Remote elementary education: a comparative analysis of learner development (part 1)

Stefan Kleinke

COA Graduate Studies, Embry-Riddle Aeronautical University Worldwide and Online, Daytona Beach, Florida, USA, and

David Cross

COA, Embry-Riddle Aeronautical University Worldwide and Online, Daytona Beach, Florida, USA

Abstract

Purpose – The purpose of this two-part research was to investigate the effect of remote learning on student progress in elementary education. Part one, presented in this paper, examined achievement differences between learners in a fully remote learning environment and those in a hybrid setting.

Design/methodology/approach – A quantitative, quasi-experimental study with factorial design was used to investigate group differences in student achievement between the different learning environments. Ex-post-facto data from standardized test scores were utilized to examine in which ways the learning environment may have affected learner progress in two distinct subject areas crucial to elementary education: English language (ELA) and math.

Findings – Findings revealed a significant difference between the two learning environments in both subject areas. While preexisting group differences, selection biases and testing inconsistencies could be effectively ruled out as potential causes for the observed differences, other factors such as developmental and environmental differences between the learning environments seemed to be influential. Therefore, the follow-on research aimed at further investigating and confirming the influence of such factors and will be presented in a Part 2 paper.

Practical implications – Knowledge of the observed differences in learning achievements between the different environments, as well as the factors likely causing them, may aid educators and school administrators in their decision processes when faced with difficult circumstances such as during the pandemic.

Originality/value – When the SARS-CoV-2 virus started to rapidly spread around the globe, educators across the world were looking for alternatives to classroom instruction. Remote learning became an essential tool. 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. Lacking this type of information, it seems that educators and administrators are facing difficult decisions when trying to align the often conflicting demands of public health, local politics and parent pressure with what may be best for student learning.

Keywords Elementary education, Remote learning, Learning assessment

Paper type Research paper

When, in the spring of 2020, the SARS-CoV-2 virus started to rapidly spread around the globe, public health officials and governments in many countries reacted by implementing social distancing strategies aimed at reducing community transmission. Such strategies included, for example, cancellation of mass events, suspension of group activities, limitations to indoor gatherings as well as stay-at-home orders. While, in the USA, the response was not

necessarily unified, by April 2020, many states and local communities had enacted at least some form of restriction, most often including the suspension of in-person, indoor schooling (Storey and Slavin, 2020). Therefore, educators across the country were scrambling to provide suitable alternatives for their students that would allow continuing education without the need for in-person, in-classroom meetings. The application of remote learning strategies became an essential tool in achieving this goal (Storey and Slavin, 2020).

Nevertheless, while online learning has been an established modality in higher education for many years, with a rapid growth in application seen throughout the last two decades (Allen and Seaman, 2017), its widespread application in elementary education seems entirely new and uniquely driven by the circumstances of the global pandemic. Thus, in contrast to e-learning in postsecondary education, for which an abundance of research has been conducted to establish its effectiveness and investigate its level of efficiency, as well as any factors influencing its successful application, relatively little is known about the efficacy of remote learning approaches in elementary education or any influences that may affect the learning outcomes in such environments. Lacking this type of information seems that educators and school administrators are facing difficult decisions when trying to align the often conflicting demands of public health, local politics and parent pressure with what may be best for student learning.

Therefore, the purpose of this two-part research was to investigate the effect of remote learning on student progress in elementary education. A quantitative, quasi-experimental study with factorial design was used to investigate group differences in student achievement between the differing learning environments. Ex-post-facto data from standardized test scores were used to examine in which ways the learning environment may have affected learner progress in two distinct subject areas that are crucial to elementary education: English language (ELA) and math. For the first part of the study, presented in this paper, the following two hypotheses were sought to be tested:

H1. There are no significant differences in students' ELA development between the different learning environments.

H2. There are no significant differences in students' math development between the different learning environments.

Literature review
Recent trends show that the education system is changing. What is happening in the world has led to the fact that almost all countries were forced to switch to online education. The forced transition to distance learning raises a number of questions regarding the quality and effectiveness of education. In addition, in such conditions, the learning style is also changing, where the opportunities that the Internet and technologies provide differ from the standard teaching method. In this regard, it is advisable to conduct a literature review that will assess the effectiveness of distance learning in elementary school and the difference between verbal and computational styles of education.

Effectiveness of distance/online learning in primary school
Recently, distance learning has replaced full-time education not only in higher educational institutions but also in primary school. According to Ferri *et al.* (2020), in the context of a pandemic, online teaching is of utmost importance, since students can study from anywhere and at any time and also avoid riding in crowded public transport. Given the current situation with coronavirus around the world, these factors are important not only in the fight against the spread of the disease but also in the desire to maintain the continuity of education. In

addition, parents can save time and have the flexibility of choice (Ferri *et al.*, 2020). However, the benefits of distance education do not make it an effective way of teaching children, especially when it comes to using the approach in relation to primary school students who are at risk of various problems (Cardullo *et al.*, 2021). Younger learners may find it more difficult to be in a remote environment and require additional help with technology from parents who are not always able to be with them throughout the learning process. Moreover, a number of other problems accompany distance teaching that indicates the ineffectiveness of such an alternative way of education. Finally, the results of the McKinsey survey showed that in most countries, teachers rate online learning at five points out of ten. Moreover, teachers from all countries report that the greatest losses in education are in primary grades (Chen *et al.*, 2021). Thus, the experience of the last year shows the low efficiency of distance learning, as evidenced by many factors.

In a pandemic, that has affected the whole world, online learning has allowed teachers and students to adapt and made education accessible. Nevertheless, Domina *et al.* (2021) believed that the first signs indicate that such an approach to teaching is ineffective, since according to preliminary data, only 9% of students constantly go through distance learning, and in the opinion of parents, children acquire less knowledge, in contrast to full-time education. On the one hand, the authors found a positive association between engagement and technology use among high school and college students. However, on the other hand, researchers argue that school closures have dire consequences for the social and academic development of primary school students (Domina *et al.*, 2021). Social inefficiency is associated with risks of isolation and lack of proper communication between school and families. The low level of interaction interferes with the accumulation of social capital that serves as a resource for children and their social development (Domina *et al.*, 2021). In addition, a number of other reasons of why students show a low level of participation in online learning include the lack of adequate resources and infrastructure in some school districts, as well as the degree to which low-income families and minorities have access to technology (Domina *et al.*, 2021). According to the researchers, these factors affect the involvement of primary school students in the distance learning process.

It is important to mention that technological factors are not limited to resource constraints. Ferri *et al.* (2020) emphasize such a problem as the poor knowledge of young children of digital devices and programs, the use of which is required during online learning. In addition, the authors note the concern of parents about the long time spent by children in front of their phones and other means of communication to watch lessons that sometimes last many hours due to the lack of optimized content. At the same time, the authors pay attention to pedagogical challenges that have an impact on online learning. The new teaching format presupposes approaches to education that differ from the strategies used in face-to-face lessons. First, teachers should use more innovative methods that will keep the attention and participation of children for a long period during online lessons (Ferri *et al.*, 2020). Therefore, the experience of teachers plays a major role in the success of distance learning in primary schools where students have a lower level of perseverance and understanding of the whole process. However, as shown by a study of Cardullo *et al.* (2021) that conducted a survey of teachers, including those of primary school, only 9.22% of teachers had experience of online education at the time of its introduction in 2020. In addition, according to Ferri *et al.* (2020), it was possible to identify problems associated with the lack of teachers' skills in the use of technology that prevents more effective planning and implementation of the pedagogical program. Consequently, the low effectiveness of distance learning is due to many technical and pedagogical reasons.

Another problem that decreases the effectiveness of distance learning is emotional loss that includes the lack of sufficient interaction not only with teachers but also with peers. According to Cardullo *et al.* (2021), distance education causes a loss of relationships that is

detrimental to the training experience of elementary school students. As a result, 60% of parents are concerned about maintaining social bonds and friendships, and 59% are worried about the emotional well-being of their children (Horowitz and Igielnik, 2020). Thus, the transition to distance learning contributed to the emotional uncertainty of the participants in the educational process. Among other things, the problem of assessment and integrity of results has become another reason of the ineffectiveness of online teaching. Cardullo *et al.* (2021) stated that teachers have limited time to assess students' understanding of the material and their overall knowledge. In addition, distance learning does not provide an opportunity to control the performance of assigned tasks by children, since many teachers claim that parents often do the work for their children (Cardullo *et al.*, 2021). While parental involvement is an important factor in student achievement, sometimes their desire to help interferes with assessing the child's ability. Garbe *et al.* (2020) explained the increased involvement of parents by the fact that with the start of online learning, parents and their children happened to be unprepared; therefore, they had to adopt new and unfamiliar roles and responsibilities that they do not always understand. Finally, it is worth mentioning that data from the USA indicate that after distance education, students lag one to three months behind in reading and math skills (Chen *et al.*, 2021). Thus, these factors reveal the ineffectiveness of distance learning in primary school.

Differences in verbal versus computational learning of remote/in class students and any factors that may exaggerate these differences
In the context of the education system, the role of teaching methods is expanding. For example, computational education has recently become very popular, while verbal learning has declined with the transition to online teaching. Although both approaches are used in both face-to-face and distance environments, there is still a big difference between them. First, verbal and computational learning differ in the way that information is transmitted to students, regardless of whether they are remotely or in the classroom. Verbal teaching involves the provision of information in the classroom through the expression of speech, sentences, spoken and written words and signs. Thus, the educational process takes place in the form of communication and interaction during which the student receives knowledge from the teacher that he/she must remember (Van Tetering *et al.*, 2018). Computational learning relies less on the traditional verbal learning and requires the use of computers and computing technologies that provide more possibilities for solving problems. The advantage of computational learning, as emphasized by Barchas-Lichtenstein *et al.* (2020), is that the approach is viewed as a way to support independent active learning. Although with the help of such computer technologies, students can learn to develop various algorithms, routines and procedures within the framework of mathematics or computer science lessons, the method aims more for the development of logical thought processes. However, while computational learning is increasingly appearing in school practices, the purpose of the method is mostly to complement the standard education system (Barchas-Lichtenstein *et al.*, 2020). Thus, both styles differ not only in the way they convey information but also has verbal learning long played the role of the main approach to in-classroom teaching.

Although computational learning has seen a tremendous increase in recent years, the role and relevance of the verbal learning style may be changing as a result of this newer learning style. Falloon (2016) noted that computing programs are widely implemented as part of basic education in compulsory education systems. Many governments view computer-based instructional approaches as valuable and beneficial to learners as the older verbal teaching approach. Learning programming that underlies computational learning and includes sequencing, understanding triggers and events, working with conditionals and other concepts is a valuable skill in the 21st century (Falloon, 2016). Hence, by receiving this kind of

knowledge, students can improve their intellectual abilities that will help in solving various problems. Consequently, with the advent of such computer programs, verbal learning is no longer as prominent as a stand-alone learning style. The evidence comes from data showing that recently, in the context of a pandemic and the transition to online education, teachers spend only half a week on verbal meetings with students (Barchas-Lichtenstein *et al.*, 2020). However, in the process of verbal learning and communication between a teacher and students, the level of their development, upbringing and education rises. By assessing verbal responses, written assignments, facial expressions, posture, gestures, tone and pauses, one can understand the level of knowledge of children. While computational learning is still at an early stage of development, it has disadvantages associated with assessing students. Grover (2017) argued that often in programming curricula, there is no function of monitoring the knowledge gained by students. Thus, different levels of ability assessment are another difference between the learning styles under consideration.

Among the factors that can exaggerate the differences between learning methods are students' ability and interest. According to Lu and Yang (2018), some students prefer to receive information from what they hear and say. Respectively, the verbal learning style is more effective and meaningful for them as it allows them to achieve better results. On the other hand, other students give preference to visual content with the help of which they best assimilate information (Lu and Yang, 2018). For these children, computational learning is more likely to meet their needs. Van Tetering *et al.* (2018) called such factors individual differences due to which students may face difficulties in learning certain subjects and that may negatively affect their motivation. Therefore, students' experience is one of the factors that can reinforce the difference between verbal and computational education. In addition, Grover *et al.* (2016) also noted that interests and attitudes towards the subject determine which approach is more meaningful for students. Finally, innovations can further increase the difference between the styles that include new technologies often changing the nature of tasks and the very concept of knowledge (Grover, 2017). Although verbal learning will always be needed, its traditional prominence (sometimes being the only option) in the classroom seems to be waning. Therefore, over time, computational learning may become a dominant learning style in the education system.

Summary
In the course of the work, an analysis of the literature related to the field of online education was conducted. Research has shown that distance learning is less effective both in the opinion of teachers and parents, as well as in students' outcomes. Online learning has many disadvantages and problems that elementary school children can face. In addition, changes in the education system have led to the fact that computational learning is becoming more and more relevant, in contrast to verbal learning that differs in the way information is transmitted and in the level of assessment of student's knowledge. Additionally, learners' individual abilities, experiences, interests, attitudes and a complete shift towards the use of technology in teaching can further exaggerate such differences.

Research method
For the first part of this two-part research, a group comparison was conducted 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). Since this quasi-experimental research was based on preexisting assessment data and the group assignment was, therefore,

non-random, the between-subject group comparison was further enhanced by using a within-subject paired pretest–posttest design, effectively permitting to account for individual pretreatment differences among students and groups and allowing to assess learner development rather than just their final achievement. Thus, students' pretest scores (from before the pandemic) served as covariate moderating any observed differences in the posttest assessments. For each student, pretest and posttest scores were independently collected and analyzed for the two subject areas math and ELA.

To further scrutinize any findings for this overall comparison, the assessment data were additionally broken down and analyzed by individual grade levels, which also allowed adding a grade-based control (classroom learning group, coded C) to the design. These control group data were derived from the corresponding grade-level test scores of previous years not spent under any pandemic-related conditions, such as remote or hybrid learning and, thus, allowed comparing student development during the pandemic in each grade level (K through 8) and group (H or R) against a prepandemic standard for that respective grade level. The analysis and discussion of this breakdown by grade level with control will be presented in a follow-on Part 2 paper.

For this study, assessment data for a total sample size of $N = 904$ students from grade levels K through 8 were examined, though individual sample sizes in the groups and grade levels varied widely during the analysis due to varying availability of assessment scores as well as exclusion criteria applied during preparation and treatment of the data. 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 comparison (as will be discussed in further detail). As a setting for the study, a medium-sized charter school in North Central Colorado was selected based on the researcher's familiarity with the school's assessment methods and detailed knowledge about its pandemic-related policy implementations. This particular school 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 USA (Bunch, 2017; Curriculum Associates, 2021; Ezzelle, 2017). Its recurring, computer-based application made it ideal for this research since students in all learning environments experienced the test in the same ways and under the same conditions, ensuring that no differences in measurement unduly influenced the results. Furthermore, since the test was already regularly administered three times per year (September, January and May) before the pandemic, control group data readily existed to allow comparison to the current test results. 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 pretest tool for this research. Thus, students' mid-term (i.e. January) test scores in ELA and math for 2019, 2020 and 2021 were used to establish the pretest–posttest design (i.e. between the 2020 and 2021 scores) as well as the control group data (i.e. utilizing the 2019 and 2020 student scores for each corresponding grade level).

Here, another limitation to the study's 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 control groups across all grade levels lacked pretest covariate data, making the assessment of student development in math the more robust analysis of the two subject areas under consideration. 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 first grade pretest data in ELA (i.e. students that were in 1st grade during posttest assessments in 2021, were, obviously, in K during pretest in 2020). Similarly, since control group data were derived by utilizing current students' 2019 and 2020 scores and recoding them for use as control at one grade level lower, no such conversion could be made when examining the entire dataset as a whole. Therefore, no independent control group data was, obviously, available when conducting analyses on the entire dataset.

Such test on the entire set of math scores was done first to gain an initial idea about any differences between the two pandemic-induced learning environments R and H. An analysis of covariance (ANCOVA) was conducted on the posttest i-Ready Math scores (2021) using the two-level (R or H) grouping variable as a factor and the pretest (2020) scores as a covariate. To further examine how any such difference may have played out across the various grade levels and also to compare results against a control, the overall test was followed up with individual analyses of covariance (ANCOVA) on the posttest i-Ready scores (2021 scores for groups R and H and 2020 scores for group C) for each grade level (K through 8), utilizing the three-level (R, H or C) grouping variable as a 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 pretest data were not available (see previous limitations discussion), the treatment was reduced to an analysis of variance (ANOVA) among the three groups only. For the ELA subject area, for which control group pretest data were not available (see previous limitations discussion), both an ANCOVA between only the two groups H and R as well as a posttest-only ANOVA between all three groups were conducted.

All statistical analysis was performed using the jamovi software tools (The Jamovi Project, 2021), including applicable plug-in packages and visualizations (Ben-Shachar and Lüdecke, 2020; Fox and Weisberg, 2020; Galluci, 2019; Length, 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. Based on this pretreatment of the data, scores outside three standard deviations (SDs) were, for example, excluded from the analysis. Similarly, during the main analysis, assumptions were continuously reexamined, for example, investigating the proper fit of the underlying model by thoroughly inspecting residuals outside 2.5 SD. In addition to the main analysis and those 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 student achievements in the different learning environments during the pandemic compared to, and progressed from, the year preceding it.

A particular emphasis during the preparation and treatment of the data was put on identifying and eliminating any potential effects of differences in testing. Since remote students took their i-Ready assessments at home, a potential risk to validity was that some of these students may have received unauthorized help during the testing. Therefore, special attention was given to any unusual distribution or combination of score data within the groups, and all main ANCOVA/ANOVA analyses were further followed up with examinations of the underlying general linear regression modeling. For example, in-depth examination of the occurrence of inhomogeneous regression slopes between groups in the grade 2 math scores (also apparent by a significant interaction effect during ANCOVA testing; Field, 2009) revealed a potentially undue influence of a single pretest–posttest score combination (i.e. a student advancing from a relatively low pretest score to an unusually high posttest score) in one of the groups, affecting the overall outcome of the analysis. Thus, once

accounted for such abnormalities (e.g. by stepwise developing the regression model to account for interactions), the observed effect on estimated marginal means reduced and confidence intervals widened somewhat, providing a more reliable picture of group differences. In essence, the goal during follow-up analysis was to establish which part of any observed differences was a genuine effect (i.e. a difference in intercept of parallel regression lines between groups indicating a different performance of the group as a whole) and which part was potentially caused by unusual score distributions within groups (i.e. differences in slopes of regression lines between the groups indicating unusual activity within parts of one or more groups; Kleinke, 2020).

To directly further clarify and delimitate, the purpose of this study was not to identify any individual differences in students' achievement scores or determine their potential causes (e.g. cheating). Rather, the goal was to eliminate any such outliers and undue influences on results from the analysis to arrive at conclusions about the level of effectiveness of the different, nonconventional learning approaches in elementary education during the unusual circumstances brought about by the pandemic. Thus, the objective was to discover generalizable lessons learned from the recent experience and identify potential factors to consider when having to make administrative decisions under difficult circumstances. Accordingly, 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 achievement. Therefore, in the following discussion of results and conclusions, besides identifying any significant differences between groups, a particular emphasis will also be put on quantifying the effect size that the grouping factor had on student performance.

Lastly, it should also be clearly delimitated 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 both 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.

Results
The findings from this study will be presented in two parts. In this first paper, results from the overall group comparison between remote and hybrid learners will be shown and findings and tentative conclusions will be discussed. A second, follow-on paper will present the breakdown of the analysis by grade levels, further nuancing any generalizable findings from the overall comparison. The Part 2 study will also allow validating any initial conclusions and add a control through an additional comparison to the conventional classroom setting. These added test data from prepandemic years in the second paper will also allow presenting an analysis of student development over time (i.e. results from a time series ANOVA).

For the overall comparison presented in this paper, the ANCOVA on the entire set of 2021 $Math_{Posttest}$ scores ($N_{Total,K-8,Math} = 539$) is depicted in Table 1 and confirmed that the 2020 $Math_{Pretest}$ score covariate was indeed a significant predictor for 2021 performance in math, $F(1, 536) = 2389.7, p < 0.001, r = 0.90$. It also revealed that grouping (H or R) was a significant factor, $F(1, 536) = 25.9, p < 0.001$, with a small to medium effect size (*partial* $\eta^2 = 0.046$; Cohen as cited in Wuensch, 2015; Green and Salkind, 2014). Thus, while roughly 80% of variance in posttest math scores were explainable by variance in pretest scores, about 5%

were related to group association. Further post-hoc testing of this grouping factor indicated that, when accounting for 2020 $Math_{Pretest}$ scores as a covariate, 2021 $Math_{Posttest}$ scores in the R group ($N_{R,K-8,Math}$ = 155, M_R = 470, SE = 1.384) were significantly higher ($t_{R-H}[536]$ = 5.09, p < 0.001) than those in the H group ($N_{H,K-8,Math}$ = 384, M_H = 461, SE = 0.879). Figure 1 visualizes this relationship.

Similarly, ANCOVA testing on the entire set of 2021 $ELA_{Posttest}$ scores ($N_{Total,K-8,ELA}$ = 394) is depicted in Table 2 and confirmed that the 2020 $ELA_{Pretest}$ score covariate was as well a significant predictor for 2021 performance in ELA, $F(1, 391)$ = 742.8, p < 0.001, r = 0.80. It also revealed that grouping was a significant factor, $F(1, 391)$ = 14.6, p < 0.001, with a small to medium effect size (*partial* η^2 = 0.036; Cohen as cited in Wuensch, 2015; Green and Salkind, 2014). Thus while roughly 65% of variance in posttest ELA scores were explainable by variance in pretest scores, about 4% were related to group association. Further post-hoc testing of the grouping factor indicated that, when accounting for 2020 $ELA_{Pretest}$ scores as a covariate, 2021 $ELA_{Posttest}$ scores in the R group ($N_{R,K-8,ELA}$ = 114, M_R = 590, SE = 2.75) were significantly higher ($t_{R-H}[391]$ = 3.82, p < 0.001) than those in the H group ($N_{H,K-8,ELA}$ = 280, M_H = 577, SE = 1.76). Figure 2 visualizes this relationship.

Therefore, both null hypotheses were rejected.

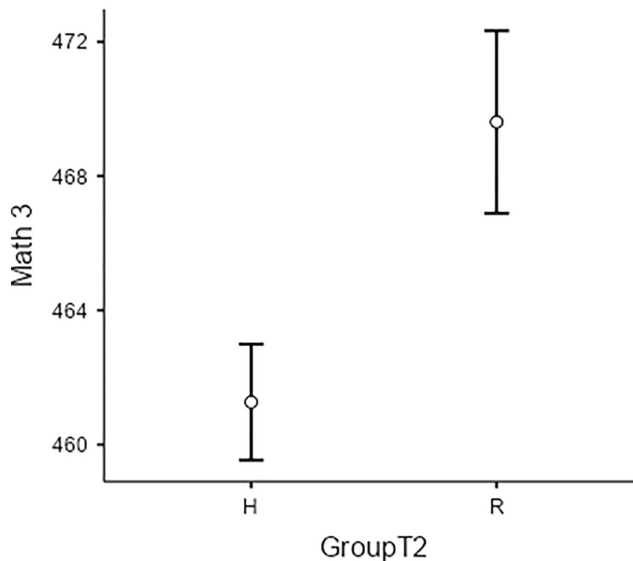| ANCOVA - $Math_{Posttest, K-8}$								
	Sum of squares	df	Mean square	F	p	η^2	$\eta^2 p$	ω^2
Overall model	716,350	2	358,175	1199.5	<0.001			
$Math_{Pretest,K-8}$	708,671	1	708,671	2389.7	<0.001	0.810	0.817	0.809
Group (R, H)	7,679	1	7,679	25.9	<0.001	0.009	0.046	0.008
Residuals	158,954	536	297					
Source(s): Fox and Weisberg (2020), The Jamovi Project (2021)								

Table 1.
Grouping effect on Math scores

Figure 1.
Estimated marginal mean of $Math_{Posttest}$ (2021) scores for the H and R groups when accounting for $Math_{Pretest}$ (2020) scores as a covariate

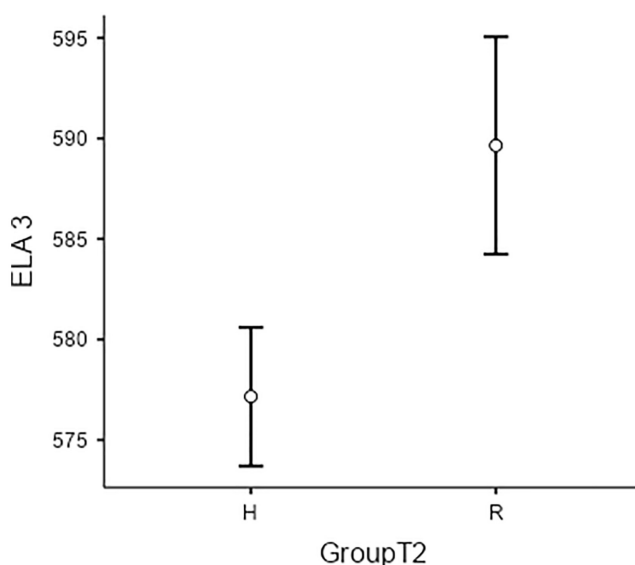**Source(s):** Lenth (2020), The Jamovi Project (2021)

Discussion

Based on the above presented results of the analysis, there was a clear difference between learning conditions notable, with standardized test scores of fully remote students significantly exceeding those of their hybrid-learning peers, in particular, once accounting for any pretest differences in learners. These overall findings seemed to disagree with Chen et al.'s (2021) observations and did not confirm the concerns expressed by other researchers (e.g. Cardullo et al., 2021; Garbe et al., 2020). Furthermore, in contrast to the distinctions found in the literature between verbal and computational learning (Van Tetering et al., 2018), the results did not indicate any differences between the two subject areas (ELA and math) under consideration. That is, the observed difference between remote and hybrid environments was consistently apparent regardless whether the assessed subject area was verbally related (i.e. ELA) or more computationally oriented (i.e. math). The results did, however, seem to agree with Barchas-Lichtenstein et al.,'s 2020 research viewing computations learning (i.e. relying predominantly on the computer in the remote group) as a way to support independent active learning. Nevertheless, the observed differences between the H and R environments also directly raised an important additional question that sparked the follow-on research to this initial overall analysis of group differences, which will be presented in Part 2 of this paper:

ANCOVA–$ELA_{Postest, K-8}$

	Sum of squares	df	Mean square	F	p	η^2	$\eta^2 p$	ω^2
Overall model	653,290	2	326,645	374.9	<0.001			
$ELA_{Pretest,K-8}$	640,673	1	640,673	742.8	<0.001	0.647	0.655	0.645
Group (R, H)	12,616	1	12,616	14.6	<0.001	0.013	0.036	0.012
Residuals	337,220	391	862					
Source(s): Fox and Weisberg (2020), The Jamovi Project (2021)								

Table 2.
Grouping effect on
ELA scores

Source(s): Lenth (2020), The Jamovi Project (2021)

Figure 2.
Estimated marginal
means of $ELA_{Posttest}$
(2021) scores in the H
and R groups with
accounting for
$ELA_{Pretest}$ (2020) scores
as covariate

How did learner performance in the two pandemic-induced groups (R and H) compared to student achievements in prepandemic conditions (i.e. to a classroom control)?

While, from the results, it may seem obvious that the fully remote learners outperformed their hybrid-learning peers, the question is whether they also increased in performance when compared to the traditional classroom conditions (C). Such comparison to a control could also help to further dispel the worries about testing integrity expressed by Cardullo *et al.* (2021). So, it was interesting to further investigate whether the observed differences between H and R were due to the students in group R advancing or due to the students in group H regressing in comparison to the control C. Furthermore, there could also be the possibility that both groups H and R improved or declined in comparison to prepandemic conditions C. Thus, the introduction of a control in the follow-on analysis, which will be presented in the second paper, was vital to further explaining the observed difference between groups. From this first part of the analysis for which the results are shown above, all that can be concluded is that during the pandemic, remote learner (group R) performed significantly better on math and ELA standardized testing.

So, a second important follow-up question that was also already touched on in the discussion is what potential reasons may have existed that could explain the observed differences. For one, as already mentioned, testing integrity in the R group may have been compromised, for example, by receiving inappropriate help at home (Cardullo *et al.*, 2021). Nevertheless, such undue influence on the test results would be expected to benefit weaker students more than strong ones, thus flattening the slope of the regression line for the correlation between pretest and posttest scores in the R group, when compared to the H group during analysis. Such differences in regression slopes between H and R were not seen in the results, making the unauthorized-help explanation unlikely the main cause for the observed differences. Similarly, since rigorous prescreening of the data was conducted to eliminate any outliers and undue influences, it is also unlikely that a few exceptional scores in the R group (which would be another indication for possible testing irregularities) may have skewed the results. However, to completely rule out testing integrity issues as a cause for the observed differences, the already discussed follow-on comparison to a control (as it will be presented in Part 2) may also serve as a confirmation that the differences found in this first part were genuinely attributable to the different environments.

Another possible explanation for the observed group differences in the results is that the groups were not equal to start with. Such differences in group composition are especially plausible since the sample selection and group assignment were not random. In short, students in the remote learning group may have chosen to stay fully remotely because they were already stronger students in the first place. Such self-selection bias could explain higher posttest scores in the R group. However, since the ANCOVA analysis also incorporated students' pretest scores as a covariate, any potentially preexisting differences between groups were accounted for. In fact, from the data (see Appendix, Table A1 and A2 and Figures A1 and A5), it seems average learner pretest scores in the R group were actually slightly lower than in the H group, effectively ruling this form of self-selection bias as an explanation for the observed differences out.

Nonetheless, another form of self-selection bias may have existed and might provide a first glimpse into a group of factors that could have more plausibly influenced results. By giving students (and their parents) the choice to either convert to the hybrid option or continue in the remote learning environment, students that felt comfortable with the remote model (and the support received at home) may have been more likely to stay in this learning environment. Thus, as observed by Lu and Yang (2018) and Van Tetering *et al.* (2018), catering to students' preferred learning styles and circumstances might have positively influenced their learning performance, which may be what is reflected in the analyzed data. This preliminary finding points to a whole group of factors related to the students' environment, their individual

development and their motivations. Aspects such as comfort, familiarity, attention, care and support, as well as required discipline, independence and responsibility, may have been influential to learner performance. Thus, matching the learning environment to individual student needs and expectations may have been more completely achieved in one environment (R) than in the other (H).

Finally and closely related to the above discussion of learner needs, a factor that may have strongly supported the learners in the R group is consistency. Since these students continued for the whole year in their remote learning environment, they experienced a more consistent learning setting than their hybrid peers who went through multiple iterations of changing rules and conditions. Such consistency may have been one of the most influential factors to learning performance. To confirm this tentative theory, our follow-on analysis that will be presented in a Part 2 paper strongly focused on comparing the observations about student development during the pandemic to those under prepandemic conditions. If consistency was a major factor, then it should be expected that remote learner development continued at prepandemic levels, while hybrid learner development did not, which would explain the observed significant differences in this first part of the analysis. Through such confirmation, consistency could be established as a major influence on learner performance, providing teachers and school administrators with an important consideration when reacting to changing circumstances: consistency in any given learning environment may be more important than finding the best environment.

Nevertheless, besides the significance of the difference between groups in these findings, a short discussion of observed effect sizes, in general, seems to be in order here. Based on the observed *partial η^2* in both the math and the ELA comparisons, it can be concluded that the effect of the learning environment was only small to medium. More precisely, approximately 4–5% of observed variance in testing scores could be attributed to the grouping factor (i.e. whether the student learned fully remotely or in a hybrid environment). By far the vast majority of variance in test scores rather seemed to be explainable by variances in students' pretest scores, accounting for roughly 82% in math and 65% in ELA. Thus, while roughly two-thirds of students' performance in ELA and over three-quarters of their performance in Math were predicted by their past performance in each respective subject area, less than 5% could be attributed to the differences in the learning environment. Or in other words, regardless of the learning condition, students seemed to mostly progress as expected along their learning path, which, given the circumstances, is really great news and another major finding of this study. Therefore, for the follow-on research that will be presented in the second paper, it was further interesting to analyze whether a similar predicted growth could also be established for the previous years, in which students encountered normal classroom conditions, and how this expected development from prepandemic conditions held up during the pandemic.

Conclusion
Our findings from this first part of the analysis indicated a significant difference between the two learning environments (fully remote and hybrid) that were employed during the pandemic. Student achievements in the remote group (R) exceeded those in the hybrid one (H) in both subject areas (math and ELA) considered. The research design and analysis mostly eliminated preexisting group differences, selection biases and testing irregularities as potential causes for the observed differences. Therefore, the findings seem to suggest that environmental and developmental factors may have played a major role in student performance in the different learning environments during the pandemic. Our tentative theory from these findings points to consistency and individualized support in the learning environment as major influences on students' academic development. Therefore, our recommendation for further research directly aligns with the continued analysis of the collected data that we conducted and that will be presented in a Part 2 to this study.

This follow-on research aimed to confirm the tentative findings presented here, by adding a comparison to prepandemic classroom conditions to the analysis. It will also provide further detail about the effect that consistency in the learning environment may have had on the differently aged learners in each grade level. Armed with such insights, teachers and school administrators may have a better basis for decision-making when faced with rapidly changing conditions, such as during the pandemic: For the benefit of academic development, they may, for example, elect to keep their younger learners in a consistent environment rather than trying to adapt too quickly to the constantly changing public health guidelines.

References

Allen, I.E. and Seaman, J. (2017), "Digital learning compass: distance education enrollment report 2017", available at: https://files.eric.ed.gov/fulltext/ED580868.pdf.

Barchas-Lichtenstein, J., Brucker, J.L., Nock, K., Gupta, R. and Flinner, K. (2020), *Education in the Pandemic and the Potential for Computational Thinking*, Knology Publication, available at: https://www.datocms-assets.com/15254/1601924291-ct-pandemic-white-paperinfact2020-10-05.pdf.

Ben-Shachar, M. and Lüdecke (2020), "Compute and interpret indices of effect size", available at: https://easystats.github.io/effectsize/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].

Cardullo, V., Wang, C., Burton, M. and Dong, J. (2021), "K-12 teachers' remote teaching self-efficacy during the pandemic", *Journal of Research in Innovative Teaching and Learning*, Vol. 14 No. 1, pp. 32-45, doi: 10.1108/JRIT-10-2020-0055.

Chen, L.-K., Dorn, E., Sarakatsannis, J. and Wiesinger, A. (2021), *Teacher Survey: Learning Loss Is Global — and Significant*, McKinsey and Company, available at: https://www.mckinsey.com/~/media/mckinsey/industries/public%20and%20social%20sector/our%20insights/teacher%20survey%20learning%20loss%20is%20global%20and%20significant/teacher-survey-learning-loss-is-global-and-significant.pdf?shouldIndex=false.

Curriculum Associates (2021), "iReady", available at: https://www.curriculumassociates.com/products/i-ready.

Domina, T., Renzulli, L., Murray, B., Garza, A.N. and Perez, L. (2021), "Remote or removed: predicting successful engagement with online learning during COVID-19", *Socius: Sociological Research for a Dynamic World*, Vol. 7, pp. 1-15, doi: 10.1177/2378023120988200.

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].

Falloon, G. (2016), "An analysis of young students' thinking when completing basic coding tasks using Scratch Jnr. On the iPad", *Journal of Computer Assisted Learning*, Vol. 32 No. 6, pp. 576-593, doi: 10.1111/jcal.12155.

Ferri, F., Grifoni, P. and Guzzo, T. (2020), "Online learning and emergency remote teaching: opportunities and challenges in emergency situations", *Societies*, Vol. 10 No. 86, doi: 10.3390/soc10040086.

Field, A. (2009), *Discovering Statistics Using SPSS*, 3rd ed., Sage Publication, London.

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/.

Garbe, A., Ogurlu, U., Logan, N. and Cook, P. (2020), "COVID-19 and remote learning: experiences of parents with children during the pandemic", *American Journal of Qualitative Research*, Vol. 4 No. 3, pp. 45-65, doi: 10.29333/ajqr/8471.

Green, S.B. and Salkind, N.J. (2014), *Using SPSS for Windows and Macintosh: Analyzing and Understanding Data*, 7th ed., Pearson, Upper Saddle River, NJ.

Grover, S. (2017), "Assessing algorithmic and computational thinking in K-12: lessons from a middle school classroom", in Rich, P. and Hodges, C. (Eds), *Emerging Research, Practice, and Policy on Computational Thinking. Educational Communications and Technology: Issues and Innovations*, Springer, Cham. doi: 10.1007/978-3-319-52691-1_17.

Grover, S., Pea, R. and Cooper, S. (2016), "Factors influencing computer science learning in middle school", *Proceedings of the 47th ACM Technical Symposium on Computing Science Education – SIGCSE*, Vol. 16, pp. 552-557, doi: 10.1145/2839509.2844564.

Horowitz, J. and Igielnik, R. (2020), "Most parents of K-12 students learning online worry about them falling behind", *Pew Research Center*, available at: https://www.pewresearch.org/social-trends/wp-content/uploads/sites/3/2020/10/PSDT_10.29.20_kids.edu_.full_.pdf.

Kleinke, S. (2020), *Cheating Risk in Fully Online Course Activities: A Quantitative Quasi-Experimental Factorial Design Study*, (Publication No. 28256665) [Doctoral dissertation, Northcentral University, La Jolla, CA], ProQuest Dissertation and Thesis Global.

Lenth, R. (2020), "Emmeans: estimated marginal means, aka least-squares means", [R package], available at: https://cran.r-project.org/package=emmeans.

Lu, T. and Yang, X. (2018), "Effects of the visual/verbal learning style on concentration and achievement in mobile learning", *Eurasia Journal of Mathematics, Science and Technology Education*, Vol. 14 No. 5, pp. 1719-1729, doi: 10.29333/ejmste/85110.

R Core Team (2021), "R: a Language and environment for statistical computing", (Version 4.0) [Computer software], (R packages retrieved from MRAN snapshot 2021-04-01), available at: https://cran.r-project.org.

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, doi: 10.15354/bece.20.or027.

The Jamovi Project (2021), "Jamovi", (Version 1.8) [Computer Software], available at: https://www.jamovi.org.

Van Tetering, M.A.J., de Groot, R.H.M. and Jolles, J. (2018), "Boy–girl differences in pictorial verbal learning in students aged 8-12 years and the influence of parental education", *Frontiers in Psychology*, Vol. 9, doi: 10.3389/fpsyg.2018.01380.

Wuensch, K. (2015), "Cohen's conventions for small, medium, and large effects", available at: http://core.ecu.edu/psyc/wuenschk/docs30/EffectSizeConventions.pdf.

Appendix

	Group T2	Math 2	Math 3
N	H	334	384
	R	155	155
Missing	H	0	0
	R	0	0
Mean	H	445	452
	R	441	467
Std. error mean	H	2.25	2.02
	R	3.55	3.36
95% CI mean lower bound	H	441	458
	R	434	461
95% CI mean upper bound	H	449	456
	R	448	474
Standard deviation	H	44.0	39.5
	R	44.3	41.8
Minimum	H	336	325
	R	341	382
Maximum	H	545	591
	R	532	596

Table A1.
Descriptive statistics for $Math_{Pretest}$ (2020; Math2) and $Math_{Posttest}$ (2021, Math3) scores

Source(s): The Jamovi Project (2021)
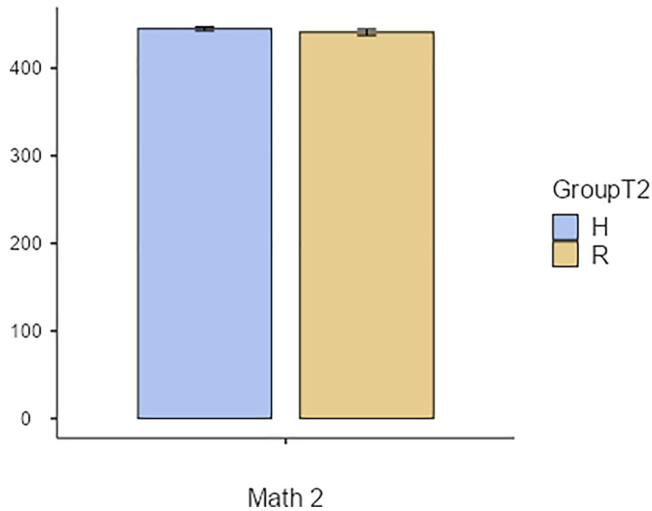
Figure A1.
Bar graph of $Math_{Pretest}$ (2020; Math2) scores

Math 2

Source(s): The Jamovi Project (2021)

Figure A2.
Histogram of
$Math_{Pretest}$ (2020;
Math2) scores
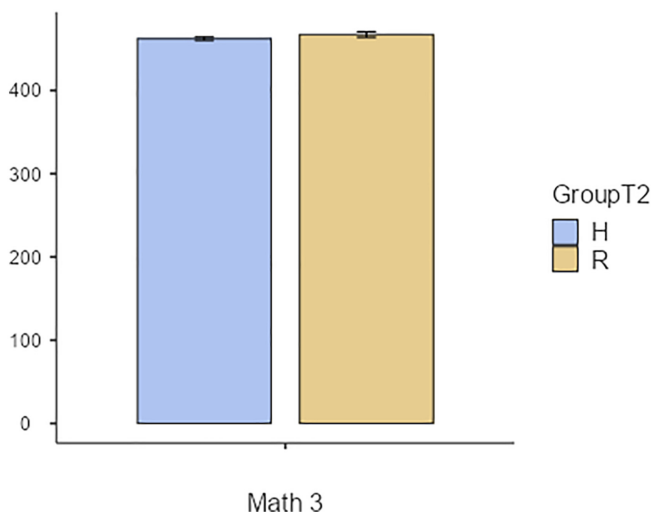
Source(s): The Jamovi Project (2021)

Figure A3.
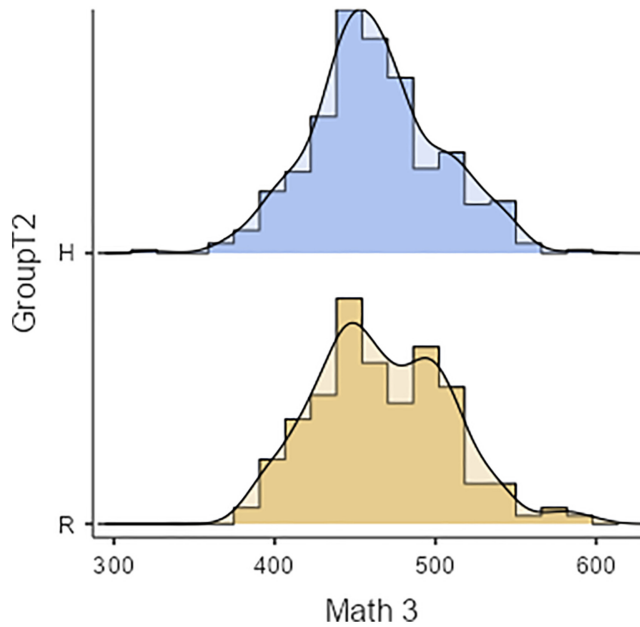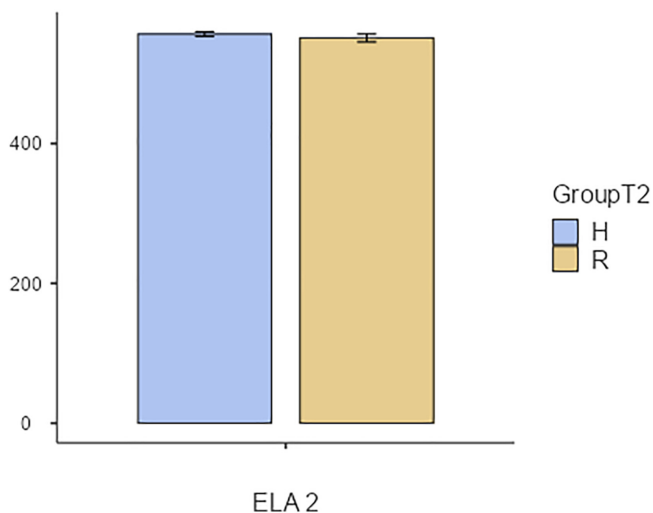Bar graph of
$Math_{Postest}$ (2020;
Math3) scores

Source(s): The Jamovi Project (2021)

Figure A4.
Histogram of
$Math_{Postest}$ (2020;
Math3) Scores

Source(s): The Jamovi Project (2021)

	GroupT2	ELA 2	ELA 3
N	H	280	280
	R	114	114
Missing	H	0	0
	R	0	0
Mean	H	557	578
	R	551	587
Std. error mean	H	3.36	3.00
	R	5.66	4.63
95% CI mean lower bound	H	550	572
	R	540	578
95%CI mean upper bound	H	563	584
	R	562	596
Standard deviation	H	56.2	50.2
	R	60.4	49.4
Minimum	H	372	409
	R	585	450
Maximum	H	830	706
	R	674	686

Table A2.
Descriptive Statistics
for $ELA_{Pretest}$ (2020;
ELA2) and $ELA_{Posttest}$
(2021, ELA3) scores

Source(s): The Jamovi Project (2021)
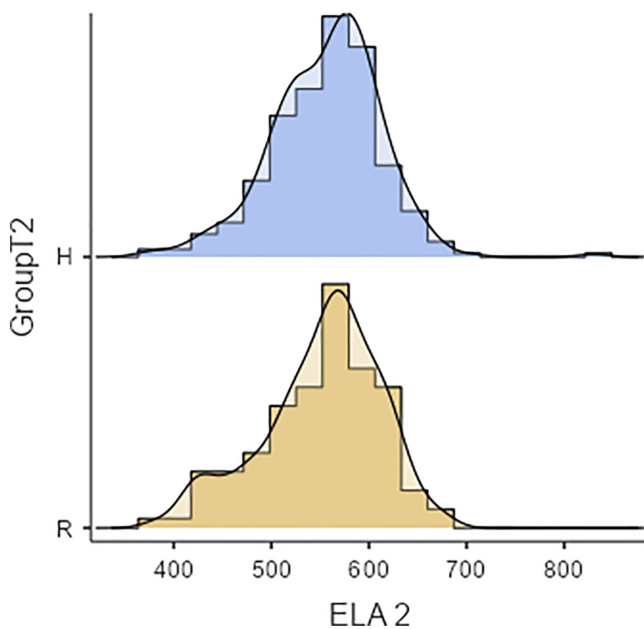
Source(s): The Jamovi Project (2021)

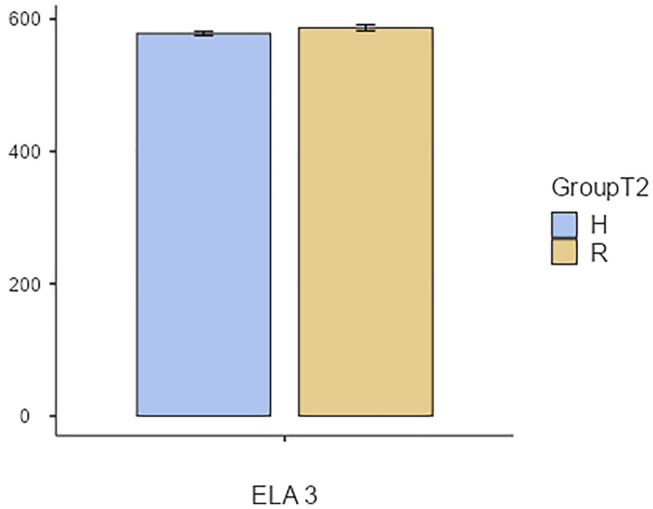**Source(s):** The Jamovi Project (2021)

Figure A7.
Bar Graph and
Histogram of
$ELA_{Postest}$ (2020; ELA3)
Scores

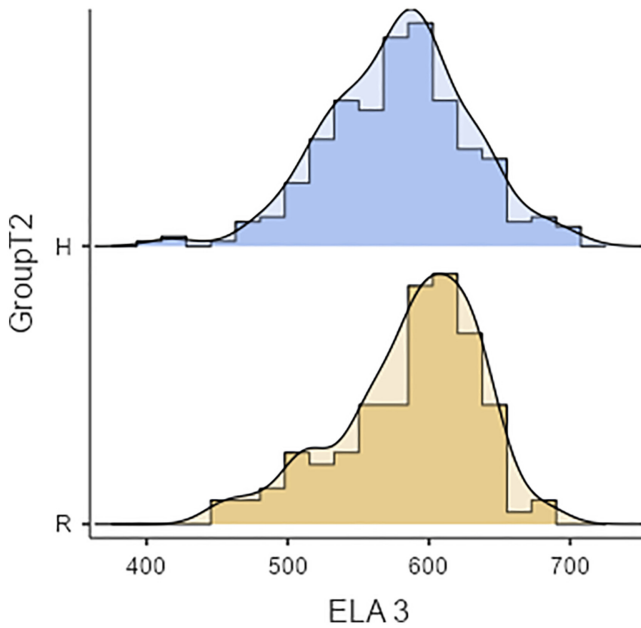**Source(s):** The Jamovi Project (2021)

Figure A8.
Histogram of
$ELA_{Postest}$ (2020; ELA3)
Scores

Source(s): The Jamovi Project (2021)

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