms typically have great opportunities for autonomy and competence but limited opportunities to relate to other adults. In summary, when teachers' basic psychological needs are not met, for instance because of ill-defined or poorly divided roles, they may perform below their potential; only when those needs are met can teachers realize their full potential.

Professional differences – promoting professional development?

Collaborating teachers' *compatibility* (Pratt, 2014) is closely related to the distribution and perception of roles. The qualitative results allow us to identify two rather different perceptions of what differences in teaching experience, personality, or professional competence bring to co-teaching.

Two of the dyads in low-performing classes are generally negative in their evaluation of the consequences of their differences. Both Bridget (B1) and Beth (B3) strongly emphasize the difference in experience between themselves and their co-teacher, devoting no attention to the hypothetical possibility that Arthur (B1) or August (B3) might bring new perspectives or updated knowledge from their teacher training and early career. Rather, the large difference in experience is presented as the justification for Beth's taking the lead in her dyad. This also seems to be the case for Bridget, but as an element of a more complex picture where other, more important, elements are Arthur's behavior and Bridget's conception of what a homeroom teacher ought to be. Both Arthur and August, for their part, report having picked up some tricks and ideas from observing Bridget and Beth, but still they devote little attention to the significance of the homeroom teachers' experience, which thus seems to be of limited value in their view. In light of Pratt's (2014) theory of achieving symbiosis in co-teaching, the absence or limited presence of the characteristics of the *fulfillment* stage is tangible: strategies for handling challenges smoothly are more or less absent; there is no work on differences to make the relationship stronger; parity, trust, and rapport are damaged or at best partially fulfilled; and shared reflection is sparse.

By contrast, all six teachers in high-performing classes highlight the positive contribution to the collaboration made by their differences in teaching experience, personality, or professional competence. They stress that such differences are enriching, bringing something new to the classroom and to the teacher's own professional development. However, these enriching differences do not seem to relate to more profound pedagogical beliefs (cf. Brownell et al., 2006); rather, the teachers emphasize that they shared core values within their dyad. In light of Pratt's (2014) theory, it seems reasonable to claim that the emphasis placed by these six teachers from high-performing classes on the professionally enriching character of their collaboration, coupled with the fact that they devoted very little attention in their interviews to strategies for handling challenges smoothly (addressed only in response to explicit questions), suggests that these dyads are in a stage of *fulfillment*, where the collaboration is characterized by teachers who are open-minded, use open communication, find (implicitly or explicitly) common ground, value the relationship, and are compatible. Keeping in mind the previously noted challenges relating to external factors that seemed to somewhat hinder the collaboration between the teachers in the low-performing classes, it is interesting to note that such factors are not mentioned as hindering, either explicitly or implicitly, the collaboration in the high-performing classes. Rather, the dyad as such is given significance as a "we" that strengthens the teachers' voices whenever they are standing up for what they believe is for their students' best, when facing colleagues, school management, or politicians. It seems fair to assume that compatibility, equity, and trust within a dyad are prerequisites for

such a position. This finding aligns with, and extends, the conclusion of Krammer and colleagues (2018) that effective teacher pairs exhibited a collective self-efficacy in terms of communicating their needs and securing the appropriate resources.

An overarching and striking difference between the two groups of dyads in the interviews concerns what might be referred to as *valuing the relationship* (cf. Pratt, 2014). Whereas appreciation for the co-teaching relationship permeates the interviews with the teachers in high-performing classes, it is absent or at least weak in the other three dyads. Further, placing a high value on the other's competence is associated with an ongoing open dialogue about all aspects of co-teaching, whereas a low value is associated with limited dialogue – either because of limited opportunities or because of more monologic patterns of collaboration. In this light, it is obvious why the teachers in high-performing classes characterize their collaboration as boosting their professional development while the teachers in low-performing classes do not.

We again use SDT (Ryan & Deci, 2000) to consider this theme from the viewpoint of the individual teacher. Here, however, rather than examining only the need to achieve growth, we must also consider a key assumption of SDT: a person's need for growth drives their behavior. Hence gaining mastery over new challenges is essential for achieving a robust sense of self. An individual in a collaborative-teaching situation could therefore view the situation as an opportunity to enrich their professional knowledge through observation, discussion, and trying new approaches. Such an individual would probably be highly engaged in the collaborative and learning process, as is evidenced in the analysis of dyads in high-performing classes described above. However, if basic needs are unfulfilled or an orientation toward self-growth is not already in place, teachers may not be motivated to exploit the learning opportunity provided by collaborative teaching. It must also be kept in mind that both teachers here have similar formal qualifications (as classroom teachers) rather than represent the traditional pairing of a special educator with a classroom teacher, meaning that the unique strengths that each person brings to the situation may be less evident than where teachers have different professional backgrounds. Hence more curiosity and discussion may be required for the pairs to recognize the potential contributions of each person.

Conclusion and implications

The main contribution of this study is the establishment of an association between degree of shared responsibility and student achievements' as well as the use of the explanatory mixed methods design, which enabled qualitative insight into the reasons for the association between shared responsibility and student outcomes. The explanatory design also allowed for pragmatic and situated knowledge that can guide teacher implications and professional development. Taken together, our results support previous research regarding challenges met by co-teaching teams and characteristics of well-functioning teams.

In summary, if the two collaborating teachers truly engage in co-teaching, give each other meaningful roles, and embrace the potential to learn from each other's competence, then "one plus one teacher" may equal more than the competence of

"two teachers" and result in a situation with much growth and synergy. By contrast, when that is not the case, one teacher may withdraw while the other feels burdened, and the metaphorical sum may be less than "two teachers".

Based on our study, we would recommend a mindset shift stating that both professionals, in any co-teaching arrangement, bring strengths and knowledge that would benefit all children. In fact, in leadership research, team heterogeneity has long been associated with innovation (Jackson, 1996) and creativity (Shih & Zhou, 2007) but only when all members are engaged in the process. Second, we would recommend intentionality and planning in order to create opportunities for cycles of reflection, goal setting and planning. To guide such processes, we have created questionnaires that can guide co-teaching dyads throughout their year of co-teaching (McTigue et al., 2022).

Limitations

Although this study relies upon a large, quantitative data set combined with in-depth individual interviews, it is important to bear in mind its methodological limitations. First, the degree of shared responsibility is self-reported by the teachers. Hence the study does not identify objective observations of parity. Also, the self-report only addresses the degree of shared responsibility on a general level, and consequently we have no detailed notions of *how* the two teachers interacted in the large sample. Second, only homeroom teachers responded to the survey item characterizing the degree of shared responsibility. Given that one of the preformulated response options (delegating) seems not to distinguish between delegation that encourages the co-teacher's autonomy and detailed delegation that deprives the co-teacher of such autonomy, it would have been informative to have co-teachers' responses to the same item. This limitation is observed and to a large extent met through the complexities of responsibility-sharing's nature revealed in the interviews. Third, the purposeful sampling of 6 extreme dyads among 148 may cast doubt on the relevance of the qualitative findings for less extreme dyads. However, since aspects previously found to be important among co-teaching teams characterized as effective (Pratt, 2014) are also prominent in the dyads chosen based on their students' strong progress, it seems reasonable to consider the results relevant for all dyads. Finally, it should be noted that two of the three dyads in low-performing classes were co-taught only during the second year of the intervention, whereas the other four dyads collaborated both years. Hence the characteristics of co-teaching during first grade in two classes are not taken into account in the qualitative part of this study.

Appendix

Interview guide

[Warm-up: make sure you obtain information about whether the two teachers worked together in both years and whether they are still working together, and also possibly try to find out something about why they chose to become teachers.]

Introduction

It's been six months since the end of the Two teachers intervention – since you were two teachers sharing responsibility for reading and writing instruction within the Two teachers context. The reason why we'd now like to interview you, [co-teacher's name] and other dyads of teachers who worked together in Two teachers is that we know that a lot of things influence how co-teaching works, how it's experienced and to what extent it leads to enhanced learning outcomes for students. In these interviews I meet teacher pairs whose classes benefited differently from the presence of two teachers in Norwegian classes, enabling us to learn more about the complexity of working together. I'd like to hear about your experiences with teacher collaboration and with Two teachers, but first of all I'd like to hear a little bit about you as a teacher.

Theme 1: Teachers' figured worlds with regard to good instruction and to the roles of students and teachers

1. Do you remember a Norwegian lesson you were particularly pleased with? Would you like to tell me about it?[6]
    a. What was it about that lesson that made it good?
    b. Are the characteristics of that lesson in line with your idea about what good instruction is? What characterizes your ideal of good instruction?
    c. Do you remember anything similar from beginner instruction/year 1 or 2?

2. How would you describe the roles of the students and the teacher in your instruction?
    a. What do you think is your role in the classroom?
    b. What kind of place do you expect and want students to have and take in the classroom?
    c. What aspect of the division of roles in your class are you the most pleased with? And what do you think is the most challenging aspect?

3. How would you describe your work on reading and writing instruction in your teaching?
    a. [Link this to all the teachers have said about his or her more fundamental views.]

[6] The interviewer always asked all *numbered* questions (1–5) within each of the three themes – to the extent that they were applicable to the interviewee's project condition and that the interviewee had not already informed the topic in previous responses. By contrast, the items marked with *characters* (a–h) or *Roman numerals* (i–iii) were intended as guidance for possible follow-up questions and so were asked only as appropriate.

4. [For conditions 1 and 2]: During the two years of the intervention, you worked with *Language Tracks*.[7] Would you like to tell me a little about that work? Is there anything about working with *Language Tracks* that you remember particularly well? Please tell me about it.

    a. The work with *Language Tracks* was supposed to take the school's own experienced needs as its starting point. Do you remember what needs were identified and how you followed up on them? (For example, did you come back to this later on to evaluate the work so far, was it mentioned as having guided the choice between work packages 2 and 3, …?)
    b. Did your work with *Language Tracks* change the way you think about teaching and instruction in any way? Please tell me about it.
    c. In what ways did you and your co-teacher, as a team, process and adapt the work you did with *Language Tracks*? Did it make experimentation in instruction easier?

5. [For condition 2]: For those of you who were in group 3 in Two teachers, your schools committed to following specific guidelines (that is, to carry out reading conferences, guided reading at least once a week, sessions where students read aloud to teachers at least once a week and extra support for students struggling with letter knowledge and/or lagging behind in reading). Did these mandatory working methods – and your duty to use them – change your instruction in any way? Please tell me about it.

    a. Did you perceive this as useful/instructive/liberating, for example because it required adaptations on the part of school management – or did you experience it as a straitjacket depriving teachers of instructional ownership…?

Theme 2: Co-teaching

1. Now I'd like to hear a little about your experience of co-teaching with [co-teacher's name] in first and second grade. [Pause.] How did you experience teaching together with [him/her]? What's the best part of your experience being two teachers in the class? Please tell me about it.

    a. Could you give me an example of something you were very pleased with?
    b. Could you give me an example of a time when collaboration did not work as you expected/wanted.
    c. How did you divide work and roles during planning and during instruction? Could you describe your role and [co-teacher's name]'s role?

[7] *Language Track* is the name of a large open-source TPD from where the schools and teachers chose relevant TPD modules.

i. How did you distribute roles during instructional planning?
ii. How did you distribute roles during instruction: what organizational approach was used in the class and how did you perceive the distribution of roles? Did the distribution of roles change over time?
iii. How much time/how many lessons did the co-teacher spend in the class in addition to the lessons funded by Two teachers? Did the homeroom teacher spend a lot more time with the students because she taught almost all subjects while the co-teacher was present in the class only for the dedicated lessons?

d. Did co-teaching make it possible to meet individual students' needs better? What kinds of needs?
e. If you disagreed about priorities or about how to do something, how did you manage such disagreement (of a subject-specific, didactical or pedagogical nature)?
f. How would you describe the "chemistry" between the two of you? How important is good chemistry when co-teaching?
g. How did you perceive the role of school management in the project? Did they follow the project closely? Did they appreciate the project? Did they encourage experimentation?
h. Did you change your views on co-teaching based on your experience from Two teachers? Did you develop a more positive/negative attitude? Why?

Theme 3: teachers' understandings of students' literacy development and the roles played by various factors for that development

1. You've now followed a class for [two or three] years. How would you characterize that class compared with your previous classes? (The students' (literacy) development as a class/group, any gaps within the group.)
2. [Show the two graphs representing the classes' development in (i) decoding and (ii) reading comprehension.] Here you can see how your class has developed in word reading/decoding and reading comprehension from the start of first grade to the end of second grade, compared with other classes who had two teachers in Two teachers. [Explain how to read the graph.]

a. Looking at the development of your class compared with other classes in Two teachers, we can see that their reading development is a lot [stronger/weaker] than average.

i. Does this surprise you?
ii. What do you think may explain this result? [Possible follow-up questions: school management, school culture, collaboration between school and parents, the students' backgrounds and circumstances (in a broad sense), co-teaching, methods of instruction, relationship between the students and the teacher…]

Acknowledgments We thank students and teachers who participated in the study and especially the twelve teachers who took the time to share their experiences and reflections about the co-teaching experience with us.

Author contibutions The Research Council of Norway had no role in designing the study; collecting, analysing or interpretating the data; writing the report; or deciding to submit the manuscript for publication.

Funding This work was supported by the Research Council of Norway (RCN), grant number 256197. Open access funding provided by University Of Stavanger.

References

Alexander, R. (1997). *Policy and practice in primary education: Local initiative, national agenda*. Routledge.

Andersen, S. C., Beuchert, L., Nielsen, H. S., & Thomsen, M. K. (2020). The effect of teacher's aides in the classroom: Evidence from a randomized trial. *Journal of the European Economic Association, 18*(1), 469–505. https://doi.org/10.1093/jeea/jvy048

Atkinson, J. M., & Heritage, J. (1999). Transcript notation – Structures of social action: Studies in conversation analysis. *Aphasiology, 13*(4–5), 243–249. https://doi.org/10.1080/026870399402073

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Brownell, M. T., Adams, A., Sindelar, P., Waldron, N., & Vanhover, S. (2006). Learning from collaboration: The role of teacher qualities. *Exceptional Children, 72*(2), 169–185. https://doi.org/10.1177/001440290607200203

Conderman, G. (2011). Middle school co-teaching: Effective practices and student reflections. *Middle School Journal, 42*(4), 24–31. https://doi.org/10.1080/00940771.2011.11461771

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/psychologie Canadienne, 49*(3), 182–185. https://doi.org/10.1037/a0012801

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry, 12*(2), 219–245. https://doi.org/10.1177/1077800405284363

Friend, M. (2008). Co-teaching: A simple solution that isn't simple after all. *Journal of Curriculum and Instruction, 2*(2), 9–19. https://doi.org/10.3776/joci.2008.v2n2p9-19

Friend, M. (2021). *Interactions collaboration skills for school professionals* (9th ed.). Pearson.

Friend, M., Embury, D. C., & Clarke, L. (2015). Co-Teaching versus apprentice teaching: an analysis of similarities and differences. *Teacher Education and Special Education, 38*(2), 79–87. https://doi.org/10.1177/0888406414529308

Friend, M., Reising, M., & Cook, L. (1993). Co-teaching: An overview of the past, a glimpse at the present, and considerations for the future. *Preventing School Failure: Alternative Education for Children and Youth, 37*(4), 6–10. https://doi.org/10.1080/1045988X.1993.9944611

Gourvennec, A. F. (2021). Digging into the extremes: A case study of figured worlds of early literacy instruction among homeroom teachers in more or less successful co-taught classrooms.

L1-Educational Studies in Language and Literature, 21, 1–30. https://doi.org/10.17239/L1ESLL-2021.21.01.19

Haaland, V. F., Rege, M., & Solheim, O. J. (2022). Do students learn more with an additional teacher in the classroom? Evidence from a field experiment. [Manuscript submitted for publication]. The Norwegian Reading Centre, University of Stavanger.

Jackson, S. E. (1996). The consequences of diversity in multidisciplinary work teams. In M. A. West (Ed.), *Handbook of work psychology* (pp. 53–75). Wiley.

Johnston, W. R., & Tsai, T. (2018). *The prevalence of collaboration among American teachers:* National findings from the American Teaching Panel. (Research Report) Santa Monica, CA: RAND Corporation. Retrieved from https://www.rand.org/content/dam/rand/pubs/research_reports/RR2200/RR227/RAND_RR2217.pdf

Jones, N. J., & Winters, M. A. (2022). Are Two Teachers Better Than One?: The Effect of Co-Teaching on Students With and Without Disabilities. *The Hournal of Human Resources*. https://doi.org/10.3368/jhr.0420-10834R3

Karten, T. J., & Murawski, W. W. (2020). *Co-Teaching Do's, Don'ts, and Do Betters*. ASCD.

Krammer, M., Gastager, A., Lisa, P., Gasteiger-Klicpera, B., & Rossmann, P. (2018). Collective self-efficacy expectations in Co-teaching teams — What are the influencing factors? *Educational Studies, 44*(1), 99–114. https://doi.org/10.1080/03055698.2017.1347489

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Testing linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Main, K. (2012). Effective middle school teacher teams: A ternary model of interdependency rather than a catch phrase. *Teachers and Teaching, 18*(1), 75–88. https://doi.org/10.1080/13540602.2011.622556

Maxwell, J. A. (2009). Designing a qualitative study. In L. Bickman & D. J. Rog (Eds.), *The SAGE handbook of applied social research methods* (2nd ed., pp. 214–253). SAGE.

McTigue, E. M., Gourvennec, A. F., & Solheim, O. J. (2022). Key question for literacy co-teachers: What possibilities do we have together? *The Reading Teacher, 76*(2), 176–190. https://doi.org/10.1002/trtr.2121

Ministry of Education and Research (2020). *The education system.* Retrieved 03.03.2022 from: https://www.regjeringen.no/en/topics/education/grunnopplaring/the-norwegian-education-system/id445118/

Mofield, E. L. (2020). Benefits and barriers to collaboration and co- teaching: Examining perspectives of gifted education teachers and general education teachers. *Gifted Child Today, 43*(1), 20–33. https://doi.org/10.1177/2F1076217519880588

Murawski, W. W., & Swanson, H. L. (2001). A Meta-analysis of co-teaching research: Where are the data? *Remedial and Special Education, 22*(5), 258–267. https://doi.org/10.1177/074193250102200501

Neale, M. D. (1997). *Neale analysis of reading ability–revised (Windsor, NFER-Nelson).*

Pesonen, H. V., Rytivaara, A., Palmu, I., & Wallin, A. (2020). Teachers' Stories on Sense of Belonging in Co-Teaching Relationship. *Scandinavian Journal of Educational Research*. https://doi.org/10.1080/00313831.2019.1705902

Pratt, S. (2014). Achieving symbiosis: Working through challenges found in co-teaching to achieve effective co-teaching relationships. *Teaching and Teacher Education, 41*, 1–12. https://doi.org/10.1016/j.tate.2014.02.006

R Core Team (2020). *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria: Retrieved from https://www.r-project.org/

Ryan, R. M., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Rytivaara, A., & Kershner, R. (2012). Co-teaching as a context for teachers professional learning and joint knowledge construction. *Teaching and Teacher Education, 28*, 999–1008. https://doi.org/10.1016/j.tate.2012.05.006

Rytivaara, A., Pulkkinen, J., & Palmu, I. (2021). Learning about students in co-teaching teams. *International Journal of Inclusive Education*. https://doi.org/10.1080/13603116.2021.1878299

Scruggs, T. E., Mastropieri, M. A., & McDuffie, K. A. (2007). Co-teaching in inclusive classrooms: A metasynthesis of qualitative research. *Exceptional Children, 73*(4), 392–416. https://doi.org/10.1177/2F001440290707300401

Shin, S. J., & Zhou, J. (2007). When is educational specialization heterogeneity related to creativity in research and development teams? Transformational leadership as a moderator. *Journal of Applied Psychology, 92*(6), 1709. https://doi.org/10.1037/0021-9010.92.6.1709

Solheim, O. J., Frijters, J. C., Lundetræ, K., & Uppstad, P. H. (2018). Effectiveness of an early reading intervention in a semi-transparent orthography: A group randomised controlled trial. *Learning and Instruction, 58*, 65–79. https://doi.org/10.1016/j.learninstruc.2018.05.004

Solheim, O. J., & Opheim, V. (2018). Beyond class size reduction: Towards more flexible ways of implementing a reduced pupil-teacher ratio. *International Journal of Educational Research*. https://doi.org/10.1016/j.ijer.2018.10.008

Solheim, O. J., Rege, M., & McTigue, E. (2017). Study protocol: "Two Teachers": A randomized controlled trial investigating individual and complementary effects of teacher-student ratio in literacy instruction and professional development for teachers. *International Journal of Educational Research, 86*, 122–130. https://doi.org/10.1016/j.ijer.2017.09.002

Somech, A., & Drach-Zahavy, A. (2007). Schools as team-based organizations: A structure-process-outcomes approach. *Group Dynamics: Theory, Research, and Practice, 11*, 305. https://doi.org/10.1037/1089-2699.11.4.305

Størksen, I., Ellingsen, I. T., Tvedt, M. S., & Idsøe, E. M. (2013). Norsk vokabulartest (NVT) for barn i overgangen mellom barnehage og skole: Psykometrisk vurdering av en nettbrettbasert test [Norwegian Vocabulary Test (NVT) for children in transition between day care center and school. Reviewing psychometric properties of a computer tablet based test]. *Spesialpedagogikks Forskningsdel, 4*(13), 40–54.

ten Braak, D., & Størksen, I. (2021). Psychometric properties and validity of the ani banani math test. *European Journal of Developmental Psychology, 18*(4), 610–628. https://doi.org/10.1080/17405629.2021.1879046

Torgersen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *TOWRE: Test of word reading efficiency*. Psychological Corporation.

Wechsler, D. (1991). *WISC-III: Wechsler intelligence scale for children: Manual*. Psychological Corporation.

Weiss, M. P. (2004). Co-teaching as science in the schoolhouse: More questions than answers. *Journal of Learning Disabilities, 37*(3), 218–223. https://doi.org/10.1177/2F00222194040370030601

Weiss, M. P., & Brigham, F. J. (2000). Co-teaching and the model of shared responsibility: What does the research support? *Advances in Learning and Behavioral Disabilities, 14*, 217–246.

Zhang, X., Korponen, T., Räsänen, P., Aunola, K., Lerkkanen, M.-K., & Nurmi, J.-E. (2014). Linguistic and spatial skills predict early arithmetic development via counting sequence knowledge. *Child Development, 85*(3), 1091–1107. https://doi.org/10.1111/cdev.12173

Authors and Affiliations

Aslaug Fodstad Gourvennec[1] · Oddny Judith Solheim[1] · Njål Foldnes[1] · Per Henning Uppstad[1] · Erin M. McTigue[1]

Oddny Judith Solheim
oddny.j.solheim@uis.no

Njål Foldnes
njal.foldnes@uis.no

Per Henning Uppstad
per.h.uppstad@uis.no

Erin M. McTigue
erin.m.mctigue@uis.no; emmctigue@gmail.com

[1] Norwegian Centre for Reading Education and Research, University of Stavanger, Postboks 8600 Forus, 4036 Stavanger, Norway