

Learning Loss During the COVID-19 Pandemic and the Impact of Emergency Remote Instruction on First Grade Students' Writing: A Natural Experiment

Gustaf Bernhard Uno Skar¹, Steve Graham^{2,3}, and Alan Huebner⁴

¹ Department of Teacher Education, Norwegian University of Science and Technology

² Mary Lou Fulton Teachers College, Arizona State University

³ Institute for Learning Sciences and Teacher Education, Australian Catholic University, Brisbane Campus

⁴ Department of Applied and Computational Mathematics and Statistics, University of Notre Dame

The coronavirus disease 2019 (COVID-19) pandemic and the sudden cancellation of in-class instruction for many students around the world presented an unprecedented disruption in children's education. As the COVID-19 pandemic took form, multiple concerns were raised about the potential negative impact on students' learning. The current study examined this proposition for children's writing. We compared the quality of writing, handwriting fluency, and attitude toward writing of first grade Norwegian students during the COVID-19 pandemic (421 girls, 396 boys), which included emergency remote instruction for almost 7 weeks, with first grade students in the same schools a year before the pandemic began (835 girls, 801 boys). After controlling for variance due to national test scores, school size, proportion of certified teachers, students per special education teacher, school hours per student, student gender, and native language, we found that students attending first grade during the pandemic had lower scores for writing quality, handwriting fluency, and attitude toward writing than their first grade peers tested a year earlier before the COVID-19 pandemic emerged. Implications for policy and instruction as well as future research are presented.

Educational Impact and Implications Statement

In December 2019, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus was identified, and it spread across the world quickly causing the coronavirus disease 2019 (COVID-19) pandemic, presenting unprecedented challenges for the education of school-age children. This study examined the impact of the pandemic and the temporary 7 week replacement of in class learning with online instruction on the writing of first grade children in Norway. First grade students tested shortly after the end of online instruction displayed a learning loss, having lower scores on measures of writing quality, handwriting fluency, and attitude toward writing than first grade children from the same schools tested a year earlier before the start of the pandemic.

Keywords: writing, COVID-19, first grade, handwriting, motivation


In December 2019, a new virus, SARS-CoV-2, was identified. It spread across the world quickly, presenting unprecedented economic and social challenges, impacting virtually all aspects of daily life including public health, the world of work, and food systems.¹ To help slow the spread and impact of the virus, countries across the globe in March and April of 2020 moved from in class instruction to emergency remote instruction, which

included remotely delivered instruction (e.g., online, radio) or some combination of remote and in class instruction (Di Pietro et al., 2020; Hodges et al., 2020).

Norway, the country in which this study took place, was among the nations that canceled in class instruction and applied a model of emergency remote instruction that was delivered online in March of 2020. In a public statement on March 12, the Norwegian Directorate of Health announced that it had decided to close all schools immediately. The closing of schools, which was replaced by emergency remote instruction via digital meeting platforms, was to last until March 26, 2020.² Subsequently, the Norwegian government extended the period of emergency remote instruction to April 27. The

Gustaf Bernhard Uno Skar  <https://orcid.org/0000-0002-6486-396X>

Steve Graham  <https://orcid.org/0000-0002-6702-5865>

Alan Huebner  <https://orcid.org/0000-0003-2141-5793>

This study was made possible by funding from the Norwegian Research Council (Grant 288795).

Correspondence concerning this article should be addressed to Gustaf Bernhard Uno Skar, Department of Teacher Education, Norwegian University of Science and Technology, Gunnerus Gate 1, NO-7491 Trondheim, Norway. Email: gustaf.b.skar@ntnu.no

¹ Retrieved from <https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems>.

² Retrieved from <https://www.helsedirektoratet.no/nyheter/helsedirektoratet-har-vedtatt-omfattende-tiltak-for-a-hindre-spredning-av-covid-19>.

lockdown lasted for 46 days, weekends and Easter holidays included.³

This sudden cancellation of in-person instruction in Norway and across the globe presented an extraordinary educational challenge. It was estimated that 1.19 billion of students across the world were impacted (UNESCO, 2020). The world had never faced an educational disruption of this magnitude (Azoulay, 2020; Winthrop, 2020).

The application of emergency remote instruction, including the use of online platforms to promote such instruction, provided a mechanism for ensuring the continuity of schooling around the world (Di Pietro et al., 2020). Even so, many experts expressed concerns about the effects of this sudden transition on students' learning (e.g., Daniel, 2020). While data on the actual learning loss experienced by students is sorely needed (the focal point of this investigation), predictions about the level of learning loss and its subsequent consequences proliferated. Citing forthcoming research by Hanushek and Wosseman, Schleicher (2020) noted that a loss of one third of a year of schooling due to the COVID-19 pandemic would represent a 1.5% loss of current gross domestic product (GDP) for a typical country for the rest of this century. For the United States, this would translate to a total loss of \$15,300,000,000,000 by the year 2100. Consequently, it is critical to acquire information on students' learning loss as a result of the COVID-19 pandemic and the resulting emergency remote instruction. There is little reason to expect all aspects of student learning were impacted equally, as learning effects differ from one academic domain to the next (Alexander, 2003). As a result, this study focused on one aspect of students' learning: writing.

Why is it important to examine the association between learning to write and the COVID-19 pandemic and the ensuing emergency remote instruction? Writing is a foundational skill in terms of students' success at school and beyond. Writing is one of the primary means by which students' demonstrate what they know (e.g., written essays, written responses on tests). Writing about material read enhances comprehension of the ideas presented in text, and writing instruction makes students better readers (Graham & Hebert, 2011). Writing also improves learning, as writing about content material makes material more understandable (Graham et al., 2018). Just as important, writing is a versatile tool as we use it to record information, communicate, persuade, inform, chronicle experiences, think about topics, create imaginary worlds, entertain others, explore the meaning of events, and heal psychological and physiological wounds (Graham, 2006).

Why the COVID-19 Pandemic and Emergency Remote Instruction Were Likely to Reduce Student Learning

Stress

There are multiple reasons why the COVID-19 pandemic and emergency remote instruction might negatively impact learning, including learning to write. One possible determinant to students' learning involves the effects of the COVID-19 pandemic and ensuing emergency remote instruction on teachers and parents. Teachers and parents have reported increased stress as a result of schools moving to distance learning (Bonal & Gonzales, 2020; Kimm & Ashbury, 2020). As teacher stress increases, there is a corresponding decrease in student learning (e.g., Oberle & Schonert-Reichl, 2016). Additionally, parents have become proxy educators during emergency remote instruction—a responsibility they cannot easily abdicate. Davis et al. (2021) reported that stress is especially

elevated for parents of children who find distance learning challenging, as occurred in many countries such as Norway during emergency remote instruction. This increased stress is likely to impact their children's learning.

The potentially negative effects of stress on learning are not limited to teachers and parents, but extend to children too. A study by Sprang and Silman (2013) found that children who were isolated at home during previous pandemics were more likely to experience stress disorder, grief, and adjustment problems. Howard and Howard (2012) also reported children experienced increased anxiety and fear during a prior pandemic. Higher levels of stress, anxiety, and fear take a toll on students' well-being and learning, and can result in poorer learning outcomes (Di Pietro et al., 2020).

Resources for Distance Learning

Emergency remote learning that accompanied the COVID-19 pandemic required that teachers deliver instruction using distance learning methods. In some instances, this involved older technology such as radios (Kirshner, 2020), and in other instances it involved newer digital learning platforms (Fauzi & Khusuma, 2020). While digital tools can enable and support learning and writing in multiple ways (Graham, in press a), the COVID-19 pandemic and ensuing in-person school closures occurred at a time when many countries, schools, and teachers were not ready to apply these tools effectively. For example, just before the pandemic, one-fourth of principals in countries participating in the Organization for Economic and Cooperation Development (OECD) assessments indicated learning was hindered due to a shortage or inadequacy in digital tools (Schleicher, 2020). Many teachers in this OECD survey also indicated the need for additional training in the use of digital tools, and 47% of teachers did not allow students to use these tools for classwork. For these teachers, the quick transition to emergency remote instruction using digital tools that occurred as a result of the COVID-19 pandemic raises questions about their effectiveness in delivering instruction digitally and remotely.

The use of digital platforms for student learning at home during the COVID-19 pandemic not only depends on teachers' skills and school resources; they rely on parents' availability and abilities to serve as proxy educators as well as the availability of digital resources at home. Such resources are not equally available to all families. Students from less affluent homes are less likely to have access to a laptop or computer, family members are less likely to be able to telework from home, and their parents may be not be as educated (Di Pietro et al., 2020). These factors likely diminish the effectiveness of digital learning for some students.

Quality of Learning Opportunities

One potentially negative consequence of moving education to a remote learning environment as was done with emergency remote instruction during the COVID-19 pandemic is that less time may be devoted to learning (Schleicher, 2020). For instance, Huber et al. (2020) reported that weekly learning time was 4 to 8 hr less during emergency remote instruction during the COVID-19 pandemic in Austria, Switzerland, and Germany than when students attended

³ Retrieved from <https://www.udir.no/contentassets/35a1b6c984ec4114b00479cc943322d1/ekspertgruppens-rapport-konsekvenser-av-snitteverntak-i-barnehager-og-skoler-12.11.20.pdf>.

schools in person. Theoretically, learning is a consequence of time spent and time needed to learn (Carroll, 1989). Lowering instructional time for writing (Graham & Santangelo, 2014) or time spent writing (Graham et al., 2015) is likely to have a negative impact on writing.

In addition to instructional time, other factors can potentially decrease students' learning during emergency remote instruction. For example, while students can interact with others when learning digitally, this may not be as effective as the contact students experience during in class instruction (Di Pietro et al., 2020). Likewise, students may be less motivated to learn due to decreased instructional time and less direct person-to-person contact with their teachers. Finally, some complex skills like writing may be easier to learn when teachers have direct access to students.

Purpose of the Present Study

This study examined the effect of the COVID-19 pandemic and emergency remote instruction on the writing of first grade students in Norway. Because it is not ethically responsible to assign students to in-class and remote learning during a pandemic, the impact of such a situation must be studied with more natural methods. Consequently, we assessed the quality of first grade students' writing, handwriting fluency, and attitude toward writing in May/June of 2019 before the first case of the COVID-19 pandemic was reported (referred to as the before COVID-19 cohort). In May/June of 2020, we tested a second cohort of first grade students from the same schools using the same writing measures (referred to as the during COVID-19 cohort). This was after the start of the pandemic and just following the end of emergency remote instruction in Norway (students returned to in class instruction on April 28, 2020). This allowed us to assess the writing of comparable cohorts of first grade students who did and did not experience these the COVID-19 pandemic conditions. This is the first study to our knowledge that provides an estimate of the effects of the COVID-19 pandemic and emergency remote instruction on children's writing.

We focused our attention on first grade as there is a general consensus that problems that occur early in a child's education, if not corrected quickly, can become more intractable over time (Slavin et al., 1989). While a previous survey study by Håland et al. (2019) indicated that close to one in five teachers do not teach writing in the first semester of first grade, a larger study by Graham et al. (2021) just before the COVID-19 pandemic began found that virtually all first grade teachers taught writing. Teachers in this second study indicated they used a variety of instructional procedures to teach writing, and their students spent about 20 min a day writing. Thus, first grade provided an excellent point to study the effects of the COVID-19 pandemic and emergency remote instruction. This is the initial point where writing is taught to virtually all students in Norway, and negative effects on writing due to the COVID-19 pandemic and emergency remote instruction with first graders are likely to reverberate across the school years.

The theoretical framework that guided the current study was the Writer(s)-within-Community model (WWC; Graham, 2018a, 2018b, in press b). According to this model, writing and writing instruction are simultaneously and interactively shaped and bound by the writing communities in which they occur as well as the cognitive capabilities and resources of the members of these communities (teachers and students). In terms of students' capabilities and resources, the model posits that students' motivations for writing, including attitude

toward writing, fuel effort and provide the impetus for students to draw upon available cognitive control mechanisms (executive functioning processes) to regulate and process information for writing from long-term memory and external sources as well as execute the processes needed to produce writing (conceptualization, ideation, translation, transcription, and reconceptualization).

For writers in general, and beginning writers particularly, the cognitive resources specified in the WWC model are limited (Graham, 2018a, 2018b). This is especially problematic for first graders because their use of writing production processes requires conscious attention and considerable cognitive resources. As a result, young children commonly minimize conceptualization (e.g., planning) and reconceptualization process (e.g., revising) as they write, applying an approach that involves generating one idea that serves as the stimulus for the next idea, and so forth. They apply this approach, at least in part, because transcription skills (e.g., handwriting) are so slow and effort intensive that engaging in other resource intensive activities such as conceptualization and reconceptualization can result in cognitive overload.

In the current study, we included a measure of attitude toward writing as well as handwriting fluency. As noted above, the former can impact young writers' effort and persistence (Graham, 2006), whereas the latter impacts written production (Berninger, 1999). We also assessed the quality of students' writing on a functional and authentic writing task.

The WWC model (Graham, in press b), along with practical constraints, influenced our decision to focus on a single country in this study. The model posits that writing and writing instruction are shaped and bound by local context (writing purposes, available writing tools, typical writing practices, types of written products, experiences and resources of classroom members, physical and social environment, and collective history of the classroom) as well as macrofactors including cultural, social, institutional, political, and historical influences. While the model supports the proposition that no single writing classroom or its effects are exactly identical within a country, it also stands to reason that classrooms that share a similar set of cultural, social, institutional, political, and historical catalysts are more likely to have similar effects on writing than ones that operate under a different configuration of mitigating influences. We made the decision to examine the effects of the COVID-19 pandemic and emergency remote instruction in Norway because it is a country with a strong educational system (scores above average on OECD assessments) and digital infrastructure (98% of homes have Internet access), where the use of digital technology for teaching has been emphasized for decades (Blikstad-Balas et al., in press). It is also an affluent country, identified as the fourth richest nation in the world (<https://data.oecd.org/gdp/gross-domestic-product-gdp.htm>). This provided a strong test of the hypothesis of the negative effects of the COVID-19 pandemic and emergency remote instruction, as Norway had the infrastructure and resources to respond strongly to this crises.

Research Question and Prediction

Our study answered the following question: Did the COVID-19 pandemic and emergency remote instruction negatively impact the quality of first grade students' writing, handwriting fluency, and attitude toward writing?

We predicted that the cohort of students who attended first grade during the COVID-19 pandemic and emergency remote instruction would evidence lower scores for writing quality, handwriting fluency, and attitude toward writing than first grade cohort of students attending the same school a year earlier before the COVID-19 pandemic. Because no two sets of students are exactly the same from one year to the next, our analysis controlled for variance due to school performance on national tests (these are provided to older students in Norwegian schools), school size, proportion of certified teachers, number of students per special education teacher, average school hours spent teaching students, student gender, and primary languages of students (native Norwegian speaker, bilingual speaker with Norwegian as a native language, and native speaker of a language other than Norwegian). In the context of this study, bilingual refers to students who learn two or more languages as part of their normal development. For example, bilingual student would be one who learns Norwegian and Swedish as they are growing up, as both languages are spoken naturally at home.

School variables like the ones included in this study (Walberg & Ethington, 1991) as well as student gender and language predict students' writing performance (Camping et al., 2020; Cordeiro et al., 2018). In terms of language, we expected that students who were native Norwegian speakers and bilingual speakers with Norwegian as a native language would likely be better writers than native speakers of a language other than Norwegian because students were asked to write in Norwegian (see Camping et al., 2020, for evidence to support this contention). Additionally, we expected that girls' scores on all three writing measures would be higher than those of boys. Gender is associated with individual differences in writing quality, handwriting fluency, and attitude toward writing, as girls produce better text, have faster handwriting, and express more positive attitudes about writing (e.g., Ekholm et al., 2018; Graham et al., 1998; Reilly et al., 2019).

Our prediction that the COVID-19 pandemic and emergency remote instruction would depress students' writing scores was based on three arguments. One, as noted earlier, quality of learning, resources for distance learning, and stress created by the COVID-19 pandemic and the resulting emergency remote instruction provided are likely to dampen students' learning.

Two, Blikstad-Balas et al. (in press) investigated how a large sample of parents ($N = 4,642$) viewed the success of emergency remote instruction in Norway. While this study included parents of children in Grades 1 to 10, we only share findings for Grades 1 to 4 here because they are most relevant to the current investigation.

Twenty-three percent of parents in Blikstad-Balas et al. (in press) reported that their children were not expected by teachers or schools to attend classes online at all. Another 27% of parent indicated their children demonstrated attendance by completing assigned tasks. According to parents, 7%, 26%, and 21% of children had no teacher contact, contact once a week, or contact two or three times a week during emergency remote instruction. Many parents (51%) strongly agreed they had to continuously monitor their child's progress to ensure schoolwork was completed, and 76% of them indicated they had to spend more time than usual following-up on their child's schoolwork (85% of parents reported spending 1 to 2 hr more on their child's schoolwork a day). Even more important, 23% and 35% of parents indicated their child spent 50% or less than 50% of time, respectively, doing schoolwork during emergency remote instruction. Therefore,

many first graders had little contact with their teachers during emergency remote instruction, and the time devoted to learning was severely restricted for many of them. This makes it likely that students' writing development during the pandemic and resulting emergency remote instruction would be depressed. This is especially important for the writing measures included in our study, as handwriting fluency develops mainly through practice writing (Graham, 2009), quality of writing is impacted by how much students' write and the instruction they receive (Graham et al., 2012), and attitude toward writing is influenced by interactions with teachers and peers (Graham et al., 2015).

Finally, most studies examining possible decrements in writing when school is not in session for summer break have reported loss in writing skills (Jesson et al., 2009; Pacchter et al., 2015). This occurred even when the break was only 6 weeks long (Meyer et al., 2020). While the conditions between a summer break and the COVID-19 pandemic and emergency remote instruction are not identical, these findings provided support for our prediction.

To better situate the possible learning loss in writing experienced by grade one students in the during COVID-19 cohort, we draw on the findings from an earlier study where we collected data on the same three writing measures (writing quality, handwriting fluency, and attitude toward writing) with second-grade students in the same schools (Graham et al., 2021). This data was collected at the exact same time as data for the before COVID-19 sample in this investigation (May/June 2019). These data provided a benchmark for what students in the during COVID-19 cohort should have achieved, and this is explored in the Discussion.

Method

Participants

This study included 2,453 first grade students in Norway (also see Table 1). Two-thirds of these students ($N = 1,636$) attended first grade in the 2018–2019 school year (before COVID-19 cohort). This was the school year before the COVID-19 pandemic, and all students were taught at school for the full school year. The remaining participants ($N = 817$) attended first grade during the year of the COVID-19 pandemic in 2019–2020 (during COVID-19 cohort). These students were not taught at school from March 12 to April 27 due to the institution nationwide of emergency remote instruction in Norway. These two cohorts (before and during COVID-19) represented 4% of all first grade students in Norway across the 2 years combined.

The participating students came from 200 classrooms, in 60 schools, in four municipalities. Two of the municipalities included major Norwegian cities, while the other two were more rural. These schools represented 2% of all schools in Norway. The average school size for the before and during COVID-19 cohorts were 462.1 ($SD = 170$) and 439.3 ($SD = 164.2$), respectively. The average class in the before COVID-19 cohort received 55 teacher hours ($SD = 11.8$) per student, whereas the average class in the during COVID-19 cohort received 55.4 teacher hours ($SD = 8.6$). This metric is the number of instructional hours divided by the number of students. Most of the first grade teachers were certified: 96.4% in the before COVID-19 cohort ($SD = 5.1$) and 95.5% in the during COVID-19 cohort ($SD = 5.1$). On average, there were 88.2 students

Table 1
Participant Characteristics

Characteristic	Before COVID-19 cohort		During COVID-19 cohort		χ^2
	Participants (<i>N</i>)	Participants (%)	Participants (<i>N</i>)	Participants (%)	
Gender					0.035
Boys	801	49	396	48.5	
Girls	835	51	421	51.5	
Language					9.849**
Bilingual	193	11.8	129	15.8	7.270**
L1	1,353	82.7	633	77.5	9.308**
L2	90	5.5	55	6.7	1.271

Note. COVID-19 = coronavirus disease 2019; L1 = native Norwegian speakers; L2 = nonnative Norwegian speakers. The chi-square values refer to tests of independence. Specifically, the values for gender and language refer to overall tests of independence. Furthermore, the chi-square value for bilingual refers to the comparison between bilingual versus L1 and L2, and so on for the other language categories.

** indicates $.001 < p \leq .01$.

($SD = 35.23$) per special education assistant in the before COVID-19 cohort in each school and 92.9 ($SD = 34.1$) in the during COVID-19 cohort. The average score on national tests for schools for the before COVID-19 cohort students was 51.2 ($SD = 2.9$); it was 50.6 ($SD = 2.3$) for the during COVID-19 cohort. Refer to Table 2 for a summary of these statistics.

Fifty-one percent of students in the before COVID-19 cohort ($N = 835$), and 51.5% of students in the during COVID-19 cohort ($N = 421$) were girls. The proportion of girls and boy was not significantly different between the before and during cohorts ($p = .85$). The before COVID-19 cohort included 90 first grade students (5.5%) who had a language other than Norwegian as their first language (L2 group), whereas the during COVID-19 cohort included 55 such students (6.7%). There were 193 students in the before COVID-19 cohort (11.8%) who were bilingual with Norwegian as one of their first languages. For the during COVID-19 cohort there were 129 of these students (15.8%). It was not possible in Norway to collect systematically information on languages other than Norwegian that children speak, so we do not report such information here.

There was a statistically significant difference between the before and during cohorts in the proportion of students in the three language groups ($p = .007$). Follow-up analyses demonstrated that before and during COVID-19 students differed in terms of proportion of students who were native Norwegian speakers only ($p = .002$) and bilingual speakers with Norwegian as one of their first languages

($p = .007$). However, there was no significant difference in L2 speakers who had a language other than Norwegian as their first language ($p = .26$). Thus, there were more native Norwegian only speakers in the before COVID-19 cohort than the during COVID-19 cohort, and there were more bilingual students in the during COVID-19 cohort than the before COVID-19 cohort.

Multiple indices provided evidence that the participating students in the before and during COVID-19 cohorts were representative of students nationally. While there are no national, standardized tests in writing in Norway, tests of reading, mathematics, and English are administered in fifth, eighth, and ninth grade for students in compulsory school (i.e., Grades 1–10). The participating schools average scores across the two cohorts for the three subjects was 50.9 ($SD = 2.7$), where averages for the before and during COVID-19 cohorts were 51.2 ($SD = 2.9$) and 50.6 ($SD = 2.3$), respectively. The population average is 50.0, with a standard deviation of 10.

Furthermore, the proportion of boys and girls in the before and during COVID-19 cohorts were similar to the proportion in the Norwegian population. In 2018–2019, 48.6% of the population of Norway were girls, and 51% of the before COVID-19 cohort were girls (within the 95% confidence interval [CI] of the sample estimate of this study). In 2019–2020, 48.9% of the national population were girls, and 51.5% of the during COVID-19 cohort students were girls (within the 95% CI of the sample).

Although there are no publicly available registers of language backgrounds in Norway, 8% of students in 2018–2019 were entitled to extracurricular language instruction, and 7.9% of students received such instruction in 2019–2020. These statistics are relatively close to the 5.5% and 6.7% of students in the before and during COVID-19 cohorts, respectively, who did not speak Norwegian as their first language.

While the size of schools from which students in the before and during COVID-19 cohorts combined were drawn ($M = 454$; $SD = 168.4$), was larger than the national average size of schools ($M = 225$, $SD = 166$), the four municipalities participating in this study were representative of the population range of municipalities in Norway. It ranged from large municipalities (population of 673,469, representing 13% of the population) to average sized municipalities (population of 13,958, representing 3% of the population) to smaller municipalities (population of 6,882, representing 1% of the population). Additionally, the proportion of certified teachers (96% average across both cohorts) and the school hours

Table 2
School Characteristics

Characteristic	Before COVID-19 cohort		During COVID-19 cohort		Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
School size (<i>N</i> students)	462.1	170	439.3	164.2	0.136
Teaching hours	55	11.8	55.4	8.6	0.039
Proportion certified teachers (%)	96.4	5.1	95.5	5.1	0.176
Students per special education assistant	88.2	35.23	92.9	34.1	0.136
National test score	51.2	2.9	50.6	2.3	0.229

Note. COVID-19 = coronavirus disease 2019.

per student (55 hr across both cohorts) for the teachers teaching the first grade students in this study were similar to national average, which was 95.5% and 61 hr, respectively. Also, the average number of students per special education teacher (88.2 across the before and during COVID-19 cohorts) was similar to the national average, which was 83.8 ($SD = 98.4$).

Sampling Procedures

The recruitment of students was done in the context of a large scale writing intervention study (Skar, Aasen, & Jølle, 2020). First grade students in the before COVID-19 cohort participated in a single test administration in May/June of 2019. This occurred before the COVID-19 pandemic and before schools were assigned to treatment and control conditions. All of the first grade students tested at this time were included in the before COVID-19 cohort. None of these students participated in a writing intervention before they were tested.

Students in the during COVID-19 cohort were assessed on the same writing measures in May/June of 2020. Due to COVID-19, schools in Norway moved from in class to digital instruction about 5 weeks before this group of students was tested. The during COVID-19 cohort only included first grade students from the 22 schools assigned to the control condition. Including students from schools receiving the writing treatment would have biased any comparison involving first grade students from the before and during COVID-19 cohorts. Because the during COVID-19 cohort only included control students, it included about one half of the students included in the before COVID-19 cohort.

The sampling scheme was motivated by the desire to obtain the largest sample size possible, as statistical power is always a consideration. Moreover, as we explain below, multilevel regression modeling was the main statistical methodology used for this study, and we wished to ensure there were no issues with the stability of the model estimates. Thus, measurements from all students were included in the before COVID-19 cohort, as students had not yet been assigned to or received treatment (i.e., writing intervention program). Moreover, there were no statistically significant differences for writing quality, handwriting fluency, and attitude toward writing, after controlling for clustering, between students who were eventually assigned to the writing intervention and control group in the before COVID-19 cohort (all $ps > .31$). However, we only included students from the control condition in constructing the during COVID-19 cohort. Like students in the before COVID-19 cohort, children in the during COVID-19 cohort were taught writing as usual. This made these children as similar as possible to the children in the before COVID-19 cohort, as they attended the same schools and received a business as usual approach to writing, with the exception that they were enrolled in school during the COVID-19 pandemic and emergency remote instruction. This difference was the focal point of this investigation.

The schools participating in this and the larger study (Skar, Aasen, & Jølle, 2020) were recruited with the help of executive officers from different municipalities. All schools in a municipality were invited to participate. Students of participating schools were invited to participate by means of parental consent. In the before COVID-19 cohort, 2,459 students were approached, and 2,294 (93.4%) consented to participation. In the during COVID-19 cohort, 1,333 students were approached, and 1,170 (87.8%)

consented to participation. For the before and during COVID-19 cohorts, respectively, 71% and 69.8% of participating students participated in all test administration, resulting in the net sample of 1,636 and 817 for the two cohorts.

Writing Measures and Covariates

Students in both the before and during COVID-19 cohorts completed the same three writing tasks. This included writing a letter that yielded a measure of writing quality, a copying task which provided an assessment of handwriting fluency, and a short survey that measured attitude toward writing. There were a total of eight covariates included in the analysis (described below).

Writing Quality

Students wrote a letter using a prompt developed by the *Functional Writing in Primary School* project (Skar, Aasen, & Jølle, 2020). Teachers introduced this writing task by reading aloud a letter from researchers at the local university. The letter explained that these researchers were interested in children's favorite recess activities. It further indicated students could help researchers at the University to understand this topic better by writing "a letter where you describe what you usually play with during recess time."

Before writing the letter, teachers led a discussion with students about what researchers do, what students like to do during recess, why they like each activity, what it means to describe something, and what is included in a letter. Teachers also showed students a picture of a school playground with children engaged in multiple activities (e.g., kicking a ball, playing marbles, swinging, or building a sand castle). Students were informed they had 45 min to write their letter. We asked teachers not to provide students with help as they wrote.

Students' letters were scored for writing quality on the following eight traits: audience awareness, organization, content relevance, vocabulary, sentence construction, spelling, legibility, and punctuation. Each trait was scored using a separate 1 to 5 rating scale. The rating scales were developed and validated for use in Grades 1–3 by Skar, Jølle, and Aasen (2020), and also used in Skar et al. (2021), who reported that, on average, girls outperformed boys, and students in higher grades outperformed students in lower grades. Higher scores represented better quality for the trait assessed. The scores for the eight traits were summed into a single, scaled score.

The rating scale for each writing attribute had descriptors and benchmark text for each score point (1 to 5). The rating scale "audience awareness" focused on how text explicitly or implicitly addressed a reader. The descriptor for a score of 5 (the highest score) stated that: "The text addresses the reader in the assignment in a generally relevant manner and takes into account the reader's need for knowledge of participants/characters, circumstances, and events. The text may contain traces of the student's voice with reflective or evaluating utterances." For a score of 1 (the lowest score), the descriptor stated that "To understand the text, a conversation with the writer is required." The 1–5 point scale for "organization" concerned the internal structure of the text at the macro-, meso-, and microlevel. "Content relevance" focused on whether the letter presented relevant content with regards to the writing task assigned; a higher score indicated more relevant content, but

not necessarily longer text. “Vocabulary” involved the repertoire of words used in the text, with the highest level descriptor stating that: “The text contains a repertoire of words and expressions that are relevant to the task. In some cases, there is use of specialized and abstract words, and/or creative forms of expression.” “Sentence construction” concerned syntactical variation and syntactic complexity, whereas “spelling” concentrated on increasing mastery of correct spelling in text. “Legibility” focused on the legibility and shape of letters, including if letters were reversed or not. Finally, “punctuation” involved the correct use of punctuation marks. The descriptor for the lowest level stated that: “the text does not contain punctuation marks,” whereas the descriptor for the highest level stated that: “mostly correct use of different punctuation marks. The use of comma may occur.” Descriptors for scores of 1 and 5 for all scales are presented in the Appendix.

The quality of students’ letters were scored by a pool of 34 trained raters at two time points. Papers produced by students in the before COVID-19 cohort were scored after these students were tested as were papers generated by students in the second cohort (i.e., during the COVID-19 pandemic). These raters were recruited from staff and master-level students in teacher education at the University where the first author was employed. In the training sessions, the rating scales were presented and practiced, and each rater was given opportunities to address any questions that arose while practicing. Further, the raters sent their first 25 ratings to the first author who provided feedback on individual raters’ relative harshness and consistency. There was also an e-mail list to which the raters could address questions that arose during the rating session. When a question was posted on the list, ratings stopped, and raters were given directions addressing the question posed.

Each text was assessed independently by two raters. To provide an empirical link between raters and rating occasion, fifty of the students’ letters were scored by all raters. This empirical link was used to scale the raw scores from the two occasions with a many-facet Rasch model (MFRM; Linacre, 2017). The following model was used:

$$\log[P_{nij(k)}/P_{nij(k-1)}] = B_n - E_i - C_j - F_x$$

where $P_{nij(k)}$ represents the probability of student n , rated on rating scale i , by rater j , receiving a score of k , and $P_{nij(k-1)}$ represents the probability of the same student under the same conditions receiving a score of $(k - 1)$. B_n is the ability for person n , E_i is the difficulty of rating scale i , and C_j is the severity of rater j . Lastly, F_x represents the point on the logit scale where category k and $(k - 1)$ are equally probable.⁴

The data from the two occasions fit the model well. The “reliability of separation” measure, which is a MFRM analog to Cronbach’s α was .94 for papers from the before COVID-19 cohort and .96 for papers from the during COVID-19 cohort, which indicate a reliability fit for high stakes assessment in the measure of students’ writing quality. The standardized residuals also indicated good fit, as 3.90% were in the range of 2–3 and .59% exceeding 3 for papers from the before COVID-19 cohort, and 4.10% were in the range of 2–3 and .80% exceeding 3 for papers from the during COVID-19 cohort. Standardized residuals above 2 should not exceed 5%, and standardized residuals above 3 should not exceed 1% (Eckes, 2011). We used the MFRM estimates to adjust scores

using a compensatory principle where the final writing quality score was a single averaged score adjusted for varying difficulties between rating scales and differing rater severity. The FACETS software (Linacre, 2017) was used to convert the MFRM estimates to a Text Quality Scale, where scores were not limited to integers.

Handwriting Fluency

Teachers administered a copying task designed to assess students’ handwriting fluency (i.e., rate at which letters of the alphabet were produced correctly). Students were directed to copy a short paragraph as quickly and accurately as possible within a 90 s period (Graham et al., 1997). The paragraph was from the group Diagnostic Reading and Aptitude and Achievement Tests (Monroe & Sherman, 1966). Students were shown an instructional video that explained to students how the copying task was to be completed. After watching the video, the students followed along as the teacher read the paragraph aloud. With the written paragraph in front of them, students were told to copy the paragraph once the teacher said “Start,” and to stop copying when directed to do so at the end of 90 s. While the paragraph students were asked to copy was written in article, students were not asked to copy it in article or another script such as cursive. There is no specific mandate in Norway to teach article or cursive handwriting. This is determined by teachers.

The score for handwriting fluency was obtained by tallying all letters from the paragraph written correctly in the 90 s interval. The trained coders did not count as correct letters incorrectly written, skipped, or correctly written but not a match to text. To estimate interrater reliability, 10% of all tasks were rescored by an independent coder. Reliability was satisfactory ($\kappa = .812$, intra-class correlation coefficient [ICC] = .99). The score for handwriting was divided by 1.5 to provide an estimate of handwriting fluency per min.

Attitude Toward Writing

Students completed a short survey that included four items designed to measure attitude toward writing. The items were: “I liked the composition task,” “I am satisfied with my text,” “I am satisfied with my effort,” and “I like to write.” Students rated each statement using a 3-point Likert-type scale. The highest point on the scale was denoted by three stars (higher attitude toward writing), while the lowest point was a single star (lower attitude toward writing). An exploratory factor analysis, using principal axis factoring, resulted in a one-factor solution, with all items loading at .58 or higher. This factor accounted for 53.80% of the variance (coefficient $\alpha = .71$). The score for the attitude toward writing measure was the average score of all four items.

Covariates

The following covariates at the school-level were included in the analysis: national test result, school size, proportion of certified teachers, students per special education teacher, and school hours per student. Data for these variables, which were described earlier were collected from the database “Skoleporten” [The School

⁴ There were two rating occasions and scores were scaled for each occasion. To make scaled scores from each occasion comparable, we used scaled scores from the fifty texts as “anchor values” when scaling scores from the second rating occasion.

Gate], hosted by the governmental agency the Norwegian Directorate for Education and Training (<https://skoleporten.udir.no/>). Covariates at the student-level were gender and language background. At student level, we also included a student cohort membership variable (before COVID-19 cohort and during COVID-19 cohort).⁵ Data on gender and language background were collected by participating students' teachers who identified each student's gender and whether the student learned Norwegian or another language first, or both. A student was coded as having learned Norwegian first (denoted "L1"), another language than Norwegian first (denoted "L2"), or Norwegian and another language from birth (denoted "bilingual"). The teacher did not assess the students' language skills when noting language background, which means that the level of proficiency within the L2- and bilingual groups might be quite varied.

Table 3 presents the descriptive statistics for student level measures. We were unable to collect data on school-level variables for 36 students. To refrain from excluding them from the analysis a mean imputation strategy was used. We believe this strategy was reasonable given the relatively small proportion of missing data.

Procedures

Procedures to collect data were designed to fit the Norwegian educational context. As testing and assessment is controversial in Norway and formal grades are not introduced until Grade 8, we developed test procedures that would mimic regular instruction as much as possible, while allowing the collection of data under uniform conditions. Other reasons why we mimicked regular practices is that formal instruction in first grade has been a topic of public and heated debate in Norway (e.g., Lundetræ, 2018), and the transition from kindergarten to school was a topic of a relatively recent report from the Norwegian Knowledge Center for Education (Lillefjord et al., 2015). Thus, we did not want to upset teachers, students or guardians by introducing unusual formats in the classrooms.

The first author consulted with first grade teachers to develop a strategy for administering the three writing tasks that would balance the needs of contextual alignment and standardization. It was decided that writing tasks would be administered during a regular lesson, and students would be offered 45 min to complete the letter writing tasks, which is similar to the time frame students apply to tasks that are part of regular class instruction in Norway. The copying task was also administered as part of regular instruction, but only took 1.5 min to complete it.

We provided teachers with instructions for administering the three writing tasks. This included written and video instruction for each task. The video instruction for the copying task was shown to students. Task administration was counterbalanced so that one half of the teachers administered the letter writing task and attitude toward writing measure first and the copying task second, whereas the remaining teachers did just the opposite. After administering the tasks to students, teachers posted test materials on a secure site accessible only by the researchers. The data was collected on two occasions (May/June 2019 and May/June 2020). On both occasions, teachers had a 15-day window to complete data collection.

Table 3

Student-Level Outcome Variables Broken Down by Cohort, Gender, and Language Background

Characteristic	Before COVID-19 cohort		During COVID-19 cohort	
	<i>N</i>	<i>M (SD)</i>	<i>N</i>	<i>M (SD)</i>
HW fluency				
Bilingual				
Boy	94	15.5 (7.3)	70	12.8 (6.7)
Girl	99	15.9 (8.1)	59	14.6 (6.4)
L1				
Boy	667	14.4 (6.7)	297	12.2 (6.3)
Girl	686	17.2 (8.6)	336	14.3 (6.8)
L2				
Boy	40	14.7 (6.2)	29	10.6 (5.5)
Girl	50	17.7 (9.6)	26	10.7 (6.6)
Total	1,636	15.9 (7.9)	817	13.2 (6.6)
Writing quality				
Bilingual				
Boy	94	2.8 (0.4)	70	2.2 (0.7)
Girl	99	2.9 (0.4)	59	2.5 (0.6)
L1				
Boy	667	2.7 (0.4)	297	2.4 (0.6)
Girl	686	2.8 (0.4)	336	2.6 (0.6)
L2				
Boy	40	2.7 (0.4)	29	2.1 (0.7)
Girl	50	2.8 (0.4)	26	2.3 (0.7)
Total	1,636	2.8 (0.4)	817	2.5 (0.6)
Attitude				
Bilingual				
Boy	94	2.6 (0.5)	47	2.5 (0.5)
Girl	99	2.7 (0.4)	47	2.5 (0.5)
L1				
Boy	667	2.5 (0.5)	247	2.4 (0.5)
Girl	686	2.6 (0.4)	277	2.5 (0.4)
L2				
Boy	40	2.6 (0.5)	25	2.6 (0.4)
Girl	50	2.7 (0.3)	25	2.4 (0.7)
Total	1,636	2.6 (0.5)	668	2.5 (0.5)

Note. COVID-19 = coronavirus disease 2019; HW = handwriting; Attitude = attitude towards writing; bilingual = students with Norwegian and at least one other native language; L1 = native Norwegian speakers; L2 = nonnative Norwegian speakers; 149 students did not have an attitude towards writing score, resulting in $N = 2,304$ for attitude towards writing.

Results

Power Analysis

The analytical strategy applied in this investigation was multilevel, or hierarchical, regression analysis. For regression models, the estimate of statistical power may focus on inference for a coefficient, a variance component, or on the means of particular groups (Snijders, 2005). Oftentimes, the focus is on a regression coefficient, and in the current study we focused on the coefficient indicating the effect of the COVID-19 pandemic and emergency remote instruction on students' scores on each of the writing measures. In multilevel modeling, the statistical power of a test for a given regression coefficient depends on sample size at each level (total number of students, classes, and schools; Snijders & Bosker,

⁵ Unfortunately, we did not have data on students' socioeconomic background.

1999) as well as myriad other factors. Given the computational complexities introduced by multiple levels of nesting, statistical simulation is a current method of assessing statistical power. Investigations using the MLPowSim software (Browne & Golali-zadeh, 2021) to assess statistical power of the regression coefficient for detecting differences between the before COVID-19 and during COVID-19 cohorts indicated high power (>99%) for the current sample sizes. The simulation was based on estimates derived from results from a study by Skar et al. (2021), who reported on similar measures.

Analytical Strategy

For each writing measure (writing quality, handwriting fluency, and attitude toward writing; Table 3), a three-level linear hierarchical regression model was computed in which students, classrooms, and schools represented Levels 1, 2, and 3, respectively. The models allowed for random intercepts for classes and schools. For each writing measure, the following three-level hierarchical models were fit: (1) a “null” model (no predictors) to assess the correlation structure via the ICC, (2) a model with all covariates except cohort membership (before COVID-19 cohort vs. during COVID-19 cohort), and (3) a model with all covariates including the cohort membership variable. The purpose of fitting models (2) and (3) was to calculate the difference in explained variation (R^2) between the two models, which was interpreted as the increase in R^2 Provided \times Specifically \times Differences in the writing performance of students in the before COVID-19 cohort and the during COVID-19 cohort. For purposes of interpretation and computational stability, standardized versions of continuous covariates were used in the three models.

Was the COVID-19 Pandemic and Emergency Remote Instruction Associated With Lower Writing Scores?

Table 4 displays effect sizes for all three models fit for each of the three writing outcome variables (writing quality, handwriting fluency, and attitude toward writing). For each outcome variable, Table 4 displays ICCs resulting from the null model (for three-level hierarchical models, there are two ICCs: one for Level 2 and

one for Level 3) and the R^2 values from the models without and with the student cohort membership variable (before COVID-19 cohort vs. during COVID-19 cohort). It can be seen that writing quality had the highest correlation within classes and schools, while attitude toward writing had the lowest correlation. Also, of the three outcomes, the student cohort membership variable (before COVID-19 cohort vs. during COVID-19 cohort) explained the most variation for writing quality (contributed R^2 almost 5%) and the least variation for attitude toward writing (contributed R^2 less than 1%). Effect size for R^2 , $f^2 = R^2/(1 - R^2)$, is reported in parentheses next to the R^2 values in Table 4. A rule of thumb interpretations for f^2 is that values of .02, .15, and .35 represent small, medium, and large effects (Lorah, 2018). Thus, the student cohort variable (before COVID-19 cohort vs. during COVID-19 cohort) had a small to medium effect for writing and quality and handwriting fluency and a small effect for attitude toward writing.

Table 5 displays the parameter estimates and p -values for the models including the student cohort variable (before COVID-19 cohort vs. during COVID-19 cohort). For categorical covariates gender, language, and cohort, the reference categories were boy, Norwegian, and students in the before COVID-19 cohort. Thus, students in the during COVID-19 cohort scored, on average, 2.1, .24, and .06 units lower on handwriting fluency, writing quality, and attitude toward writing, respectively, than students in the before COVID-19 cohort after controlling for nesting, gender, language background, and school characteristics. The difference was statistically significant for handwriting fluency ($p = .001$) and writing quality ($p < .001$) and statistically significant for attitude toward writing ($p = .036$). Also, girls scored significantly higher than boys, on average, on all three outcomes. Regarding the continuous covariates, national testing was the only one that was statistically significant for all three writing outcomes. For every 1 SD increase in national test scores, we expect, on average, handwriting fluency, writing quality, and attitude toward writing to increase by 1.014, .075, and .043, on average after controlling for nesting, gender, language background, and school characteristics.

After the models described above were examined, we subsequently fit three additional models to examine the interaction between language and cohort, one for each response variable. The models were similar to those in Table 5, but the new models contained an interaction term for language and cohort (before COVID-19 cohort vs. during COVID-19 cohort). Given that cohort is a binary variable and language contributes two slopes, this results in two interaction terms being added to the models. For the sake of brevity, partial results for the interaction terms only are presented in Table 6. For handwriting fluency, the interaction between L2 and during COVID-19 cohort interaction was statistically significant ($p = .005$). Thus, the decrease in handwriting fluency for the L2 language group in the during COVID-19 cohort was significantly more severe than for the other language groups in this cohort. In addition, for writing quality, both the interaction between bilingual and the during COVID-19 cohort as well as the interaction between L2 and the during COVID-19 cohort were statistically significant ($p < .001$). Therefore, the decrease in writing quality for students in the before COVID-19 cohort was more severe for the L2 and bilingual students than for the reference L1

Table 4
Effect Sizes for the Three Models Fit to Each of the Three Outcome Variables

Quantity	Outcome		
	HW fluency	Writing quality	Attitude
ICC (class)	0.271	0.305	0.043
ICC (school)	0.081	0.134	0.020
R^2 , model without cohort membership (f^2)	0.037 (0.038)	0.049 (0.052)	0.023 (0.024)
R^2 , model with cohort membership (f^2)	0.058 (0.062)	0.097 (0.107)	0.027 (0.028)
R^2 , difference (f^2)	0.021 (0.021)	0.048 (0.050)	0.003 (0.003)

Note. COVID-19 = coronavirus disease 2019; ICC = intraclass correlation coefficient; HW = handwriting; Attitude = attitude towards writing; cohort membership = before COVID-19 cohort vs. during COVID-19 cohort.

Table 5
Parameter Estimates and p-Values for Models With Cohorts for All Three Outcome Variables

Parameter	Outcome					
	Handwriting fluency		Writing quality		Attitude	
	Estimate	p-value	Estimate	p-value	Estimate	p-value
Intercept	14.716	<.001	2.651	<.001	2.483	<.001
National test	1.014	.024	0.075	.029	0.043	.008
School size	0.014	.973	0.032	.302	−0.004	.765
Teacher certification	0.091	.826	0.033	.306	0.026	.098
Students/SPED teacher	−0.817	.047	−0.025	.427	0.002	.867
Instruction hours/student	0.483	.337	0.055	.149	0.023	.206
Gender						
Ref level boy						
Girl	2.542	<.001	0.181	<.001	0.130	<.001
Language						
Ref level L1						
Bilingual	−0.698	.117	−0.033	.267	0.064	.040
L2	−0.716	.274	−0.128	.003	0.060	.173
During COVID-19 cohort	−2.131	.001	−0.240	<.001	−0.056	.036

Note. COVID-19 = coronavirus disease 2019; Attitude = attitude towards writing; SPED = special education; continuous covariates were standardized LI = native Norwegian speakers; L2 = nonnative Norwegian speakers.

group for students in this cohort. There were no statistically significant interactions for attitude toward writing.

Discussion

The COVID-19 pandemic and the implementation of emergency remote instruction for students around the world is viewed by some as the largest educational disruption of all times (e.g., Azoulay, 2020; Winthrop, 2020). The stress that the COVID-19 pandemic, in person school closures, and the move to emergency remote instruction placed on educational systems, teachers, and students was a global phenomenon. Multiple organizations predicted this would negatively impact students' learning (e.g., Di Pietro et al., 2020; Schleicher, 2020). The current study examined this proposition in terms of children's writing by comparing quality of writing, handwriting fluency, and attitude toward writing of first grade Norwegian students during the COVID-19 pandemic and following emergency remote instruction with first grade students in the same schools a year earlier before the COVID-19 pandemic.

The COVID-19 Pandemic and Emergency Remote Instruction Negatively Impacted First Grade Students' Writing

Even though Norwegian students score above the average on OECD assessments, Norway has a strong digital infrastructure

(98% of homes have Internet access), and the use of digital technology for teaching has been emphasized for decades (Blikstad-Balas et al., in press), we predicted that the COVID-19 pandemic and the move to emergency remote instruction delivered online would have a negative effect on students' writing. We based this prediction on multiple factors. This included the presumed negative effects of the COVID-19 pandemic and emergency remote instruction on teaching, the stress created by this situation, and the possibility that teachers might not make optimal use of the digital resources available to them when providing emergency remote instruction online (see Di Pietro et al., 2020; Schleicher, 2020). It also included recent survey data showing that many young students in Norway (Grades 1 to 4) had little contact with their teachers and less time was devoted to schooling during emergency remote instruction than before schools were closed to in person learning (Blikstad-Balas et al., in press). Our prediction was further bolstered by evidence showing a decline in students' writing when school is discontinued over the summer break (Jesson et al., 2009; Meyer et al., 2020; Pachter et al., 2015).

The prediction that the COVID-19 pandemic and emergency remote instruction would negatively impact first grade students' writing was supported. When compared with first grade students in the same schools a year earlier before the pandemic began, first grade students educated during the pandemic and the emergency remote instruction evidenced statistically significant lower outcomes for writing, after we controlled for variance related to

Table 6
Parameter Estimates and p-Values for Interaction Terms

Parameter	Outcome					
	Handwriting fluency		Writing quality		Attitude	
	Estimate	p-value	Estimate	p-value	Estimate	p-value
Ref level L1						
Bilingual × During COVID-19 Cohort	−0.040	.965	−0.238	<.001	−0.036	.581
L2 × During COVID-19 Cohort	−3.761	.005	−3.546	<.001	−0.117	.187

Note. LI = native Norwegian speakers; L2 = nonnative Norwegian speakers; COVID-19 = coronavirus disease 2019.

national test scores, school size, proportion of certified teachers, students per special education teacher, school hours per student, as well as student gender and native language. This included lower scores for writing quality, handwriting fluency, and attitude toward writing. The effect of the COVID-19 pandemic and emergency remote instruction was most pronounced for writing quality and handwriting fluency, representing small to medium effects. They were less pronounced for attitude toward writing, as the pandemic and move to emergency remote instruction resulted in a small negative effect.

A study by Skar et al. (2021) provides a useful comparison point for thinking about the findings from the current study. Second grade Norwegian students from the same schools that participated in this study scored four-tenths of a point higher, on a 5-point scale, than their first grade counterparts when both were tested at the end of the school year (May/June 2019) before the COVID-19 pandemic. The first grade students in that study were in the before COVID-19 cohort in the current investigation. These second-grade students' handwriting fluency was also 12.5 letters per minute faster than their first grade peers. For the first grade students in the current study, who came from the same schools as the second graders and were tested with the same measures in May/June 2020 during the pandemic but just after emergency remote instruction (during COVID-19 cohort in the present study), they would need to increase writing quality and handwriting fluency gains by 175% and 130%, respectively, during the ensuing year to score at the same level of performance as their second grade counterparts did in the prepandemic year. It is highly likely that this needed increase will be even larger for students classified as L2 (i.e., native language was not Norwegian), given that the quality of these students' papers and the speed of their handwriting was lower than that of native Norwegian speakers and bilingual students.

Assuming that the findings from this study extend to other Norwegian children in first grade and beyond as well as students in other countries, policymakers and schools need to consider how they can mitigate writing losses that occurred as a result of the COVID-19 pandemic and emergency remote instruction. The losses in writing quality are particularly troublesome because students increasingly become better writers with experience and schooling (Graham, 2006), but it is especially difficult to overcome literacy issues in later grades that have their genesis early in students' educational journey (Slavin et al., 1989). One possible solution is to provide extra writing instruction in the immediate and upcoming years, as increasing how much writing instruction students receive improves writing performance (e.g., Graham & Santangelo, 2014). Another possible solution is to provide writing instruction during students' summer vacation or during other breaks in the school year. Providing instruction when students are not in school has enhanced literacy capabilities previously (e.g., Nicholson & Tiru, 2019).

The handwriting fluency losses evidenced by first grade students in this study are also troubling. Until handwriting becomes fluent and automatic, it consumes an inordinate amount of cognitive effort when writing, and it interferes with other writing processes such as content generation and planning (Graham et al., 1997). For children whose handwriting fluency is inordinately slow over time, it can also lead them to avoid writing when possible and develop a negative mindset toward writing (Berninger, 1999). This

may have contributed to some of the loss in writing quality that occurred in this investigation.

To overcome the handwriting fluency losses evidenced by first grade students in this study, schools and teachers need to provide extra instruction in how to form letters efficiently. They also need to ask children to write more frequently in second grade and possibly beyond. This is the most common means for increasing handwriting fluency (Graham, 2009). This recommendation concerning frequency of writing can further help to mitigate losses in writing quality too (see Graham et al., 2015).

For attitude toward writing, the outcome from this study was relatively neutral. While we obtained a statistical significant difference in the attitude toward writing of first grade students before and during the pandemic, attitude scores for these two groups differed by only one-tenth of a point. There are several possible reasons why the COVID-19 pandemic and emergency remote instruction did not impact attitude toward writing more strongly. One, young children may have difficulty evaluating their feelings and beliefs about writing, making such assessments less reliable and certain. If this was the case, the added noise and error of the assessment may have blurred any differences that existed between the two groups of students. Two, the average attitude toward writing scores for the two groups of children (2.6 and 2.5) were close to the ceiling of the three-point scale. This may have restricted students' estimation of their attitude toward writing, minimizing any obtained differences. Three, attitudes can be resistant to change (McGuire, 1985), and change in attitudes can depend on multiple factors including how important or relevant an issue is to an individual (Petty & Cacioppo, 1986). It is possible that first grade children do not give much consideration to how they feel about writing, reducing the likelihood of any major changes in their attitudes toward writing as a consequence of the COVID-19 pandemic and emergency remote instruction. It is also possible that potential negative effects of the COVID-19 pandemic and emergency remote instruction is countered by children's attitude toward non-school writing, which is likely to be less evaluative.

Limitations and Future Research

The two cohorts of first grade students participating in this study (before COVID-19 cohort and during COVID-19 cohort) attended the same schools, but in different years (2018/2019 and 2019/2020). While children in these two groups were similar in terms of gender, they differed in terms of their language status. First grade students from the second cohort had proportionally fewer native Norwegian speakers and more students who were bilingual (with Norwegian as a native language). Because students' language status is related to writing performance (Camping et al., 2020), this may have contributed to obtained writing differences between the before and during COVID-19 cohorts. We believe that such effects did not influence this study for two reasons. One, we controlled for language status in our analyses. Two, the proportional differences between the two groups were small. Future studies examining the effects of the COVID-19 pandemic and emergency remote instruction, however, do need to take into account language status.

Even though students in this study were generally representative of first grade students in Norway, we were unable to randomly select children from the general population. Unfortunately, this will not be the last pandemic, nor was it the first in modern history

(e.g., Howard & Howard, 2012). Consequently, politicians, policy-makers, and school systems need to be more prepared for the next pandemic. This includes determining how education will proceed and how the impact of any new pandemic is assessed.

We assessed three aspects of first grade students' writing in the current study (writing quality, handwriting fluency, and attitude toward writing). Many other aspects of children's writing could have been assessed, such as how students write, what they know about writing, and how much they write. Hopefully, subsequent investigations of the pandemic will assess a broader range of writing capabilities and students in both elementary and secondary grades.

Our study would have been enriched if we had obtained information on what aspects of writing teachers taught during emergency remote instruction, what role parents played in their children's education during this period of online learning, and how many sessions students actually attended during this close to 7-week period of remote instruction. In addition, we had no direct assessment of students' mastery of Norwegian language, whether they were native speakers, bilingual, or learning Norwegian as a second language. Such information would have helped us better situate our findings, and we hope that future research is able to collect such information.

Finally, this study involved a single country. There is a need to investigate the educational effects of the COVID-19 pandemic and new pandemics that occur across a broader array of countries, including more and less affluent nations and ones that differ culturally. It is especially important to compare how these countries respond to such pandemics, and how their responses differentially affect students, teachers, and parents.

Concluding Comments

If the presumed negative effects of the COVID-19 pandemic and ensuing emergency remote instruction on students' learning are further confirmed in other analyses, policymakers, leaders of educational systems, and individual teachers must consider how they can reduce and even surmount these effects. This is especially important in writing, as it is a complex skill that is not acquired naturally, but must be taught (Graham, 2006).

References

- Alexander, P. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, 32(8), 10–14. <https://doi.org/10.3102/0013189X032008010>
- Azoulay, A. (2020). COVID: 12 education responses. <https://rb.gy/kuckhk>
- Berninger, V. W. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly*, 22(2), 99–112. <https://doi.org/10.2307/1511269>
- Blikstad-Balas, M., Dalland, C. P., Roe, A., & Klette, K. (in press). Home-schooling in Norway during the pandemic - Digital learning with unequal access to qualified help at home and unequal learning opportunities provided by the. In F. M. Reimers (Ed.), *Education and COVID-19. Disruptions to educational opportunity during a pandemic*. Springer Nature.
- Bonal, X., & Gonzalez, S. (2020). The impact of lockdown on the learning gap: Family and school divisions in times of crisis. *International Review of Education*, 66, 635–655. <https://doi.org/10.1007/s11159-020-09860-z>
- Browne, W., & Golarizadeh, M. (2021). MLPowSim [Computer software]. <http://www.bristol.ac.uk/cmm/software/mlpowsim/>
- Camping, A., Graham, S., Ng, C., Aitken, A., Wilson, J., & Wdowin, J. (2020). Writing motivation of middle school emergent bilingual students. *Reading and Writing: An Interdisciplinary Journal*, 33(9), 2361–2390. <https://doi.org/10.1007/s11145-020-10046-0>
- Carroll, J. B. (1989). The Carroll model: A 25 year retrospective and prospective view. *Educational Researcher*, 18(1), 26–31. <https://doi.org/10.3102/0013189X018001026>
- Cordeiro, C., Castro, S. L., & Limpo, T. (2018). Examining potential sources of gender differences in writing: The role of handwriting fluency and self-efficacy beliefs. *Written Communication*, 35(4), 448–473. <https://doi.org/10.1177/0741088318788843>
- Daniel, S. J. (2020). Education and the COVID-19 pandemic. *Prospects*, 49, 91–96. <https://doi.org/10.1007/s11125-020-09464-3>
- Davis, C., Grooms, J., Ortega, A., Rubalcaba, J., & Vargas, E. (2021). Distance learning and parental mental health during COVID-19. *Educational Researcher*, 50(1), 61–64. <https://doi.org/10.3102/0013189X20978806>
- Di Pietro, D., Biagi, G., Costa, P., Karpinski, Z., & Mazza, J. (2020). *The likely impact of COVID-19 on education: Reflections based on existing literature and recent international datasets* (EUR 30275 EN). Publications Office of the European Union.
- Eckes, T. (2011). *Introduction to many-facet Rasch measurement: Analyzing and evaluating rater-mediated assessments* (1st ed.). Peter Lang. <https://doi.org/10.3726/978-3-653-04844-5>
- Ekholm, E., Zumbrunn, S., & DeBusk-Lane, M. (2018). Clarifying an elusive construct: A systemic review of writing attitude. *Educational Psychology Review*, 30(3), 827–856. <https://doi.org/10.1007/s10648-017-9423-5>
- Fauzi, I., & Khusuma, I. (2020). Teachers' elementary school in online learning of COVID-19 pandemic conditions. *Jurnal Iqra*, 5(1), 58–70. <https://doi.org/10.25217/ji.v5i1.914>
- Graham, S. (2006). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (pp. 457–478). Erlbaum.
- Graham, S. (2009). Want to improve children's writing? Don't neglect their handwriting. *American Educator*, 33(4), 20–27.
- Graham, S. (2018a). A writer(s) within community model of writing. In C. Bazerman, V. Berninger, D. Brandt, S. Graham, J. Langer, S. Murphy, & M. Schleppegrell (Eds.), *The lifespan development of writing* (pp. 258–279). National Council of English. <https://doi.org/10.1080/00461520.2018.1481406>
- Graham, S. (2018b). The writer(s)-within-community model of writing. *Educational Psychologist*, 53(4), 258–279. <https://doi.org/10.1080/00461520.2018.1481406>
- Graham, S. (in press a). Teaching writing in the digital age. In T. L. Good & M. McCaslin (Eds.), *Educational psychology section; D. Fisher (Ed.). Routledge encyclopedia of education* (Online). Taylor & Francis.
- Graham, S. (in press b). Writer(s)-Within-Community model of writing as a lens for studying the teaching of writing. In R. Horowitz (Ed.), *The Routledge Handbook of International Research on Writing* (Vol. II). Routledge.
- Graham, S., & Hebert, M. (2011). Writing-to-read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81(4), 710–744. <https://doi.org/10.17763/haer.81.4.t2k0m13756113566>
- Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, 27(9), 1703–1743. <https://doi.org/10.1007/s11145-014-9517-0>
- Graham, S., Berninger, V., Abbott, R., Abbott, S., & Whitaker, D. (1997). The role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, 89(1), 170–182. <https://doi.org/10.1037/0022-0663.89.1.170>

- Graham, S., Berninger, V., Weintraub, N., & Schafer, W. (1998). The development of handwriting fluency and legibility in grades 1 through 9. *The Journal of Educational Research*, 92(1), 42–52. <https://doi.org/10.1080/00220679809597574>
- Graham, S., Harris, K. R., & Santangelo, T. (2015). Research-based writing practices and the common core: Meta-analysis and meta-synthesis. *The Elementary School Journal*, 115(4), 498–522. <https://doi.org/10.1086/681964>
- Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2018). Reading for writing: A meta-analysis of the impact of reading and reading instruction on writing. *Review of Educational Research*, 88(2), 243–284. <https://doi.org/10.3102/0034654317746927>
- Graham, S., McKeown, D., Kiuahara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, 104(4), 879–896. <https://doi.org/10.1037/a0029185>
- Graham, S., Skar, G., & Falk, D. (2021). Teaching writing in the primary grades in Norway: A national survey. *Reading and Writing*, 34, 529–563. <https://doi.org/10.1007/s11145-020-10080-y>
- Håland, A., Hoem, T. F., & McTigue, E. M. (2019). Writing in first grade: The quantity and quality of practices in Norwegian classrooms. *Early Childhood Education Journal*, 47(1), 63–74. <https://doi.org/10.1007/s10643-018-0908-8>
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *EDUCAUSE Review*, 27, 1–12.
- Howard, P., & Howard, J. (2012). Pandemic and pedagogy: Elementary school teachers' experiences of H1N1 influenza in the classroom. *Phenomenology & Practice*, 6(1), 18–35. <https://doi.org/10.29173/pandpr19852>
- Huber, S., Gunther, P., Schneider, N., Helm, C., Schwander, M., Schneider, J., & Pruitt, J. (2020). COVID-19 und aktuelle Herausforderungen in Schule und Bildung [COVID-19 and current challenges in schooling and learning]. <https://www.waxmann.com/index.php?eID=download&buchnr=4216>
- Jesson, R., McNaughton, S., Lai, M., Hsiao, S., & Leonard, A. (2009, April). Longitudinal patterns and problem of transfer in surface and deep features of writing in a schooling improvement context [Paper presentation]. *Annual Meeting of the American Educational Research Association, San Diego, CA*.
- Kimm, L., & Ashbury, K. (2020). Like a rug has been pulled under you: The impact of COVID-19 on teachers in England during the first six weeks of the U.K. lockdown. *Psyarxiv*. <https://psyarxiv.com/xn9ey/>
- Kirshner, J. (2020). School radio: Finding innovation in reaching remote learners in Belize. *International Journal of Education and Literacy Studies*, 8(3), 90–97. <https://doi.org/10.7575/aiac.ijels.v.8n.3p.90>
- Lillefjord, S., Børte, K., Halvorsrud, K., Ruud, E., & Freyr, T. (2015). *Til tak med positiv innvirkning på barns overgang fra barnehage til skole: En systematisk kunnskapsoversikt* [Programs with positive effect on the transmission from kindergarten to school: A systematic review.] (p. 72). Kunnskapssenter for Utdanning. <https://utdanningsforskning.no/contentassets/d66fcf1950f34b46a79a5fc647e40774/kunnskapssenter-overgangbarnehage-web.pdf>
- Linacre, J. M. (2017). *A user's guide to FACETS. Rasch-model computer programs* [Program manual 3.80.0]. <http://www.winsteps.com/a/Facets-ManualPDF.zip>
- Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, 6(1), 8. <https://doi.org/10.1186/s40536-018-0061-2>
- Lundetræ, K. (2018, November 1). Jo, norske elever er blitt flinkere [Norwegian students have become more proficient]. *Verdens Gang*.
- McGuire, W. (1985). Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., Vol. 2, pp. 233–346). Random House.
- Meyer, F., Yao, E., & Meissel, K. (2020). The summer learning effects in writing in New Zealand. *Reading and Writing*, 33(1), 1183–1210. <https://doi.org/10.1007/s11145-019-10003-6>
- Monroe, M., & Sherman, E. (1966). *Group diagnostic reading aptitude and achievement tests*. Nevins.
- Nicholson, T., & Tiru, S. (2019). Preventing a summer slide in reading—The effects of summer school. *Australian Journal of Learning Difficulties*, 24(2), 109–130. <https://doi.org/10.1080/19404158.2019.1635499>
- Oberle, E., & Schonert-Reichl, K. A. (2016). Stress contagion in the classroom? The link between classroom teacher burnout and morning cortisol in elementary school students. *Social Science & Medicine*, 159, 30–37. <https://doi.org/10.1016/j.socscimed.2016.04.031>
- Pachter, M., Luttenberger, S., Macher, D., Berding, F., Papousek, I., & Weiss, E. (2015). The effects of 9-week summer vacation: Losses in mathematics and gains in reading. *Eurasia Journal of Mathematics, Science and Technology Education*, 11, 1339–1413.
- Petty, R., & Cacioppo, J. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19(1), 123–205. [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender differences in reading and writing achievement: Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*, 74(4), 445–458. <https://doi.org/10.1037/amp0000356>
- Schleicher, A. (2020). *The impact of COVID-19 on education: Insights from education at a glance*. OECD Publishing. <https://www.oecd.org/education/education-at-a-glance/>
- Skar, G. B., Aasen, A. J., & Jølle, L. (2020). Functional writing in the primary years: Protocol for a mixed-methods writing intervention study. *Nordic Journal of Literacy Research*, 6(1), 201–216. <https://doi.org/10.23865/njlr.v6.2040>
- Skar, G. B., Jølle, L., & Aasen, A. J. (2020). Establishing rating scales to assess writing proficiency development in young learners. *Acta Didactica Norden*, 14(1), 1–30. <https://doi.org/10.5617/adno.7909>
- Skar, G. B., Lei, P., Graham, S., Aasen, A. J., Johansen, M. B., & Kvistad, A. H. (2021). Handwriting fluency and the quality of primary grade students' writing. *Reading and Writing*. Advance online publication. <https://doi.org/10.1007/s11145-021-10185-y>
- Slavin, R., Madden, N., & Karweit, N. (1989). Effective programs for students at risk: Conclusions for practice and policy. In R. Slavin, N. Karweit, & N. Madden (Eds.), *Effective programs for students at risk* (pp. 21–54). Allyn & Bacon.
- Snijders, T. (2005). Power and sample size in multilevel linear models. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. 3, pp. 1570–1573). Wiley. <https://doi.org/10.1002/0470013192.bsa492>
- Snijders, T., & Bosker, R. (1999). *Multilevel Analysis: An introduction to basic and advanced multilevel modeling*. Sage.
- Sprang, G., & Silman, M. (2013). Posttraumatic stress disorder in parents and youth after health-related disasters. *Disaster Medicine and Public Health Preparedness*, 7(1), 105–110. <https://doi.org/10.1017/dmp.2013.22>
- Unesco. (2020). *COVID: 12 education responses*. <https://rb.gy/kuckhk>
- Walberg, H., & Ethington, C. (1991). Correlates of writing performance and interest: A U.S. National assessment study. *The Journal of Educational Research*, 84(4), 198–203. <https://doi.org/10.1080/00220671.1991.10886016>
- Winthrop, R. (2020). *COVID and school closures: What can countries learn from past closures?* Brookings. <https://rb.gy/ctwu4n>

(Appendix follows)

Appendix

Writing Quality Anchor Points for Scores of 1 and 5

Trait	Score of 1	Score of 5
Audience awareness	To understand the text, a conversation with the writer is required.	The text addresses the reader in the assignment in a generally relevant manner and takes into account the reader's need for knowledge of participants/characters, circumstances, and events. The text may contain traces of the student's voice with reflective or evaluating utterances.
Organization	The text consists of individual letters/words/characters/drawings.	The text has a complete global structure with several elaborated elements arranged in a logical or otherwise appropriate order. The text may show examples of comparisons, classifications, chronology. The text contains connectors that are used suitably and purposefully.
Content relevance	The part of the verbal text that is a relevant answer to the task corresponds to a sentence or less.	The part of the verbal text that is a relevant answer to the task corresponds to approximately one and a half A4 pages or more (75+ words).
Vocabulary	The text consists of individual letters/words/characters/drawings.	The text contains a repertoire of words and expressions that are relevant to the task. In some cases, there is use of specialized and abstract words, and/or creative forms of expression.
Sentence construction	The text consists of individual letters/words/characters/drawings.	The text has for the most part appropriate syntactic variation, and it has some developed phrases and/or paragraphs.
Spelling	There may be letters in the text and/or there may be scribbles (imitating writing).	There are a number of examples of nonphonetic words written correctly.
Legibility	Letters may be difficult to decipher (if any).	The letters are crafted in a conventional and legible manner. For the most part, there is conventional use of "bokstavhuset".
Punctuation	The text does not use punctuation.	Usually follows conventions for use of upper- and lower- case Functional use of various forms of punctuations. The use of a comma may occur.

Note. The descriptors for Levels 2–4 are omitted. Please refer to Skar, Jølle, and Aasen (2020) for a full description.

Received March 11, 2021
Revision received June 3, 2021
Accepted June 7, 2021 ■