Primary school mathematics during Covid-19: No evidence of learning gaps in adaptive practicing results

Martijn Meeter,

LEARN! Research Institute,
Vrije Universiteit Amsterdam

Abstract

During the spring of 2020, the Covid-19 pandemic induced many governments to close schools for many months. In the Netherlands, primary schools were closed during a lockdown lasting from March until June. While education was continued online, evidence so far suggests that learning has suffered from the lockdown. Here, we report results from Dutch primary schools that relied on adaptive practicing software for teaching mathematics. The software, installed on tablets that could be taken home by the pupils, continued to be used during school closures. Performance was tracked throughout the lockdown for 53,656 pupils in grades 2 through 6 from 810 schools, and compared to performance of similar pupils in the preceding year. While performance up to the lockdown was similar for school year 2019-20, during the lockdown learning was faster that it had been in 2018-19, contradicting results reported so far. These enhanced gains were correlated with increased use, and remained after the lockdown ended. Pupils ended the year with a higher proficiency score than was reached in earlier years. This was the case for all grades but more so for lower grades, for strong and weak students but more so for weak students, and less so for students in schools with disadvantaged populations than in schools with non-disadvantaged populations.

Introduction

While the Covid-19 pandemic is in first instance a medical emergency, it has had vast consequences in many sectors. For education, lockdowns included the closure of schools in many countries (Hirsch, 2020; Hsiang et al., 2020). While schools have typically moved to forms of distance learning, there is the general suspicion that the school closures have led to decrements in learning. Moreover, these decrements were projected to be unevenly distributed, with students from less privileged backgrounds hit harder than others (Kuhfeld, Soland, et al., 2020).

The first data on the closures in the spring of 2020 have now appeared. Most collecting data on learning in primary school pupils, typically using standardized, computerized tests that are administered digitally to very large samples. These studies have generally found that, on average, pupils had learned less during the semester including the school closures than their peers did in the same period in previous years. A large American study of grade 3-8 students found performance on the MAP Growth assessments at normal levels in October 2020 (i.e., after school closures), but math scores in the same assessments some 5-10 percentile scores lower than in previous years (Kuhfeld, Tarasawa, Johnson, Ruzek, & Lewis, 2020). Decrements were larger for mathematics scores than for English scores, and larger for lower grades (i.e., 3-5) -although also higher grades (6-8) did not perform as well in 2020 than they did in earlier years. Two studies done on Dutch pupils compared standardized tests administered just before a lockdown period (February 2020) to performance after an eight-week school closure period (June 2020). One study concluded that the gap in learning compared to earlier years was of a size (3 percentile points) consistent with that the weeks of online education had been a vacation in which no learning had occurred (Engzell, Frey, & Verhagen, 2020). Another study, using even more data (Lek, Feskens, & Keuning, 2020) showed that decrements were seen for all grades and all levels of prior proficiency, most strongly for reading comprehension but also for mathematics and spelling. A study of Flemish grade-6 exam results showed a drop equal to 0.2 standard deviations compared to previous years in mathematics, and a drop equal to 0.3 standard deviations in language scores (Maldonado & De Witte, 2020).

School closures were also feared to lead to an increase in educational inequality (Kuhfeld, Soland, et al., 2020). Findings in this regard were somewhat mixed, depending on what data was available to correlate with test scores. Engzell et al. (2020), Lek et al (2020), and Kuhfeld, Tarasawa et el. (2020) all noted that decrements in learning were not correlated with previous attainment – in other words, student who had lower scores preceding the lockdown did not suffer more during it. At the school level, Engzell et al. (2020) and Maldonado and De Witte (2020) observed strong variation between schools in how much their students had suffered from the lockdown, which Maldonado and De Witte (2020) could link to the student population: schools with more disadvantaged students in their population were more likely to show decrements in results after the school closures. Engzell et al. (2020) did not have such school-level data, but they could couple results to parental educational attainment. They showed that in their sample, students with parents with less educational attainment had learning decrements that were up to 55% larger than students with highly educated parents. Kuhlfeld, Tarasawa et al. (2020) could not confirm any of these patterns, but they did see worrying signs of mounting inequality in their attrition analyses: Compared to other years, many more students with a disadvantaged background were not tested at all.

All studies discussed above rely on a comparison of results from standardized tests from 2020 with previous years. The use of such data from standardized tests has several advantages, such as that comparability is usually guaranteed by the rigorous procedures used to produced them, it also has disadvantages. First, it lacks all temporal precision, with typically one or measurements a year. This means that any effect from the period of school closures cannot be disentangled from the period when schools had reopened, but education was in all likelihood not equal to how it would have been in a normal year. Second, and related, the critical test in the comparison was administered in June or October 2020, when schools were still heavily affected by the Covid pandemic. In all studies it was noted that not all schools had administered the test, and that nonresponse was not randomly distributed. Moreover, the setting of the test may have been abnormal, schools may have prepared less or differently for it than in other years, and students may have worked differently than they would otherwise have (e.g., because emotional upheaval as a results of the pandemic).

Here, we present results from a different source of data, adaptive practicing software. Such software typically computes proficiency scores on a fine-grained time scale (up to real-time, updated after every exercise), which allows us to plot learning throughout the whole school closure period and the period thereafter. With this data, we aimed to answer two research questions:

- How did learning decrements build up during the school closures, and the period immediately thereafter?
- Is there any evidence of increasing inequality in scores, either as a function of prior learning or of background, during the school closures?

Methods

Context

Snappet is a digital learning environment that is primarily used for teaching and adaptive practice in mathematics, language and spelling within classroom contexts. It is aimed at primary schools, and is used by a sizeable number of schools in both the Netherlands and Spain (only Dutch data was used).

Snappet comes pre-installed on tablets that schools can hand out to their students. Data from each student is then collected on a dashboard for the teacher, who can use real-time data to adjust their instruction or give personal feedback to individual students (Molenaar et al., 2016). Internally, Snappet computes estimates of student achievement using item response theory, both at the level of individual skills and of general mathematical, language and spelling achievement. Item response theory is a framework where the difficulty of each item and the ‘trait score’ of each respondent are estimated on the basis of the responses given (Embretson & Reise, 2000). Within Snappet, the ‘trait score’ is not a fixed trait, but instead achievement as can be derived from performance over a certain period. We used weekly estimates computed each Sunday morning, using data from students of the week before.

Participants

Of the 6333 primary schools in the Netherlands, around a third is currently using Snappet, of these 810 (36%) consented to using their data on a pseudonymous basis for scientific research.

To have some handle on population in the schools, we used the ‘disadvantaged population’ scores computed by the Dutch national bureau of statistics (CBS) for the purpose of disbursing remediation funds to schools with many pupils from disadvantaged households. The score is derived from a regression, predicting standardized test scores from demographic characteristics. Variables that predict lower results are then added to the score of a school, such as pupils stemming from household with low parental educational attainment, low income, migration background, single-parent household, etc. These scores were subdivided into deciles (for 12% of consenting schools, this index was not available). Figure 1 shows the proportion of schools that fall within each decile, both of all Snappet schools and the schools that consented to data use. The distribution of consenting schools over deciles was not different from all Snappet schools ($\chi^2(9)=6.35$, $p=0.70$), but it was different from all Dutch schools ($\chi^2(9)=23.39$, $p=0.005$) due to a slight overrepresentation of schools with a more disadvantaged population.

Within the participating schools, 100,471 students used Snappet sufficiently to have achievement estimates for the periods under study. Due to schools using Snappet in some grades but not others, or starting and stopping to use it for different cohorts of students, there was sizeable variation in student numbers per grade (see Table 1). Few schools used Snappet in first grade – this grade was therefore dropped in the analyses. There was some increase in student numbers from school year 2018-19 to 2019-20, probably resulting from the fact that some consenting schools started using Snappet only in 2019-20 (while schools that had used it in 2018-19 but not in 2019-20 were not asked for their consent in using the data).

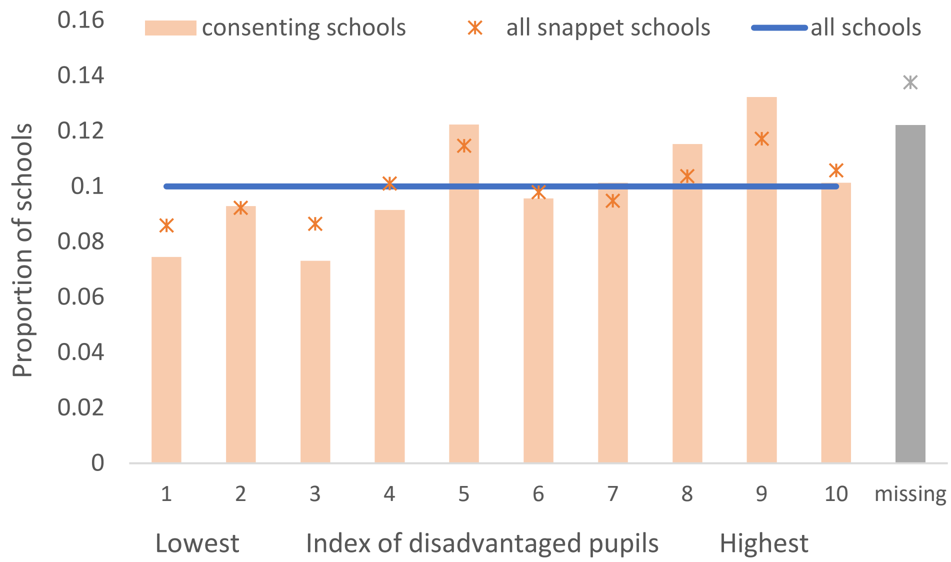

Figure 1: distribution of schools consenting to data use as a function of the index of disadvantaged pupils in their population. All Dutch primary schools were subdivided into deciles on the basis of this index (with the first decile having very few disadvantaged students, and the 10th many). Missing refers to the proportion of schools for which this index was missing – all other proportions were computed with reference to the number of schools where the index was not missing.

Table 1: included students using Snappet in each grade, in school years 2018-19 and 2019-2020, and the same periods in school year 2018-19. Data from grade 1 students (193 and 253 in the two years) was not used.

grade	2018-19	2019-20
2	6,845	7,755
3	9,633	11,975
4	12,722	12,614
5	6,338	14,979
6	11,277	6,333
Total	**46,815**	**53,656**

Data analysis

Data analysis was run on Snappet servers with a script provided by the researchers. These did not have access to raw data, out of privacy concerns. Only data was used from the schools that consented to pseudonymous use for research. Only data was used from students who had at least 16 weekly estimates (i.e., who had used Snappet for at least 16 weeks, or 40% of the year), and who were in classes with at least 10 users (to exclude classes in which Snappet was used either only for remediation or as optional material).

The weekly estimates of mathematics achievement were divided, for the school years 2018-19 (control year) and 2019-20 (year of Covid-19), into three periods: the pre-lockdown period (i.e., from the start of the school year up to March 16), lockdown period (March 14 – May 11th), and post-lockdown period (from May 11th up to the end of the school year). Weekly estimates of each student were fit with linear regression for each period separately. Users with only one weekly estimate in one of the periods were excluded for that period, since no regression line could be computed for these users. The regression analyses led to estimates for starting level (intercept parameter), weekly learning (slope parameter), and attainment at the closing of the period (endpoint of the regression line) for each individual student. These estimates were the basis of our analyses. We also analyzed the number of coherent sets each student played within a week (where each coherent set of exercises would take a few minutes to complete).

In our main analysis, we then compared the estimates of 2019-20 with those from 2018-19, for each of the three periods. No markers necessary for hierarchical modeling were available. We therefore ran traditional ANOVAs on the comparisons, but used a stricter alpha of 0.005 to compensate for not being able to model the hierarchical structure of the data. Moreover, to study differential effects as a function of previous attainment, we split students into three equal bins on the basis of their results in the first half of the year. This was done on a per-class basis, to ensure that any effect of bin was not confounded with school- or class-level differences (binning the whole cohort in one go yielded only numerically different results). School bin, as defined by the 'disadvantaged population' scores, was also used in the analyses. To pinpoint results to specific bins or grades, we computed Cohen's d as the difference between 2018-19 and 2019-20 for a certain variable, for each bin or grade. 99% confidence intervals were computed around the effect size for statistical evaluation.

Results

Figure 2 shows average estimated attainment as a function of school week and grade, separately for the year of Covid-19 and the control year 2018-19. For most grades there was a slight advantage of 2018-19 students over their 2019-20 peers before the lockdown. However, once the lockdown started, the two years started to diverge. Learning was stronger for the lockdown year than the year before, and this effect remained when the lockdown ended, although the lines of the two years converged towards the end of the year.

Figure 3 shows the same data but recomputed as the difference between 2018-19 and 2019-20 expressed as an effect size (Cohen's d). Panels a and c show average learning and attainment as a function of student bin and period. Pre-lockdown, there was slightly stronger learning in 2019 than 2018 (positive effect size for all three bins), in which 2019-20 students were making up for a lower level of attainment at the start of the year (negative effect size for pre-lockdown period in panel c, of attainment). During the lockdown learning was much stronger in 2019-20 (positive effect sizes in panel a), leading to higher attainment at the end of the period (positive effect sizes in panel c). Higher attainment was still visible at the end of the post-lockdown period, even though weaker learning in this period (negative effect sizes for the post-lockdown period in panel a) mitigated some of it.

These patterns were confirmed with a set of ANOVAs. For the pre-lockdown period, there was a main effect of year on learning, in favor of 2019-20, F(1, 100456)=94,05, p<.001. There was also a main effect of bin, reflecting somewhat stronger learning for the lowest bin, F(2, 100456)=91.13, p<.001, but no interaction between the two factors, F(2, 100456)=2.09, p=.13. For the lockdown period, there was a main effect of year in favor of 2019-20, F(1, 100456)=6844, p<.001, of bin, again in favor of the lower bin, F(2, 75459)=1052, which was qualified by an interaction between year and bin, F(2, 100456)=368, p<.001): stronger learning during the lockdown was particularly pronounced for the lowest bin, with the higher bins benefiting progressively less (see Figure 3, panel a). In the post-lockdown period, effects were reversed. The main effect of year, F(1, 100456)=1128, p<.001, now favored 2018-19, while the interaction between bin and year, F(2, 100456)=53.19, p<.001, showed that the higher bin progressed more strongly after the 2020 lockdown that the lower bins. Only the main effect of bin, F(2, 100456)=56.54, p<.001, continued to favor the lower bin.

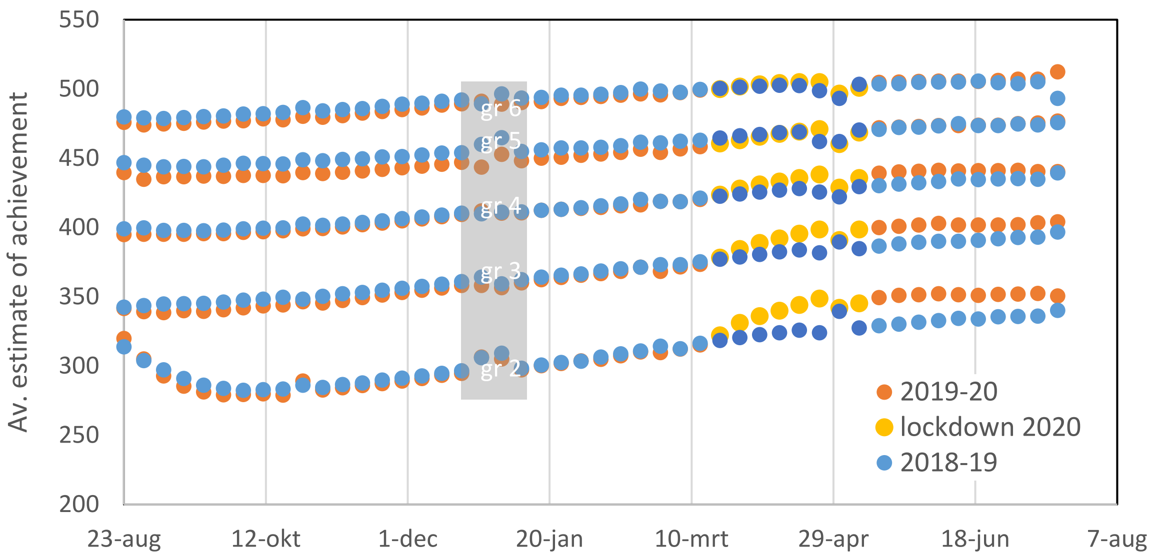

Figure 2: weekly estimates of achievement for 2018-19 and 2019-20, separately for grades 2 to 6. The lockdown period is marked in yellow. Most students start using the software in grade 2, and it takes one to two weeks for the estimates to go from their default starting level to the true level of the student. Since the start of the school year is staggered in the Netherlands, this effect is smeared out over time. Weeks with sudden jumps up or down (e.g., the weeks underneath the gray box, the first May week and the last week) are vacation weeks in which only few students use the software.

Moving to attainment, an ANOVA for the pre-lockdown period showed a main effect of year, F(1, 100456)=34.38, p<.001, showed somewhat better attainment in 2018-19 than in 2019-20, with a main effect of bin, F(1, 100456)=21,151, p<.001 (obvious since bins were defined on the basis of attainment), and no interaction between these two factors, F(1, 100456)<1. After the lockdown, the main effect of year had reversed, F(1, 100456)= 102, p<.001 with better attainment, now, for 2019-20. The main effect of bin was still there, F(1, 100456)= 19,158, p<.001, and now there was an interaction between bin and year, F(1, 100456)= 23.79, p<.001: attainment was especially higher in 2019-20 for the students in the lowest bin (see Figure 3, panel c). Due to missing information no ANOVA of attainment at the end of the post-lockdown period was possible, but Figure 3, panel c shows that for both the lowest and the middle student bin, the confidence intervals around the effect size show that attainment was still higher for these groups than for their peers in the preceding year. This was not the case, however, for the highest bin, where attainment at the end of the year was not different from that of peers in the preceding year.

Figure 3, panel b shows the effect on learning during lockdown and post-lockdown split out per grade. Stronger learning during lockdown was especially prominent for the lower grades, and less so for the higher grades (this can also be seen in Figure 2). However, the slowing of learning post-lockdown was also especially prominent for those grades. The suggestion of a tradeoff between strong learning during the lockdown and weak learning after it was further strengthened by the small, negative correlation between learning estimates during and after the lockdown (r=-0.25, p<.001). No such negative correlation was present between learning estimates for the pre-lockdown and lockdown periods (r=.06, p<.001). We come back to this in the discussion.

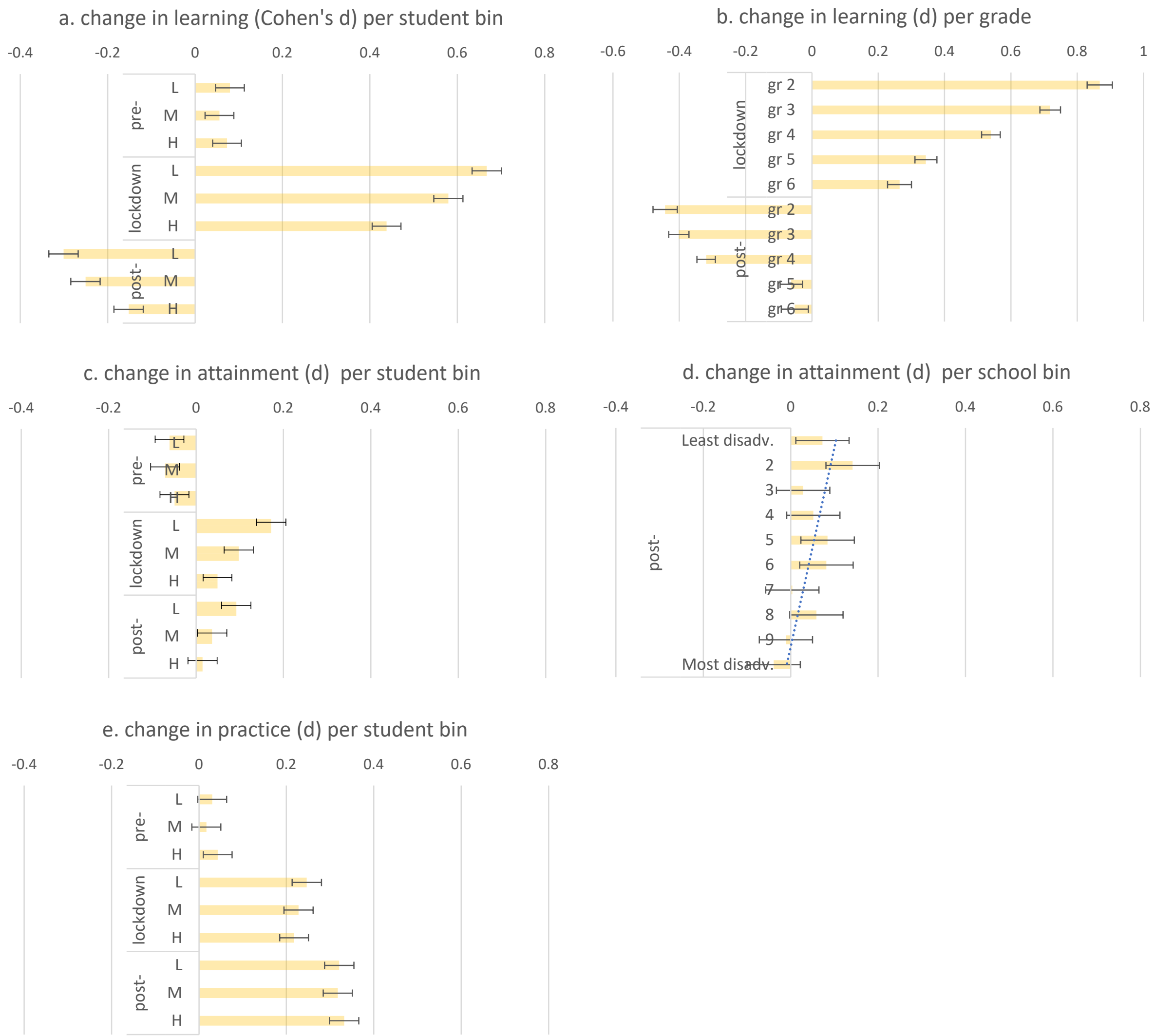

Figure 3: Comparison of the Covid-19 year with the preceding control year (2018-19). Each bar shows an effect size (Cohen's d) computed from the comparison. The error bars denote the 99% confidence interval around the estimate. (a). Change in average weekly learning for the pre-, lockdown, and post-lockdown periods, split up for the lowest- (L), middle- (M) and highest-achieving students, as defined by their average performance in the pre-lockdown period. (b) Same as preceding panel, but now split up per grade. (c) Change in estimated attainment at the end of each period for each of the three student bins. (d) Same, but only for the post-lockdown period (i.e., end of the year) and binning based on school disadvantaged student population scores. In blue the line of best fit through the ten estimates. (e). Change in practice, as measured by the number of exercises finished in a week by students, again separately for each period and student bin.

Figure 3, *panel d* shows attainment at the end of the school year as a function of school bin, where schools were binned on the basis of having a disadvantaged student population. Although there is some variation, stronger learning in the Covid-19 year was especially evident for the school bins with less disadvantaged populations. The two bins with the most disadvantaged populations were also two of the ones where the mean effect was not significantly different from 0 (as derived from the confidence intervals shown in the figure).

One possible explanation for the stronger progress during the lockdown in 2019-20 is simply more practicing. Figure 3, *panel e* shows that while practice was about equal for the two years in the pre-lockdown period, students finished more exercises on Snappet during and after the lockdown in 2019-20 than their peers had done in 2018-19. However, the different was particularly stark in the post-lockdown as opposed to the lockdown period. In the lockdown period, practicing was some 15% higher than it had been in the same period in 2018-19. In the post-lockdown period the difference was 30%, mostly because usage dropped in that period in 2018-19 but not in 2019-20. Increased use correlated with stronger learning, though weakly so ($r=.15$, $p<.001$). This correlation was about the same strength for the pre-lockdown period ($r=.145$, $p<.001$ between pre-lockdown practice and learning), but was absent for the post-lockdown period ($r=-.001$, $p=.07$).

Discussion

Several studies have shown, also for Dutch primary education (Engzell et al., 2020; Lek et al., 2020), that, after the school closures, students had fallen behind compared to previous years. Here, using data from adaptive practice software, we found no evidence of any learning decrements. During the period of school closures, students progressed more strongly than their peers in previous years had. This was especially true for younger students (grades 4 and 5), students who had weaker results before the lockdown, and students in schools with less disadvantaged student populations. These gains diminished in the weeks after the closures, when education returned to normal. Why the advantage for 2020 students dissipated during the return to normal education is unclear. Education was messy in these weeks, with students first being taught in half classes for half of the week. This may have led to lower gains. Perhaps schools made a conscious choice to concentrate on student welfare or neglected skills.

Nevertheless, for lower grades and weaker students, some gains were still present at the end of the school year. There are several possible explanations for the contrast between these results and those of other studies:

Altered testing circumstances: it is possible that the decrements in standardized test scores observed in other studies do not truly reflect learning deficits. Instead, they may reflect altered testing circumstances, such less preparation for the tests than in other years. The test analyzed by Engzell et al. (2020) and Lek et al. (2020) is formative, which may mean that schools may have not put too much emphasis on optimal preparation, optimal administration or optimal concentration by their students. However, this would not explain why the results on a high stakes test from a neighboring country would yield very similar results (Maldonado & De Witte, 2020).

Concentrating on the core: it may have been that the school closures induced teachers to focus on core skills that, they felt, had to be covered or could be covered at a distance better than other topics. This might have led to strong learning on those skills (covered by Snappet) at the detriment of other skills that are not emphasized by Snappet but were part of the curriculum. Indeed, practice was more intense during the 2020 school closures than in the same period a year earlier (although the difference was not large, and did not align well with where the gains were seen). While this may explain the surprising finding of stronger learning in the lockdown than in the preceding year, it does not provide a good explanation for the mismatch between current findings and those of Engzell et al. (2020) and Lek et al. (2020), since Snappet covers the same topics as standardized mathematics tests.

Reading comprehension: One difference that does exist between Snappet exercises and those in standardized tests is that those in standardized tests tend to include many word problems, for which good reading skills are needed (Korpershoek, Kuyper, & van der Werf, 2015). Both Engzell et al. (2020) and Lek et al. (2020) showed stronger decrements in reading comprehension than in mathematics after the school closures (this was also found in Flanders but not the US). Perhaps those decrements in mathematics scores were at least partly a result of deficits in reading comprehension sustained by the school closures, and less of deficits in mathematical skills.

More effective teaching. The difference between our current results and those of Engzell et al. (2020) and Lek et al. (2020) may also reflect a true difference between schools relying on Snappet for mathematics teaching, and other schools. The effectiveness of Snappet at boosting mathematics scores was shown in two large quasi-experiments (Faber, Luyten, & Visscher, 2017; Molenaar, van Campen, & van Gorp, 2016). Faber et al. (2017) quantified it as 1.5 month of additional gain for primary school students using Snappet for four months. Schools in the current sample all relied at least partly on Snappet for mathematics teaching, while Little is known about how schools in the sample of Engzell et al. (2020) taught mathematics – presumably through of mix of methods that may not all have been as effective.

More effective distance education. Benefits of Snappet and learning environments like it may be accentuated by school closures. The breakdown of routine and the lack of social contacts during school closures may lead to demotivation and problems keeping oneself at work, even in college students (Meeter et al., 2020). Our current results show that students using Snappet increased the time spent practicing with it. It is unclear whether other forms of mathematics distance education lead to the same amount of practice, but one study suggests that they are not always very effective (Woodworth et al., 2015). This may have been because of the rewarding nature of practice and seeing oneself improve, or the fact that education at home could rely on routines with Snappet tablets established in school.

Inequality

With regard to educational inequality, the current replicate those of Engzell et al. (2020), Lek et al. (2020) and Kuhfeld et al. (2020) in that students with lower prior results did not suffer more from the lockdown than other students. To the contrary, weak learners seemed to catch up to their more advanced peers during the school closures, to a larger extent than did peers in the preceding year.

Part of this catching up was reversed as students returned to schools. This may be explained by findings that when Snappet it used in the classroom, more proficient students tend to benefit relative to students in classrooms where no adaptive practicing software was used. Those at the highest level of proficiency barely made any progress during the five-month study in normal classrooms, but this relative lack of progress was much less pronounced in classes using adaptive practice, purportedly because adaptive practice allowed proficient students to practice at their own level (Faber et al., 2017; Molenaar et al., 2016). However, what then explains the finding of strong progress for weaker students during the school closures? This remains to be explored, but a reason may be that adaptive practice software mostly incorporates mastery learning (Ritter, Yudelson, Fancsali, & Berman, 2016) , which may be more effective if unmoored from classroom routines.

While the results thus contain some good news, there are also indications for increased inequality as a function of student background. Replicating Maldonado and De Witte (2020), we found that schools with more disadvantaged student populations were characterized by less strong learning gains than schools with more advantaged populations. This adds to strong evidence that unequal outcomes are more a function of student background than of prior attainment (Engzell et al., 2020). In a study performed in Great Britain, it was found that students from more disadvantaged backgrounds spent fewer hours learning at home during school closures than did their more fortunate peers (Pensiero, Kelly, & Bokhove, 2020), while a Dutch survey of parents found that parents with more education were more likely to state that they could help their children with distance education than parents with less education, even though both valued it equally (Bol, 2020). Meanwhile, in the United States, it was found that schools with poorer student populations were more likely to close than schools with less poor student populations (Parolin & Lee, 2020).

Limitations

The study has several strengths, such as that it shows the time course of learning during the school closures, but also several limitations. The analyses were concentrated on mathematics skills, while other studies (e.g., Maldonado & De Witte, 2020; Lek et al., 2020) found larger decrements in reading comprehension than in mathematics in primary school students. Moreover, privacy concerns led to data being aggregated in a way that precluded multilevel analysis. This may have increased the likelihood of spurious positive findings. To compensate for this, a much smaller alpha was used, but given the sample size all effects easily exceed this higher threshold. However, multilevel analyses might still give a better representation of effects, for example by allowing to split variance at the school, class and individual level.

Conclusion

Here, learning in mathematics was analyzed during the Covid-induced school closures of spring 2020 using adaptive practice software. To our surprise, stronger learning was found during the school closures than in the year before. These gains were stronger for lower grades, and for students with weaker previous learning. However, while students in schools with more disadvantaged populations benefited, they benefited less. The study thus adds to those to suggest increased educational inequality as a result of the Covid pandemic. However, more positively, it suggests that adaptive practice software may be a way to attenuate learning losses due to school closures, or even reverse them.

References

Bol, T. (2020). *Inequality in homeschooling during the Corona crisis in the Netherlands. First results from the LISS Panel*. SocArxiv, https://doi.org/10.31235/osf.io/hf32q.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, N.J.: L. Erlbaum Associates.

Engzell, P., Frey, A., & Verhagen, M. D. (2020). *Learning inequality during the COVID-19 pandemic*. (https://osf.io/download/5f995b4687b7df03233b06fe/). OSF preprint.

Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers & Education, 106*, 83-96.

Hirsch, C. (2020). Europ's coronavirus lockdown compared *Politico Europe, 31-3-2020*.

Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., . . . Wu, T. (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature, 584*, 262-267.

Korpershoek, H., Kuyper, H., & van der Werf, G. (2015). The Relation between Students' Math and Reading Ability and Their Mathematics, Physics, and Chemistry Examination Grades in Secondary Education. *International Journal of Science and Mathematics Education, 13*, 1013-1037.

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the Potential Impact of COVID-19 School Closures on Academic Achievement. *Educational Researcher, 49*, 549-565.

Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020). Learning during COVID-19: Initial findings on students' reading and math achievement and growth *NWEA Research Brief* (Vol. https://www.nwea.org/research/publication/learning-during-covid-19-initial-findings-on-students-reading-and-math-achievement-and-growth). Portland, OR: Northwest Evaluation Association.

Lek, K., Feskens, R., & Keuning, J. (2020). Het effect van afstandsonderwijs op leerresultaten in het PO [The effect of distance education on learning results in primary education] (Vol. https://www.cito.nl/kennis-en-innovatie/kennisbank/p348-onderzoeksrapport-het-effect-van-afstandsonderwijs-op-leerresultaten). Enschede, the Netherlands: Cito.

Maldonado, J., & De Witte, K. (2020). The effect of school closures on standardised student test. *FEB Research Report* (Vol. https://lirias.kuleuven.be/retrieve/589583). Leuven (Belgium): KU Leuven Department of Economics.

Meeter, M., Bele, T., den Hartogh, C., Bakker, T., de Vries, R. E., & Plak, S. (2020). *College students' motivation and study results after COVID-19 stay-at-home orders.* PsyArxiv, https://doi.org/10.31234/osf.io/kn6v9.

Molenaar, I., van Campen, C. A. N., & van Gorp, K. (2016). Onderzoek naar Snappet; gebruik en effectiviteit. Zoetermeer (Netherlands): Kennisnet.

Parolin, Z., & Lee, E. (2020). *Large Socio-Economic, Geographic, and Demographic Disparities Exist in Exposure to School Closures and Distance Learning*. OSF Preprints, https://doi.org/10.31219/osf.io/cr6gq.

Pensiero, N., Kelly, A., & Bokhove, C. (2020). Learning inequalities during the Covid-19 pandemic: how families cope with home-schooling Southampton, UK: University of Southampton.

Ritter, S., Yudelson, M., Fancsali, S. E., & Berman, S. R. (2016). *How mastery learning works at scale.* Paper presented at the Proceedings of 3rd ACM Conference on Learning@ Scale, https://doi.org/10.1145/2876034.2876039.

Woodworth, J. L., Raymond, M. E., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W., & Van Donge, Y. (2015). Online Charter School Study. Stanford, CA: Center for Research on Educational Outcomes, https://credo.stanford.edu/sites/g/files/sbiybj6481/f/online_charter_study_final.pdf.