

Remote elementary education (Part 2): grade-level comparative analysis with control

Remote
elementary
education 2

259

Stefan Kleinke and David Cross
*COA Graduate Studies,
Embry-Riddle Aeronautical University Worldwide and Online,
Daytona Beach, Florida, USA*

Received 6 March 2022
Revised 25 May 2022
Accepted 7 June 2022

Abstract

Purpose – The purpose of this two-part research was to investigate the effect of remote learning on student progress in elementary education. Part 2, presented in this paper, is a follow-up study to examine how student progression in the two pandemic-induced environments compared to the pre-pandemic conditions.

Design/methodology/approach – The authors expanded the quantitative, quasi-experimental factorial design of the authors' initial study with additional ex-post-facto standardized test score data from before the pandemic to enhance the group comparison with a control: the conventional pre-pandemic classroom environment. Thus, the authors were able to examine in which ways the two pandemic-induced learning environments (remote and hybrid) may have affected learner progress in the two subject areas: English Language (ELA) and Math. Additionally, the authors provided a grade-by-grade breakdown of analysis results.

Findings – Findings revealed significant group differences in grade levels at or below 6th grade. In the majority of analyzed comparisons, learner achievement in the hybrid group was significantly lower than those in either the remote or the classroom group, or both.

Research limitations/implications – The additional findings further supported the authors' initial hypotheses: Differences in the consistency and continuity of educational approaches, as well as potential differences in learner predispositions and the availability of home support systems may have influenced observed results. Thus, this research also contributes to the general knowledge about learner needs in elementary education.

Originality/value – During the pandemic, remote learning became ubiquitous. However, in contrast to e-learning in postsecondary education, for which an abundance of research has been conducted, relatively little is known about the efficacy of such approaches in elementary education.

Keywords Elementary education, Learning environment, Remote learning, Learning assessment, Control group design

Paper type Research paper

In previous research (Kleinke and Cross, 2021), we investigated the effect of remote learning on student progression in two subject areas crucial to elementary education: English Language (ELA) and Math. Brought on by the circumstances of the global pandemic, many schools across various nations were forced to suspend in-person, in-classroom learning and adopt some form of remote education (Storey and Slavin, 2020). Once gradually returning to in-person learning, educators were faced with difficult decisions about how to best balance the constantly changing public health guidelines, parent expectations and student needs, which often resulted in schools offering learning simultaneously in different modalities (Kleinke and Cross, 2021).

© Stefan Kleinke and David Cross. Published in *Journal of Research in Innovative Teaching & Learning*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>



Therefore, our previous quasi-experimental research (Kleinke and Cross, 2021) focused on the analysis of group differences in standardized test scores between fully remote and hybrid learners. We found a significant difference in learning progression between the two groups for both of the investigated subject areas. Furthermore, the findings suggested that fully remote students outperformed their hybrid-learning peers, which we tentatively attributed to the greater consistency in the fully remote environment. While students in the fully-remote environment continued their learning through pandemic lock-down and gradual re-opening unchanged, students in the hybrid environment experienced frequent changes to learning conditions and school policies, which may have negatively affected their learning achievements (Kleinke and Cross, 2021).

Thus, the question that arose from our initial research (Kleinke and Cross, 2021) was whether the observed group differences were indeed caused by negative effects in the hybrid environment and not attributable to performance increases in the fully remote environment. Accordingly, the purpose of this follow-on study was to further examine how student progression in the two pandemic-induced environments compared to the pre-pandemic conditions. Therefore, we expanded the quantitative, quasi-experimental factorial design of our initial study with additional ex-post-facto test score data from before the pandemic to create a third condition for comparison: the conventional classroom environment. Based on our tentative theory from our initial findings (Kleinke and Cross, 2021), we hypothesized:

H1_o. There are no significant differences in student development among the three different learning environments.

H1_a. Student development in the hybrid learning environment (*H*) is significantly less than under both fully remote (*R*) and conventional classroom (*C*) conditions.

Furthermore, the revised research design also allowed us to investigate whether any observed group differences would be consistent across all grade levels, which would suggest that learners are equally affected by the conditions regardless of age. We, therefore, also hypothesized that:

H2_o. The distribution of any group differences is consistent across all grade levels.

H2_a. Group differences vary across grade levels with younger learners more affected by learning conditions.

Literature review

Consistency and continuity in teaching and learning have long been identified as important to student success. Founded on Dewey's (1938) principle of continuity of experiences, learning environments, pedagogical approaches and academic curricular traditionally sought to provide a consistent learning experience in order to maximize student achievement (Cantley et al., 2021). While his philosophy also emphasized the importance of personal experiences, Dewey (1938) stressed that individual interactions are not enough if not carefully choreographed into a consistent learning experience by a skilled educator. Nevertheless, research conducted at the beginning of the new millennium (e.g. Driessen and Slegers, 2000; Jamieson and Wikeley, 2000) started to challenge this conventional view of the role of consistency as essential to effective learning, maintaining that adaptation to individual needs may be more crucial. For example, Driessen and Slegers (2000), in their study of 7,410 8th grade students from 567 schools in the Netherlands found that consistency of the teaching approach was of no relevance to the measured levels of achievement, and socio-ethnic background and the prevalence of disadvantaged students were much more important predictors for performance.

At the same time, the home learning environment (HLE) has also been increasingly identified by researchers as an important factor for elementary children's academic and social development (Lehrl *et al.*, 2020a, b). For example, through multilevel stepwise regression analysis, Dimosthenous *et al.* (2020) found that, in addition to teacher factors, certain HLE aspects such as home-learning activities and home-learning materials had a significant short-term, as well as some long-term, effect on student achievements. They also concluded that some of the potential deficits in a student's HLE may be offset by teacher factors (i.e. that a good teacher can make up for some of the negative effects of a less than ideal learning environment at home; Dimosthenous *et al.*, 2020). Furthermore, these studies seem to point to the need for increased collaboration between a child's home and their school to positively affect learning outcomes (Lehrl *et al.*, 2020a, b), seemingly supporting both concepts: individualized attention as well as consistent delivery.

When applied to remote learning, educators must adopt effective methods of instruction and reach consistency in their teachings. Consistency may help teachers as well as students to feel more organized and less stressed (O'Leary *et al.*, 2021). Consistency may also help students to feel more relaxed and comfortable in the classroom, which might increase their participation and engagement. The more consistent a teacher can be, the more trust and respect students and their parents may have for that teacher and expectations can become mutual. Thus, consistency seems especially important during the early years of education when learning effectiveness depends on the above-discussed home-to-school connections (Lehrl *et al.*, 2020a, b). Van Ameijde *et al.* (2018) reported that consistency of teaching methods and styles during remote learning seems to result in more effective mastery of math and language skills in the majority of students in primary school; therefore, particularly in an elementary school remote learning environment, such as during the pandemic, teachers need to adopt effective approaches that help them to reach consistency (Gunzenhauser *et al.*, 2021).

To reach consistency in remote learning, it is important to curate resources in one centralized location or database. The more time students and parents spend looking for assignments and resources, the less likely it is that learning will occur. In a sense, in a remote learning environment, the previously mentioned HLE aspects of home-learning activities and home-learning materials (Dimosthenous *et al.*, 2020) seem to become a shared responsibility between the remote facilitators and the students' home support system. In this regard, school leaders also play an important role in the process, as they have to encourage their teachers to curate the content they already use in a way that is easy for parents and students to reference (Murai and Muramatsu, 2020). Singh *et al.* (2021) investigated the preparedness of faculty with little or no experience teaching online courses during the COVID-19 pandemic. Using a fishbone analysis, the authors were able to identify root problems which were then used to create a strength, weaknesses, opportunities and challenges (SWOT) matrix for their online teaching experiences. Singh *et al.* (2021) concluded that as online and hybrid teaching will increase in the future, not only is better teacher preparation needed, but support from other stakeholders, including administrators, parents and students, seems crucial to provide meaningful learning experiences. Together, all stakeholders should aim to reach consistency and continuity in remote elementary learning.

Especially in times of a pandemic, when remote learning becomes a good, if not essential, alternative to the conventional classroom setting (Ferri *et al.*, 2020; Storey and Slavin, 2020), one size may not fit all when it comes to instruction, and teachers should have the flexibility to make in-the-moment decisions to personalize learning. However, as Schafer *et al.* (2021) point out, in times like these, it is often more important to be consistent than it is to be creative. Teachers involved in remote learning should establish daily parameters for learning. Both students and parents are still adapting to working and learning from home. To aid in that transition, educators can set daily schedules and clear learning objectives to easily identify what students should be focusing on each day. As educators navigate these changes in

instructional practices, it is critically important for both administrators and teachers to present consistent information. If leaders do not have a clear vision for instruction, teachers may be left to their own devices, which could cause students and parents to try keeping track of multiple different platforms for daily learning. Dhawan (2020) researched educational institutions in India and their response to the shift away from traditional teaching methods during the COVID-19 pandemic. Conducting a SWOT analysis, Dhawan (2020) concluded that a strong IT infrastructure and e-learning platform are critical to successful implementation of online learning. Once the infrastructure and platform are established, people seem more accepting of the online teaching environment.

Daily teaching routines should also look consistent, especially for students at the primary grade levels who may not be using digital resources on a daily basis (Schafer *et al.*, 2021). Students report that communication and materials delivery become an issue during a remote study if teachers change their teaching styles. Thus, consistency in teaching methods further seems important because it helps to compensate for a lack of personal connection. Especially during a pandemic, each community may face different challenges, making it difficult to remain consistent across the board (Fontenelle-Tereshchuk, 2021). However, consistency may help to ensure equity and access for all students regardless of grade level or socioeconomic status. The application of remote learning strategies offers opportunities for students and educators to learn without the direct, face-to-face interaction (Armfield *et al.*, 2021). Nevertheless, in such a situation, educators need to have effective strategies, and providing consistency and continuity in the learning environment seems to be one important aspect.

Research method

For this follow-on research, the initial group comparison (Kleinke and Cross, 2021) between elementary students that spent the entire pandemic year under remote learning conditions (remote learning group, coded as *R* for the grouping variable) and those who returned to some form of in-person learning, albeit with restrictions such as mask mandates, at some point during the assessment cycle (hybrid learning group, coded *H*) was expanded by introducing a third learning condition: the conventional pre-pandemic classroom (classroom learning group, coded as *C* for the grouping variable and serving as the control). To measure student achievements in the three groups, pre-existing assessment data from the three consecutive years 2019, 2020, and 2021 were utilized, which allowed establishing within-subject paired pretest-posttest samples for each group. Thereby, pre-test scores (at the beginning of each assessment period) served as covariate moderating any observed differences in the posttest assessments (at the end of each assessment period). The 2019 and 2020 test results (both assessments from before the pandemic) served respectively as pretest and posttest variables for the *C* group, while 2020 and 2021 test results (one assessment from just before pandemic lockdowns and one from after gradual reopening) provided pretest and posttest scores for the two treatment groups *R* and *H*. In this way, the assessment data were collected and tracked for both subject areas Math and ELA.

By further splitting the data into grade levels, a comparison of *R* and *H* to a control *C* became possible: For example, the pre-pandemic 2019/2020 assessment data for all students that were in the 2020/2021 academic year in Grade 2 became the conventional classroom control *C* against which the 2020/2021 Grade 1 students in the *R* and *H* groups were compared. Thus, the analysis design accounted for grade-level-specific achievement advances in an academic year. Similarly, the within-subject paired pretest-posttest design in each group took individual pre-treatment differences among students and across conditions into consideration, allowing to assess relative learner development rather than just their absolute achievements. As previously discussed in our initial research (Kleinke and Cross, 2021), such design was important for ruling out any influences of (self-) selection biases since group assignments were essentially non-random.

As for the first part of our study (Kleinke and Cross, 2021), the setting for this research was a medium-sized charter school in north-central Colorado, which offers K through 12 education and regularly conducts standardized assessments of student performance utilizing Curriculum Associates' (2021) i-Ready test. This i-Ready assessment is a grade-level-based, norm-referenced, adaptive test for students' ELA and Math abilities that is widely used in elementary schools across the US (Bunch, 2017; Curriculum Associates, 2021; Ezzelle, 2017). Its recurrent, computer-based application was beneficial to the study since it ensured consistent testing conditions across all three learning environments as well as regular assessment data (three times per academic year: September, January and May) over multiple years. Another advantage of using i-Ready as the research instrument was that the January 2020 test administration happened just immediately before schools were forced to suspend in-person learning, making it a valuable pre-test tool for this research (Kleinke and Cross, 2021).

The sample for this study consisted of the assessment data for $N = 904$ students from grade levels K through 8. While our initial study (Kleinke and Cross, 2021) focused on analyzing group differences between R and H students across the entire data set, for this follow-on research we further broke down the assessment data into grade levels, which also allowed us to add pre-pandemic conventional classroom comparison data. Nevertheless, due to the varying availability of these assessment scores in the different grade levels, individual sample sizes varied widely during the analysis. For example, while the C group for 3rd grade Math scores contained $N_{C,3,Math} = 81$ paired pretest-posttest samples, the 8th grade R group in ELA had only $N_{R,8,ELA} = 6$ samples available. Thus, differences in group sizes as well as relatively small sample sizes for 7th and 8th grade were some of the limitations to the study that potentially affected the confidence in results in some of the groups and comparisons.

As outlined in our initial research methodology (Kleinke and Cross, 2021), all statistical analysis was performed using the jamovi software tools (The jamovi project, 2021), including applicable plug-in packages and visualizations (Ben-Shachar *et al.*, 2020; Fox and Weisberg, 2020; Gallucci, 2019; Lenth, 2020; R core Team, 2021; Singmann, 2018) and the preparation of the data included thorough examinations of descriptive statistics to identify outliers and test for analyses assumptions. A particular emphasis during the preparation and treatment of the data was put on identifying and eliminating any potential effects of differences in testing. An analysis of covariance (ANCOVA) on the posttest i-Ready scores (2021 scores for groups R and H and 2020 scores for group C) was conducted for each grade level (K through 8), utilizing the three-level (R , H , or C) grouping variable as factor and the paired pretest scores (Jan 2020 scores for groups R and H and Jan 2019 scores for the C group) as the covariate. For those grade levels for which pre-test data were not available (see discussion of limitations), the treatment was reduced to an analysis of variance (ANOVA) among the three groups. Similarly, for the ELA subject area, for which control group pretest data were not available (see discussion of limitations), both an ANCOVA between only the two groups H and R as well as a posttest-only ANOVA between all three groups was conducted for each grade level. In addition to the analyses for each grade level and subject, Math scores across all grade levels combined were also analyzed for their development over time in the H and R groups via a repeated measure ANOVA of 2019, 2020 and 2021 scores. This time series, while less specific than individual grade level analyses, served to gain an overall picture of how students' achievements in the different learning environments during the pandemic compared to, and progressed from, the year preceding it.

Assumptions, limitations and delimitations

As previously highlighted, one limitation of the methodology we used to follow up our initial analysis (Kleinke and Cross, 2021) was that the grade-level breakdown resulted in uneven

group sizes as well as small samples sizes in the upper grade levels. Another limitation to our design was that no 2019 ELA data existed since the school did not use i-Ready for assessments of ELA development during that year. Therefore, ELA classroom control groups (C) across all grade levels lacked pretest covariate data, making the assessment of student development in Math the more robust analysis of the two. Furthermore, obviously, there were also no pretest data available for the K grade level since these students just started their education during the pandemic, and no ELA testing was conducted at the K grade level during any year. Therefore, analysis of K grade level data was mainly a measure of achievement rather than student development and was limited to Math only. Additionally, this lack of K-grade ELA data also, obviously, affected availability of 1st-grade pretest data in ELA, highlighting again that ELA was the less robust of the two measures in our analysis (Kleinke and Cross, 2021).

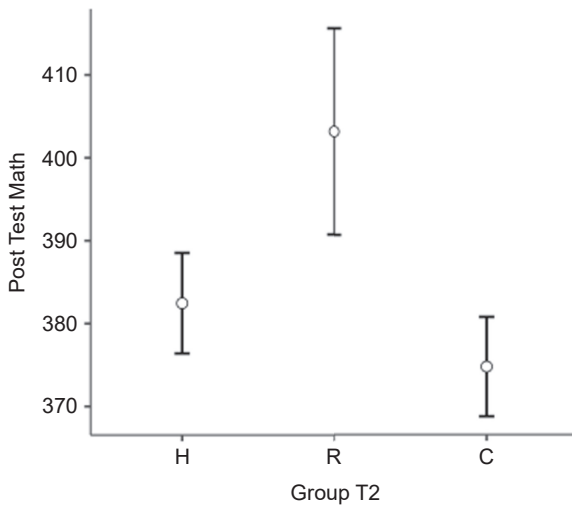
As also already delimited in our initial study (Kleinke and Cross, 2021), the purpose of this research was not to identify any individual differences in students' achievement scores or determine their potential causes (e.g. cheating). Rather, the selected research design and analysis methods were intended to eliminate or control, as much as possible, for any differences in groups not associated with the different learning environments, essentially aiming to isolate and quantify the true effect that the learning environment may have had on student development. In this way, generalizable conclusions can be drawn that allow for better administrative decision-making under difficult circumstances such as the pandemic. It should also be clearly delimited again that this research solely concentrated on the effects the learning environment may have had on academic achievements in the selected two subject areas. Other effects such as on social-emotional well-being or peer interactions were not within the scope of this research but may be valid considerations when making educational decisions that have holistic student development in mind. Furthermore, findings of this research may be specific to the particular setting, in which the school district and individual administrators were able to provide an adequate level of technical support to all learner groups. Therefore, the study did not take into account any potential issues or differences in accessibility or support for the remote learners, and results may be vastly different in environments in which such disparities exist (Kleinke and Cross, 2021).

Results

As previously outlined in the methodology discussion, to follow up our initial investigation of differences between *R* and *H* groups in the combined data (Kleinke and Cross, 2021), i-Ready score relationships were further examined at individual grade levels (K through 8), including a third comparison group *C* as the prior year control where available (see discussion of limitations). This detailed analysis by grade level and subject area indicated:

There was a significant difference between groups for the posttest-only ANOVA in K grade level Math, $F(2, 149) = 8.36, p < 0.001, \eta^2 = 0.101$, with pairwise comparison in post-hoc testing showing a marginally insignificant difference ($t_{HC}[149] = 1.75, p = \text{n.s.}$) between the H ($N_{H,K,Math} = 67, M_H = 382, SE = 3.08$) and C ($N_{C,K,Math} = 69, M_C = 375, SE = 3.03$) groups, while the R group ($N_{R,K,Math} = 16, M_R = 403, SE = 6.29$) scored significantly higher than both the H group ($t_{RH}[149] = 2.96, p = 0.004$) and the C group ($t_{RC}[149] = 4.05, p < 0.001$). Though, as mentioned, R group size was relatively small and no pretest data were available to confirm group equivalency or account for pretest differences via covariate. As discussed, there were also not sufficient ELA data available at the K grade level to conduct any analysis for that subject area (see Figure 1).

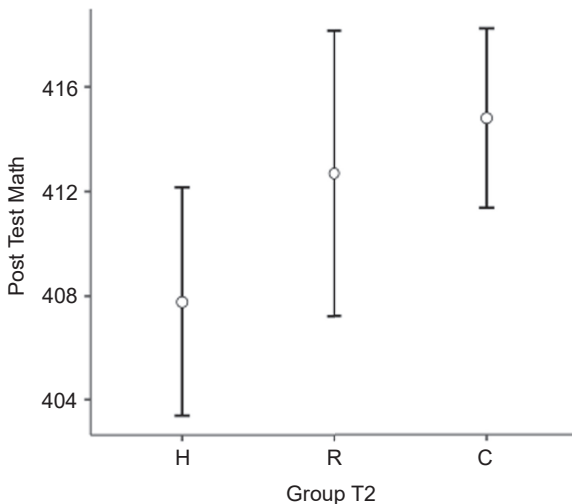
For the ANCOVA testing of 1st grade Math, the 2020 $Math_{Pretest}$ scores were a significant predictor of 2021 $Math_{Posttest}$ scores, $F(1, 122) = 216.48, p < 0.001, r = 0.79$, and grouping was, as well, a significant factor, $F(2, 122) = 3.68, p = 0.028$. Nevertheless, there was also a



Source(s): (Lenth, 2020; The jamovi project, 2021)

Figure 1. Estimated marginal means of test scores in the different groups from the posttest-only ANOVA of grade K $Math_{Posttest}$ (2021) scores

significant interaction effect between group and pretest, indicating inhomogeneous regression slopes (Field, 2009). Once accounted for, pairwise post-hoc testing revealed no significant difference between the R ($N_{R,1,Math} = 26$, $M_R = 413$, $SE = 2.76$) and C ($N_{C,1,Math} = 62$, $M_C = 415$, $SE = 1.73$) groups ($t_{R-C}[120] = -0.649$, $p = n.s.$); however, the H group ($N_{H,1,Math} = 38$, $M_H = 408$, $SE = 2.22$) scored significantly lower than the C group ($t_{H-C}[120] = -2.498$, $p = 0.014$). The difference between R and H groups was insignificant ($t_{H-R}[120] = -1.39$, $p = n.s.$). Overall, the grouping factor had a medium effect size (*partial* $\eta^2 = 0.122$; see Figure 2).



Source(s): (Lenth, 2020; The jamovi project, 2021)

Figure 2. Estimated marginal means of test scores in the different groups from the ANCOVA testing of 1st grade $Math_{Posttest}$ (2021) scores with $Math_{Pretest}$ (2020) scores as covariate

As previously discussed, testing for ELA in 1st grade was again posttest ANOVA only and revealed a significant difference in scores between groups, $F(2, 104) = 10.8, p < 0.001, \eta^2 = 0.172$. Pairwise post-hoc comparison indicated that the H group ($N_{H,1,ELA} = 57, M_H = 441, SE = 6.53$) scored significantly lower ($t_{H,R}[104] = -4.649, p < 0.001$) than the R group ($N_{R,1,ELA} = 40, M_R = 488, SE = 7.79$). Differences between the C group ($N_{C,1,ELA} = 10, M_C = 415, SE = 1.73$) and both other groups were insignificant ($t_{H,C}[104] = -0.976, p = n.s.; t_{R,C}[104] = 1.766, p = n.s.$); though, the sample size for C was relatively small, leading to a larger uncertainty in the confidence interval (see Figure 3).

For the ANCOVA testing of 2nd grade Math, the 2020 pretest scores were again a significant predictor for 2021 posttest scores, $F(1, 155) = 171.2, p < 0.001, r = 0.70$, and grouping was, as well, a significant factor, $F(2, 155) = 12.0, p < 0.001$. Nevertheless, as previously discussed for this example, there was also a significant interaction effect between group and pretest, likely caused by an outlier in the R group data. Once accounted for, pairwise post-hoc comparison indicated no significant differences between the R ($N_{R,2,Math} = 18, M_R = 435, SE = 3.32$) and C ($N_{C,2,Math} = 81, M_C = 430, SE = 1.31$) groups ($t_{R-C}[153] = 1.383, p = n.s.$), as well as H ($N_{H,2,Math} = 60, M_H = 437, SE = 1.53$) and R groups ($t_{H,R}[153] = 0.363, p = n.s.$); however, the H group scored significantly higher than the C group ($t_{H,C}[153] = 3.112, p = 0.002$). Overall, the grouping factor had a medium to large effect size (*partial* $\eta^2 = 0.231$; see Figure 4).

For the posttest-only ANOVA of 2nd grade $ELA_{Posttest}$ (2021) scores, grouping was not significant, $F(2, 182) = 2.24, p = n.s., \eta^2 = 0.024$. However, pretest group equivalency could not be verified since insufficient $ELA_{Pretest}$ data existed, not just for the C group but in all groups. Therefore, there was also no ANCOVA performed for 2nd grade ELA.

While, the lower grade analyses up to this point all had some drawbacks and inconsistencies and, consequently, only allowed limited conclusions, the following results for the middle grade levels in the data set were the most complete and, therefore, most fundamental to arriving at generalizable conclusions that may explain the overall observed effect. In general, assessment data for 3rd through 6th grade were the most comprehensive, providing adequate sample sizes and best satisfying statistical assumptions. Results, here,

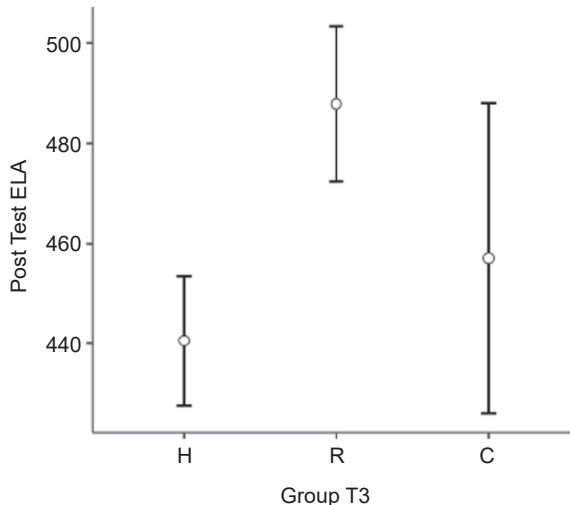


Figure 3. Estimated marginal means of test scores in the different groups from the posttest-only ANOVA of 1st grade $Math_{Posttest}$ (2021) scores

Source(s): (Lenth, 2020; The jamovi project, 2021)

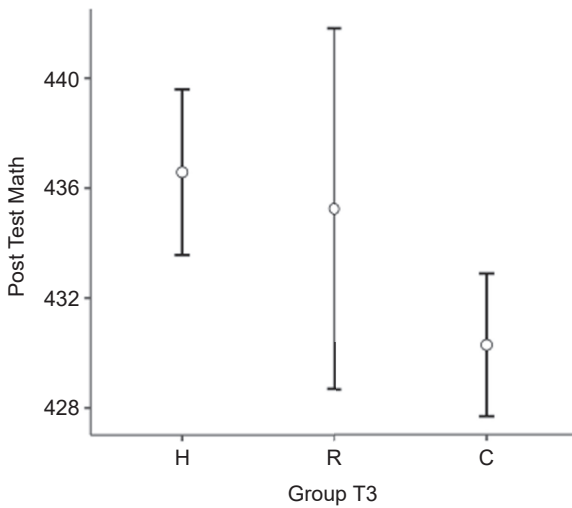


Figure 4. Estimated marginal means of test scores in the different groups from the ANCOVA testing of 2nd grade $Math_{Posttest}$ (2021) scores with $Math_{Pretest}$ (2020) scores as covariate

seemed to most consistently highlight the observed trends across grade levels and subject areas.

For 3rd grade Math, the ANCOVA testing indicated that the 2020 $Math_{Pretest}$ scores were also here a significant predictor of 2021 $Math_{Posttest}$ scores, $F(1, 164) = 207.99, p < 0.001, r = 0.73$, and grouping was, as well, a significant factor, $F(2, 164) = 6.78, p < 0.001$. Pairwise post-hoc comparison indicated no significant difference between the R ($N_{R,3,Math} = 29, M_R = 458, SE = 2.44$) and C ($N_{C,3,Math} = 81, M_C = 453, SE = 1.46$) groups ($t_{R,C}[164] = 1.65, p = n.s.$). However, the H group ($N_{H,3,Math} = 58, M_H = 448, SE = 1.73$) scored significantly lower than both the C group ($t_{H,C}[164] = -2.56, p = 0.011$) as well as the R group ($t_{H,R}[164] = -3.51, p < 0.001$). Overall, the grouping factor had a small to medium effect size ($partial \eta^2 = 0.076$; see Figure 5).

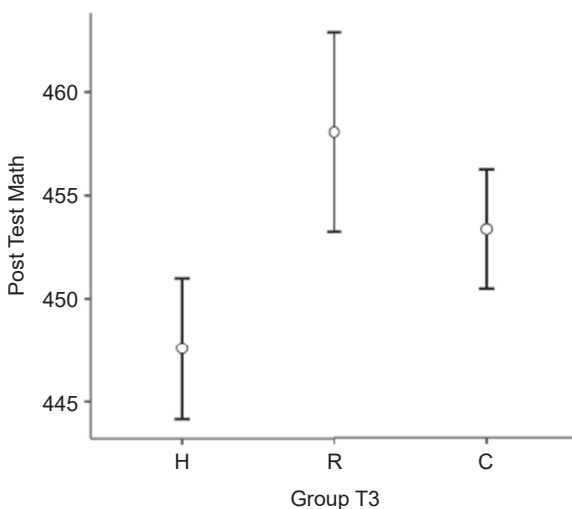


Figure 5. Estimated marginal means of test scores in the different groups from the ANCOVA testing of 3rd grade $Math_{Posttest}$ (2021) scores with $Math_{Pretest}$ (2020) scores as covariate

Similarly, ANCOVA testing of 3rd grade $ELA_{Posttest}$ scores between H and R groups indicated that the 2020 $ELA_{Pretest}$ scores were also a significant predictor, $F(1, 84) = 69.13, p < 0.001, r = 0.65$, and grouping was a significant factor, $F(1, 84) = 8.30, p = 0.005$. Post-hoc testing indicated again that the H group ($N_{H,3,ELA} = 58, M_H = 534, SE = 4.25$) scored significantly lower ($t_{H,R}[84] = -2.88, p = 0.005$) than the R group ($N_{R,3,ELA} = 29, M_R = 555, SE = 6.02$). Overall, the grouping factor had a medium effect size ($partial \eta^2 = 0.090$; see Figure 6).

Further examining $ELA_{posttest}$ scores in the two groups against the control, the posttest-only ANOVA revealed a marginally significant difference between groups, $F(2, 180) = 3.09, p = 0.048, \eta^2 = 0.033$. Pairwise post-hoc comparison indicated no significant differences between the R group ($N_{R,3,ELA} = 31, M_R = 545, SE = 7.55$; note that for some students only posttest but no pretest ELA scores existed, hence, the slightly increased sample sizes in the posttest-only ANOVA) and both other groups ($t_{H-R} [180] = -0.938, p = n.s.; t_{R-C} [180] = -0.979, p = n.s.$). Nevertheless, the H group ($N_{H,3,ELA} = 60, M_H = 536, SE = 5.43$) $ELA_{Posttest}$ scores were significantly lower ($t_{H,C}[180] = -2.476, p = 0.014$) than in the C group ($N_{C,3,ELA} = 92, M_C = 553, SE = 4.38$; see Figure 7).

For 4th grade Math, the ANCOVA testing, again, indicated that the 2020 $Math_{Pretest}$ scores were a significant predictor of 2021 performance, $F(1, 166) = 459.53, p < 0.001, r = 0.85$ and grouping was also a significant factor, $F(2, 166) = 6.57, p = 0.002$. Though, there was also a significant interaction effect between group and pretest, indicating inhomogeneous regression slopes (Field, 2009). Once accounted for, pairwise post-hoc comparison revealed no significant difference between the R ($N_{R,4,Math} = 27, M_R = 475, SE = 2.43$) and C ($N_{C,4,Math} = 80, M_C = 474, SE = 1.37$) groups ($t_{R-C}[164] = 0.208, p = n.s.$). However, the H group ($N_{H,4,Math} = 63, M_H = 467, SE = 1.58$) scored again significantly lower than both the C group ($t_{H-C}[164] = -3.182, p = 0.002$) as well as the R group ($t_{H-R}[164] = -2.498, p = 0.013$). Overall, the grouping factor had a small to medium effect size ($partial \eta^2 = 0.049$; see Figure 8).

ANCOVA testing of 4th grade $ELA_{Posttest}$ scores between H and R groups indicated that the 2020 $ELA_{Pretest}$ score was, as well, a significant predictor, $F(1, 83) = 102.39, p < 0.001, r = 0.74$, but grouping was marginally insignificant, $F(1, 83) = 3.15, p = n.s.$ Further examining the $ELA_{posttest}$ scores for 4th grade in all three groups (H, R, and C), the posttest-

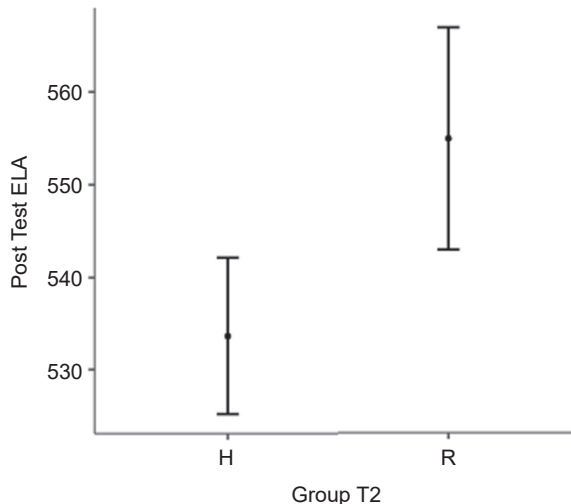
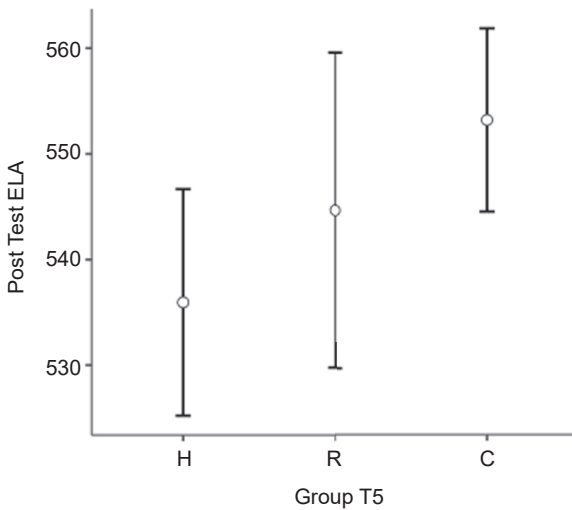


Figure 6. Estimated marginal means of $ELA_{Posttest}$ (2021) scores in the 3rd grade H and R groups when accounting for $ELA_{Pretest}$ (2020) scores as covariate

Source(s): (Lenth, 2020; The jamovi project, 2021)



Source(s): (Lenth, 2020; The jamovi project, 2021)

Figure 7. Estimated marginal means of test scores in the different groups from the posttest-only ANOVA of 3rd grade $ELA_{Posttest}$ (2021) scores

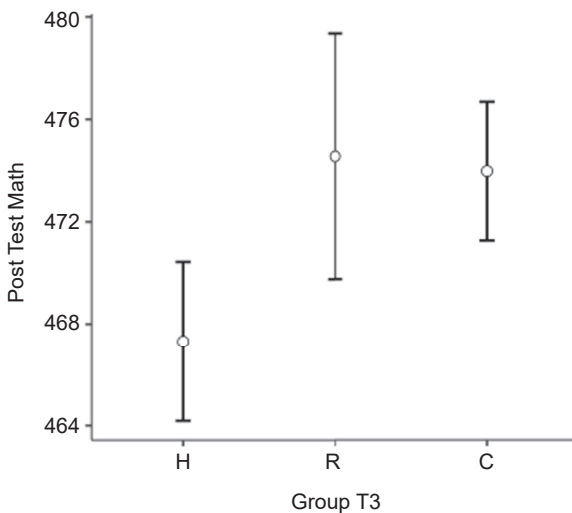
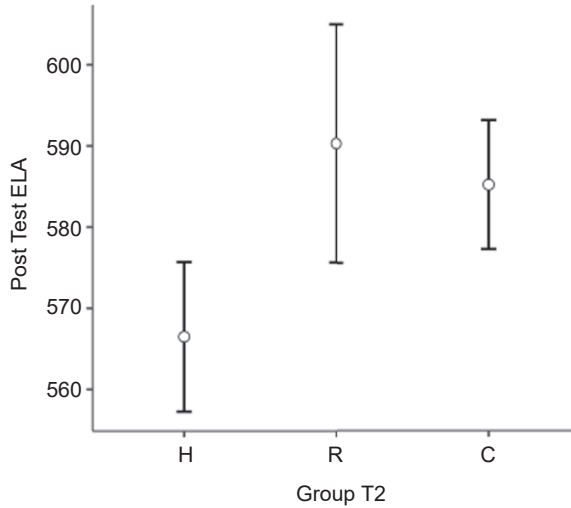


Figure 8. Estimated marginal means of test scores in the different groups from the ANCOVA testing of 4th grade $Math_{Posttest}$ (2021) scores with $Math_{Pretest}$ (2020) scores as covariate

only ANOVA revealed a significant difference between groups, $F(2, 178) = 5.97, p = 0.003, \eta^2 = 0.063$. Pairwise post-hoc comparison indicated no significant difference (t_{RC} [178] = 0.595, $p = \text{n.s.}$) between the R group ($N_{R,4,ELA} = 26, M_R = 590, SE = 7.43$) and the C group ($N_{C,4,ELA} = 89, M_C = 585, SE = 4.02$). Nevertheless, H group ($N_{H,4,ELA} = 66, M_H = 566, SE = 4.67$) $ELA_{Posttest}$ scores were significantly lower than in both other groups (t_{HC} [178] = -3.053, $p = 0.003$; t_{HR} [178] = -2.714, $p = 0.007$; see Figure 9).

Testing for 5th grade data showed similar patterns: The ANCOVA of Math scores indicated that the 2020 $Math_{Pretest}$ scores were also here a significant predictor of 2021 $Math_{Posttest}$ scores, $F(1, 162) = 399.13, p < 0.001, r = 0.83$, and grouping was, as well, a

Figure 9.
Estimated marginal means of test scores in the different groups from the posttest-only ANOVA of 4th grade $ELA_{Posttest}$ (2021) scores

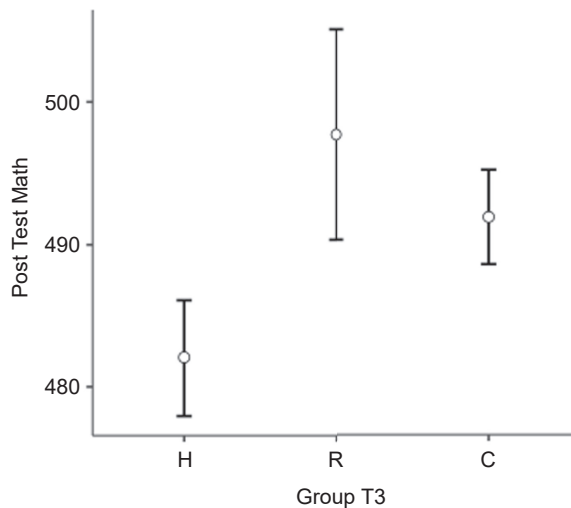


Source(s): (Lenth, 2020; The jamovi project, 2021)

significant factor, $F(2, 162) = 9.91, p < 0.001$. Pairwise post-hoc comparison showed no significant difference between the R ($N_{R,5,Math} = 18, M_R = 498, SE = 3.73$) and C ($N_{C,5,Math} = 89, M_C = 492, SE = 1.68$) groups ($t_{R,C}[162] = 1.41, p = n.s.$). However, the H group ($N_{H,5,Math} = 59, M_H = 482, SE = 2.06$) scored significantly lower than both the C group ($t_{H,C}[162] = -3.72, p < 0.001$) as well as the R group ($t_{H,R}[162] = -3.67, p < 0.001$). Overall, the grouping factor had a medium effect size ($partial \eta^2 = 0.109$; see Figure 10).

In contrast, ANCOVA testing of 5th grade $ELA_{Posttest}$ scores between H and R groups indicated that the 2020 $ELA_{Pretest}$ scores were, as well, a significant predictor for 2021 performance, $F(1, 70) = 126.19, p < 0.001, r = 0.79$, but grouping was marginally insignificant,

Figure 10.
Estimated marginal means of test scores in the different groups from the ANCOVA testing of 5th grade $Math_{Posttest}$ (2021) scores with $Math_{Pretest}$ (2020) scores as covariate

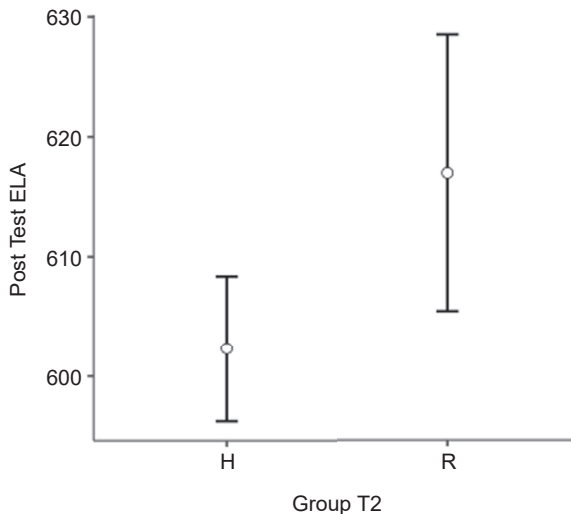


$F(1, 70) = 3.90, p = \text{n.s.}$ Nevertheless, there was also a marginally significant interaction effect between group and pretest, that, once accounted for, resulted in a significant difference ($t_{H,R} [69] = -2.25, p = 0.028$) between H group ($N_{H,5,ELA} = 57, M_H = 602, SE = 3.03$) and R group ($N_{R,5,ELA} = 16, M_H = 617, SE = 5.78$) scores (see Figure 11).

Further examining the $ELA_{posttest}$ scores for 5th grade with the posttest-only ANOVA across all three groups, there were no significant differences between groups, $F(2, 217) = 1.99, p = \text{n.s.}$ Nevertheless, also noteworthy was that sample sizes vastly differed between ANCOVA and ANOVA testing since there existed far fewer 2020 $ELA_{pretest}$ scores than 2021 $ELA_{posttest}$ ones, leading to a much smaller number of complete score pairs for ANCOVA testing than data points for the posttest-only ANOVA.

For 6th grade Math, the ANCOVA testing indicated that the 2020 $Math_{pretest}$ scores were also a significant and strong predictor of 2021 Math performance, $F(1, 112) = 268.232, p < 0.001, r = 0.84$, but grouping was not a significant factor, $F(2, 112) = 0.48, p = \text{n.s.}$ Similarly, the ANCOVA testing of 6th grade ELA data revealed no significant differences between H and R groups, $F(1, 96) = 0.327, p = \text{n.s.}$, only confirming again 2020 $ELA_{pretest}$ scores as a significant predictor, $F(1, 96) = 190.41, p < 0.001, r = 0.81$, and the posttest-only ANOVA testing across all three groups was also insignificant, $F(2, 200) = 2.27, p = \text{n.s.}$ This trend of no significant differences between groups continued into the upper grade levels; though, sample sizes also started to dramatically decrease here.

For 7th grade Math, the ANCOVA also confirmed 2020 $Math_{pretest}$ scores to be a significant and strong predictor of 2021 $Math_{posttest}$ results, $F(1, 35) = 156.63, p < 0.001, r = 0.90$, but grouping was again an insignificant factor, $F(2, 35) = 1.59, p = \text{n.s.}$ On the other hand, ANCOVA testing of 7th grade ELA data confirmed pretest scores to be a significant predictor, $F(1, 17) = 97.89, p < 0.001, r = 0.90$, but also showed a marginally significant difference between the H and R groups, $F(1, 17) = 5.33, p = 0.034, \text{partial } \eta^2 = 0.044$, with post-hoc testing indicating scores in the R group ($N_{R,7,Math} = 6, M_R = 585, SE = 9.58$) significantly exceeding ($t_{H,R} [17] = -2.31, p = 0.034$) those in the H group ($N_{H,5,Math} = 14, M_H = 559, SE = 6.26$). Nevertheless, sample sizes in this comparison were extremely small as well as



Source(s): (Lenth, 2020; The jamovi project, 2021)

Figure 11. Estimated marginal means of $ELA_{posttest}$ (2021) scores in the 5th grade H and R groups when accounting for $ELA_{pretest}$ (2020) scores as covariate

unequal. The posttest-only ANOVA of ELA scores across all three groups revealed again no significant differences, $F(2, 40) = 1.91, p = n.s.$

Similarly, ANCOVA testing for 8th grade scores revealed no significance of the group factor, neither in Math, $F(1, 16) = 2.66, p = n.s.$, nor in ELA, $F(1, 13) = 0.009, p = n.s.$ Nevertheless, sample sizes were also extremely small and, based on the way control group data were derived from prior year scores, no control group comparison was available for 8th grade scores. However, the ANCOVA testing confirmed again the usefulness of the pretest scores as covariate and posttest score predictor for both cases Math, $F(1, 16) = 26.61, p < 0.001, r = 0.77$, as well as ELA, $F(1, 13) = 27.06, p < 0.001, r = 0.82$.

Finally, a repeated measures ANOVA was conducted on the entire data set, comparing the year-over-year development in student i-Ready Math scores between the H and the R groups for the time series 2019, 2020, and 2021. Post Hoc testing revealed a marginally significant difference ($t[406] = -1.99, p = 0.047$) for the grouping factor for the 2021 results, while 2019 and 2020 did not differ significantly. The score development over time for both groups is depicted in Figure 12.

Discussion

Although, the statistical results of the time series ANOVA were not the strongest of all of the analyses that were conducted (likely because the aggregated data of all the grades contained also upper grade level data with little differences as well as year-over-year test score development differences among grade levels that contributed to increased uncertainty intervals), their graphical depiction may paint the clearest picture of what likely happened: While the score development of fully remote students essentially followed a continuous line through the three years, hybrid students' score development decreased in slope between 2020 and 2021 (i.e. during the pandemic), indicating that the pre-pandemic development of year-over-year learning performance could not be maintained for that group.

Similarly, analysis results for grades 3 through 6 (for which, as previously mentioned, the data were the most complete) consistently indicated that the hybrid group (H) performed significantly lower than the remote group (R), the prior-year classroom control group (C), or both. Furthermore, this observed trend can also be found in most of the lower grade levels below third grade, though, as previously discussed, the conclusiveness of these results was more limited due to the unavailability of either covariate or control group comparison data, or both. The only exception to this trend was in second grade Math, for which the prior-year control group (C) scores were somewhat lower than in both other groups (R and H). This specific case seemed to be an outlier in the study, which potentially could have been caused by

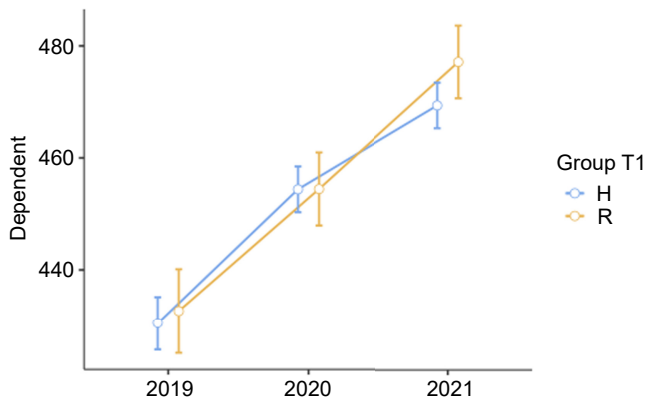


Figure 12. Estimated marginal means of math test score development over time in the two different groups from the repeated measures ANOVA

below average performance of second grade students on the standardized test in 2020 (i.e. low 2020 test results would negatively affect C scores but also inflate H and R ones due to their use as a covariate for those two groups).

Therefore, in general, the grade-wise analysis seems to indicate that the hybrid learning environment (with all its pandemic-induced inconsistencies) negatively affected student achievements in both Math and ELA for students at or below 6th grade. This observed effect seemed to diminish in grades 7 and 8, leading to the tentative conclusion that the learning environment has the greatest effect on student achievements in elementary education, though generalized conclusions about the higher grade levels in the study are more limited due to the limited availability of data in those grades. Therefore, both $H1_0$ and $H2_0$ were rejected for this study.

Furthermore, R performance seemed, overall, consistent with the achievements of the prior-year control group (C) in almost all analyses for which control group data were available. Thus, there was no indication that the fully-remote environment significantly differed from the conventional classroom environment concerning the learner development in the two subject areas Math and ELA. The only exception to this observed trend was in the K- grade Math data, for which scores in the R significantly exceeded those of both other groups (H and C; though, as previously mentioned, group size and unavailability of pre-test covariate data make this result less reliable). Such anomaly could have been caused by an exceptional teacher or the additional support that K-grade remote students received at home, or both. In either case, nowhere in this study did students in the R group perform significantly worse than the classroom control group (C), suggesting that a comparative level of subject-specific education can be achieved in fully-remote elementary education, even under adverse conditions such as during a global pandemic.

Therefore, what remains to be discussed is our perspective on the implications of the observed trends. As previously highlighted, through the research design and the analysis methods, we tried to completely and consistently eliminate any effects that the testing (e.g. abnormalities during test-taking at home) or non-random group assignments (i.e. the self-selection of the remote learning group) may have had on analysis results. By eliminating outliers and employing a control group design, as well as pre-test covariate data, we are confident that the observed group differences are, indeed, a genuine effect of the learning environment on student development. Furthermore, we observed that the biggest differences in those two learning environments during the pandemic (H and R) were likely caused by two main groups of factors:

Environmental factors, which included aspects such as comfort, familiarity/experience, attention/care/support and consistency, all of which may have contributed to students and their parents electing to continue fully remote even though a hybrid option was available. These fully remote students likely stayed fully remote because their HLE could support this modality, providing a consistent learning experience that properly coordinated responsibilities between the remote facilitators and the students' home support system. Furthermore, the school that served as the setting for this study afforded a consistent approach to remote learning, fully embracing the modality and supporting the remote teachers and learners, as well as providing the necessary support structures, equipment, and processes.

In contrast, at least some of the students returning to hybrid learning may have had less choice due to the lack of a sufficient HLE and support for remote learning. They also may have faced a higher degree of inconsistencies due to ever-changing public health guidelines for the return to in-person learning. Thus, our findings seem to be consistent with recent research that stresses the importance of a supportive HLE (e.g. [Dimosthenous et al., 2020](#); [Lehrl et al., 2020a, b](#)) and, in contrast to [Driessen and Slegers \(2000\)](#) or [Jamieson and Wikeley \(2000\)](#), we found evidence for the benefits of consistency and continuity in the learning

environment. Additional research could focus on student and parent perceptions about the available learning choices, as well as environmental factors such as the availability of home support, that may have contributed to the decision to either continue remotely or return to hybrid learning. Furthermore, follow-on research should examine whether and when the observed group differences equalize again, as learning returns to normal, allowing to also investigate any long-term effects the pandemic may have had on learners.

Developmental factors, which included aspects such as the development of responsibility, discipline, independence and individual decision-making/choice all of which may have contributed to students remaining fully remote if their learning preferences were supported in such an environment. While not the immediate focus of this study and its literature review, the ability to self-direct and self-regulate one's learning has been previously identified as crucial to online education in older learners (Association for the Study of Higher Education [ASHE], 2014). Thus, it seems plausible that similar prerequisites and learner dispositions may also hold true for elementary students in a remote learning environment. Therefore, if students chose to remain fully remote because this modality best aligned with their learning preferences, and such a fitting learning environment allowed the students to excel despite of the pandemic circumstances, it would also support the importance of individualized learning approaches that cater to student needs and predispositions. Though, further research, especially into the role of learning preferences in younger learners, seems necessary.

Conclusion

Our findings from this second part of the analysis seem to confirm our tentative conclusions presented in the first part (Kleinke and Cross, 2021), namely that environmental and developmental factors played a major role in student performance in the different learning environments during the pandemic. Consistency and continuity in the fully remote environment seemed to allow the learners to achieve a performance comparable to the pre-pandemic classroom control. In contrast, inconsistencies in the hybrid environment, as well as reduced home support, may have contributed to lower achievement levels for learners in that environment. Furthermore, the observed effects seemed to decrease with learner age and showed no significant differences in higher grade levels. More research is needed to confirm contributing factors to the observed differences between learning environments, as well as any long-term consequences of the environment choice.

References

- Armfield, J.M., Ey, L., Zufferey, C., Gnanamanickam, E.S. and Segal, L. (2021), "Educational strengths and functional resilience at the start of primary school following child maltreatment", *Child Abuse and Neglect*, Vol. 122, p. 1, doi: [10.1016/j.chiabu.2021.105301](https://doi.org/10.1016/j.chiabu.2021.105301).
- Association for the Study of Higher Education (ASHE) (2014), "Learning theories and student engagement", *ASHE Higher Education Report*, Vol. 40 No. 6, pp. 15-36.
- Ben-Shachar, M., Makowski, D. and Lüdtke, D. (2020), "Compute and interpret indices of effect size", available at: <https://easystats.github.io/effsize/index.html>.
- Bunch, M.B. (2017), "Review of i-ready K-12 diagnostic and K-8 instruction", in Carlson, J.F., Geisinger, K.F. and Jonson, J.L. (Eds), *The Twentieth Mental Measurements Yearbook*, (test.8597). EbscoHost: Mental Measurements Yearbook with Tests in Print [Database].
- Cantley, I., O'Meara, N., Prendergast, M., Harbison, L. and O'Hara, C. (2021), "Framework for analysing continuity in students' learning experiences during primary to secondary transition in mathematics", *Irish Educational Studies*, Vol. 40 No. 1, pp. 37-49, doi: [10.1080/03323315.2020.1779108](https://doi.org/10.1080/03323315.2020.1779108).
- Curriculum Associates (2021), "iReady", available at: <https://www.curriculumassociates.com/products/i-ready>.

- Dewey, J. (1938), *Experience and Education*, Collier-MacMillan Canada, Toronto.
- Dhawan, S. (2020), "Online learning: a panacea in the time of COVID-19 crisis", *Journal of Educational Technology Systems*, Vol. 59 No. 1, pp. 5-22, doi: [10.1177/0047239520934018](https://doi.org/10.1177/0047239520934018).
- Dimosthenous, A., Kyriakides, L. and Panayiotou, A. (2020), "Short- and long-term effects of the home learning environment and teachers on student achievement in mathematics: a longitudinal study", *School Effectiveness and School Improvement*, Vol. 31 No. 1, pp. 50-79, doi: [10.1080/09243453.2019.1642212](https://doi.org/10.1080/09243453.2019.1642212).
- Driessen, G. and Slegers, P. (2000), "Consistency of teaching approach and student achievement: an empirical test", *School Effectiveness and School Improvement*, Vol. 11 No. 1, pp. 57-79, doi: [10.1076/0924-3453\(200003\)11:1;1-A;FT057](https://doi.org/10.1076/0924-3453(200003)11:1;1-A;FT057).
- Ezzelle, C. (2017), "Review of the i-ready K-12 diagnostic and K-8 instruction", in Carlson, J.F., Geisinger, K.F. and Jonson, J.L. (Eds), *The Twentieth Mental Measurements Yearbook*, (test.8597). EbscoHost: Mental Measurements Yearbook with Tests in Print [Database].
- Ferri, F., Grifoni, P. and Guzzo, T. (2020), "Online learning and emergency remote teaching: opportunities and challenges in emergency situations", *Societies*, Vol. 10 No. 4, p. 86, doi: [10.3390/soc10040086](https://doi.org/10.3390/soc10040086).
- Field, A. (2009), *Discovering Statistics Using SPSS*, 3rd ed., Sage, Thousand Oaks, CA.
- Fontenelle-Tereshchuk, D. (2021), "'Homeschooling' and the COVID-19 crisis: the insights of parents on curriculum and remote learning", *Interchange*, Vol. 52 No. 2, pp. 167-191, doi: [10.1007/s10780-021-09420-w](https://doi.org/10.1007/s10780-021-09420-w).
- Fox, J. and Weisberg, S. (2020), "Car: companion to applied regression", R package, available at: <https://cran.r-project.org/package=car>.
- Gallucci, M. (2019), "GAMLj: general analyses for linear models", jamovi module, available at: <https://gamlj.github.io/>.
- Gunzenhauser, C., Enke, S., Johann, V.E., Karbach, J. and Henrik, S. (2021), "Parent and teacher support of elementary students' remote learning during the COVID-19 pandemic in Germany", *AERA Open*, Vol. 7, doi: [10.1177/23328584211065710](https://doi.org/10.1177/23328584211065710).
- Jamieson, I. and Wikeley, F. (2000), "Is consistency a necessary characteristic for effective schools?", *School Effectiveness and School Improvement*, Vol. 11 No. 4, pp. 435-452, doi: [10.1076/sesi.11.4.435.3564](https://doi.org/10.1076/sesi.11.4.435.3564).
- Klinke, S. and Cross, D. (2021), "Remote elementary education: a comparative analysis of learner development (part 1)", *Journal of Research in Innovative Teaching and Learning*, doi: [10.1108/JRIT-08-2021-0055](https://doi.org/10.1108/JRIT-08-2021-0055).
- Lehrl, S., Ebert, S., Blaurock, S., Rossbach, H.-G. and Weinert, S. (2020a), "Long-term and domain-specific relations between the early years home learning environment and students' academic outcomes in secondary school", *School Effectiveness and School Improvement*, Vol. 31 No. 1, pp. 102-124, doi: [10.1080/09243453.2019.1618346](https://doi.org/10.1080/09243453.2019.1618346).
- Lehrl, S., Evangelou, M. and Sammons, P. (2020b), "The home learning environment and its role in shaping children's educational development", *School Effectiveness and School Improvement*, Vol. 31 No. 1, pp. 1-6, doi: [10.1080/09243453.2020.1693487](https://doi.org/10.1080/09243453.2020.1693487).
- Lenth, R. (2020), "Emmeans: estimated marginal means, aka least-squares means", R package, available at: <https://cran.r-project.org/package=emmeans>.
- Murai, Y. and Muramatsu, H. (2020), "Application of creative learning principles within blended teacher professional development on integration of computer programming education into elementary and middle school classrooms", *Information and Learning Science*, Vol. 121 No. 7, pp. 665-675, doi: [10.1108/ILS-04-2020-0122](https://doi.org/10.1108/ILS-04-2020-0122).
- O'Leary, G., Schnake-Mahl, A., Vaidya, V., Bilal, U. and Kolker, J. (2021), "Indoor dining and in-person learning: a comparison of 30 US cities", *International Journal of Environmental Research and Public Health*, Vol. 18 No. 20, p. 10967, doi: [10.3390/ijerph182010967](https://doi.org/10.3390/ijerph182010967).

-
- R Core Team (2021), "R: a language and environment for statistical computing, Version 4.0", Computer software, R packages, MRAN snapshot 2021-04-01, available at: <https://cran.r-project.org>.
- Schafer, E.C., Dunn, A. and Lavi, A. (2021), "Educational challenges during the pandemic for students who have hearing loss", *Language, Speech and Hearing Services in Schools (Online)*, Vol. 52 No. 3, pp. 889-898, doi: [10.1044/2021_LSHSS-21-00027](https://doi.org/10.1044/2021_LSHSS-21-00027).
- Singh, J., Steele, K. and Lovely Singh, L. (2021), "Combining the best of online and face-to-face learning: hybrid and blended learning approach for COVID-19, post vaccine, and post-pandemic world", *Journal of Educational Technology Systems*, Vol. 50 No. 2, pp. 140-171, doi: [10.1177/00472395211047865](https://doi.org/10.1177/00472395211047865).
- Singmann, H. (2018), "afex: analysis of factorial experiments", R package, available at: <https://cran.r-project.org/package=afex>.
- Storey, N. and Slavin, R.E. (2020), "The US educational response to the COVID-19 pandemic", *Best Evidence in Chinese Education*, Vol. 5 No. 2, pp. 617-633, available at: [10.15354/bece.20.027](https://doi.org/10.15354/bece.20.027).
- The jamovi project (2021), "Jamovi (Version 1.8)", Computer Software, available at: <https://www.jamovi.org>.
- Van Ameijde, J., Weller, M. and Cross, S. (2018), "Learning design for student retention", *Journal of Perspectives in Applied Academic Practice*, Vol. 6 No. 2, pp. 2-16, doi: [10.14297/jpaap.v6i2.318](https://doi.org/10.14297/jpaap.v6i2.318).

Corresponding author

Stefan Kleinke can be contacted at: kleinkes@erau.edu