frontiers | Frontiers in Education

TYPE Brief Research Report
PUBLISHED 01 August 2022
DOI 10.3389/feduc.2022.928987

OPEN ACCESS

EDITED BY
Eva Hartell,
Royal Institute of Technology, Sweden

REVIEWED BY
Ángel Freddy Rodríguez Torres,
Central University of Ecuador, Ecuador
Julio Ciro Benegas,
National University of San Luis,
Argentina

*CORRESPONDENCE
Alina Behrendt
alina.behrendt@uni-due.de

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to STEM Education, a section of the journal Frontiers in Education

RECEIVED 26 April 2022
ACCEPTED 12 July 2022
PUBLISHED 01 August 2022

CITATION
Behrendt A, Fischer V and Walpuski M (2022) COVID-19 school closures and chemistry-related competencies: A study of German students transitioning from primary to secondary school. *Front. Educ.* 7:928987. doi: 10.3389/feduc.2022.928987

COVID-19 school closures and chemistry-related competencies: A study of German students transitioning from primary to secondary school

Alina Behrendt*†, Vanessa Fischer† and Maik Walpuski†

Department of Chemistry Education, University of Duisburg-Essen, Essen, Germany

The COVID-19 pandemic led to temporary closures of schools around the world, resulting in a change from face-to-face teaching to distance teaching, which had been practiced minimally until then. In this study, we investigated the effects of pandemic-related school closures on students' chemistry-related competencies, at the transition from primary to secondary school. We also explored the extent to which at-home or in-school data collection influenced the results. We measured the competencies of 2,262 students from grades 5 to 9 in Germany. Data collection took place before, during, and after the pandemic-related school closures, based on test booklets completed by students. The results showed that the competencies of students in Chemistry, who were taught in school before the closures, were similar to those of students who were taught *via* distance learning. Thus, students' competencies were similar before and after the school closures. The school closures led to differences not only in teaching, but also in the way the data in this study was collected. During school closures, students worked on their test booklets at home, and before and after school closures, the data were collected at school. This also enabled us to examine the effects of the different data collection designs on the test scores. We found differences between the results of the test booklets completed at home, and those completed at school, only for younger students. For students in higher grades, there were no differences.

KEYWORDS

COVID-19, chemistry-related competencies, cumulative learning, transition, primary school, secondary school

Introduction

For successfully learning science, a cumulative learning process is necessary. Cumulative learning refers to integrating new content into the knowledge acquired earlier by students (Lee, 2012). Hence, competencies acquired in primary school should be aligned with the new, more demanding requirements of secondary school (Hempel, 2010). Thus, learning should be cumulative, not only within each type of school, but also across different types of schools. This is especially important since the home and school environment shows a major impact on cumulative learning (Oludare and Alade, 2018). For cumulative learning to succeed at the transition between primary and secondary education, the curricula of various subjects play a crucial role. In these curricula, ideas that build on each other must be clearly defined, so that the development of competencies can be adequately supported over a long period of time (Shin et al., 2017).

The level to which these competencies, expected in different types of schools, are actually aligned, differs between different countries and the science-related subjects taught there. Some countries, such as the United States, teach one integrated subject, science, as shown by the Benchmarks for Science Literacy (American Association for the Advancement of Science, 2001, 2013), and the Next Generation Science Standards (National Research Council, 2013). The competencies formulated in the Benchmarks for Science Literacy span from kindergarten to grade 12, and build on each other (American Association for the Advancement of Science, 2001). Other countries do not provide curricula for one subject for science across all school levels, because science is divided into separate disciplines. In most European countries, for example, a general subject — science, is taught in primary school. This is replaced by the separate subjects of biology, physics, and chemistry in secondary school. Only a few European countries (Italy, Luxembourg, Iceland, Norway, Turkey, and parts of Belgium) teach science throughout lower secondary school as a single subject (Education, Audiovisual and Culture Executive Agency, 2011).

To assess whether scientific competencies can be developed cumulatively at the transition, despite the different subjects taught, the curricula of the various science subjects in the transition period must be compared. In Germany, the expected competencies in chemistry are of special interest for the transition between primary and secondary education. This study focuses on the situation in North Rhine-Westphalia, on behalf of all other German states. In particular, chemistry as an individual subject, only starts in grade 7 or 8 (age: 12/13 years), but is supposed to enhance the existing competencies of students from primary school ending in grade 4 (age: 10 years). The curricula in both primary and secondary school in Germany address competencies in the context of the following content: combustion, states of aggregate, substance properties, solutions, and energy (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2008, 2013; Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2021). This provides a good foundation for a cumulative learning process, because the competencies in both curricula build on each other. Nevertheless, the transition from primary to secondary school is not always successful. Various international school comparison studies show that there are differences between German students, with regard to their expected and acquired competencies. While 72 % of German 4TH graders achieve an intermediate or high competence level in science, according to the Trends in International Mathematics and Science Study 2019 (Mullis et al., 2020), the Program for International Student Assessment 2018 reveals that only 58 % of German 9TH graders achieve an intermediate or high competence level in science (Organization for Economic Cooperation and Development, 2019). Consequently, the fact that the structure of the competencies of the curricula can enable cumulative learning processes does not mean that this learning process actually occurs. Studies describing the development of science literacy across grades have been undertaken in various countries, such as the United States and the United Kingdom (American Association for the Advancement of Science, 2001; Wiser et al., 2012; Waldo, 2014). For Germany, however, such studies exist only for specific contents, such as magnetism (Möller, 2016), or they refer to specific periods within teaching chemistry, such as the strand maps with learning progressions developed by Celik and Walpuski (2018) for the first years of learning chemistry. Chemistry-related competencies, at the transition between primary and secondary school, still need further study.

During of the COVID-19 pandemic, schools were closed for several weeks, or even months, in many countries around the world, and students had to engage in distance learning. This change in teaching impacted students' learning and performance in different subjects. Hammerstein et al. (2021) summarized the effects of school closures on student achievements, in a review of several independent studies. School closures during the pandemic were found to have a predominantly negative impact on students' achievement. These findings were particularly evident for younger students, and those from families with low socioeconomic status (Hammerstein et al., 2021). For example, Tomasik et al. (2021) investigated students' achievement in mathematics and German, in primary and secondary schools. They found that the learning progress in the period before the school closures in primary school was more than twice of that during the school closures. On the contrary, they could not find this difference at the secondary level (Tomasik et al., 2021).

10.3389/feduc.2022.928987

Most of the studies summarized by Hammerstein et al. (2021) refer to students' achievement in mathematics or different languages, but Maldonado and De Witte (2020) additionally focused on science. Using standardized tests that were administered at the end of primary school in Switzerland, they showed that the test scores decreased significantly from 2019 to 2020 (Maldonado and De Witte, 2020). This result provides preliminary evidence that school closures during the COVID-19 pandemic affected scientific skills at the transition from primary to secondary education negatively; however, chemistry-specific data are missing.

The purpose of the current study was to examine students' chemistry-related competencies, at the transition between primary and secondary education, in general. However, the study was expanded amid the pandemic, to include the influences of COVID-19-related school closures. For this purpose, we collected data before, during, and after the school closures. During school closures, we could not collect the data in a face-to-face teaching environment. Instead, students completed the test booklets at home, for the assessment. The two following research questions were addressed in this study:

1. How do students' chemistry-related competencies differ at the transition between primary and secondary education, before and after COVID-19-related school closures?

2. To what extent does the design of data collection (at home or face-to-face teaching) influence measurement results?

Method

Participants and procedure

To answer the two research questions, we used a sample of 2,262 students from Germany. Parents were informed about the study by a letter to parents, and the respective school principals consented to the data collection. Since the data was collected pseudonymously, no further consent was required. In Germany, one integrated subject comprising all social and natural sciences is taught in primary school (grades 1–4; age: 6–10 years). Chemistry contents are also included in this subject. Subsequently, either biology and physics or one subject including all natural sciences are taught in secondary school from grade 5 onward. As described in the introduction, chemistry as an independent subject starts in grade 7 or 8 (age: 12/13 years). Hence, the transition from primary to secondary education, with respect to the subject of chemistry, covers the period from grades 5–7, or 8. We measured the competencies of students in grades 5–9 to investigate this transition period and their first year in chemistry.

The grade in which chemistry is taught for the first time differs between the participating schools. Therefore, we do not refer to grades, but to the following measurement points within the transition period:

Start of the **t**ransition period (ST).
Middle of the **t**ransition period **1** (MT1).
Middle of the **t**ransition period **2** (MT2).
Start of the first year in **c**hemistry (SC).
End of the first year in **c**hemistry (EC).

The measurement point MT2 exists only in schools that do not teach chemistry in grade 7, and whose transition period, therefore, comprises 3 years. **Figure 1** provides an overview of all assessments and all measurement points of data collection. The first assessment took place before the COVID-19 pandemic, in summer 2019. At that time, we had planned two assessments for each subsample, resulting in a quasi-longitudinal study comprising the beginning of secondary education until the end of the first year of chemistry education. For this reason, the first assessment comprised the measurement points ST, MT2, and SC, while the second assessment covered the measurement points MT1, SC, and EC. In the second assessment, we consulted the same sample as in the first assessment one year later for a longitudinal comparison within the transition period. The second assessment took place during the COVID-19-related school closures, in summer 2020. To obtain an overall larger sample, and more opportunities for comparison, especially regarding the closure effects, we added a third assessment after the school closures in summer 2021. In order to merge the additional data, with the data collected from the first two assessments, it was important to measure the same age groups at the measurement points. The additional, third assessment therefore includes all 5 measurement points of the first two assessments.

We analyzed different subsamples to answer the two research questions. To compare competencies before and after the school closures, we were able to use data from the first and the third assessment. In both cases, data were collected in a face-to-face teaching environment, and therefore, are comparable. Data regarding measurement points ST, MT2, and SC were available for both assessments, so we were able to use the entire sample (990 students) from the first assessment and a subsample (419 students) from the third assessment. To compare the effects of the different assessment situations (school and home), we analyzed data from the second and the third assessment. Both assessments took place after a long period of school closures, and therefore, differed with regard to the assessment situation despite their similarities. For this comparison, data regarding measurement points MT1, SC, and EC were available, so we were able to use the entire sample (496 students) from the second assessment, and a subsample (414 students) from the third assessment.

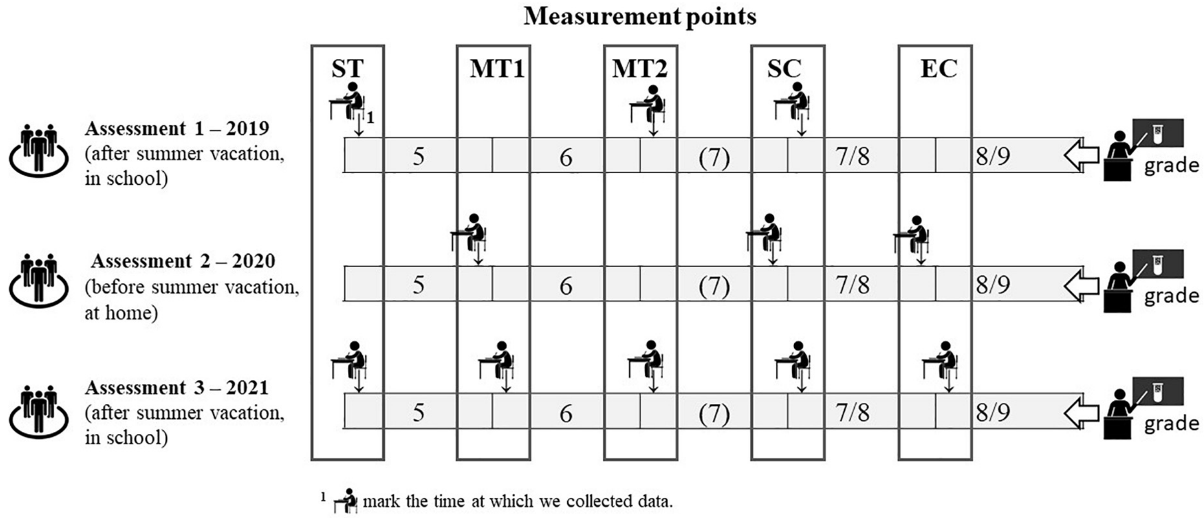

FIGURE 1
Assessments and measurement points.

Instruments

To measure students' chemistry-related competencies, we developed a new test instrument. In the first step, we identified the competencies to be measured, based on the curricula for the subjects containing chemistry-related contents in primary and secondary school. We used these competencies for developing a paper-pencil test with multiple-choice items. Each item contains six answer options. For each option, learners had to decide whether this answer was correct or incorrect, or whether they were unsure about it. Thus, one point could be scored for each answer option, and zero to six points could be scored for each item. Based on the competencies addressed in the chemistry curriculum in Germany, 24 items were developed for content knowledge (CK) with each 8 items each for the key concepts of chemical reactions, structure of matter, and energy, and for procedural knowledge (PK) with the categories of scientific inquiry, communication, and decision-making. To reduce the test time, we created test booklets with either 20 or 32 items, depending on the grade level. All test booklets contained items based on the competencies of the primary school curriculum. In the test booklets for students in the first year of learning chemistry, there were additional items based on the competencies of the chemistry curriculum. The different test booklet versions were linked by a balanced incomplete block design. Each test booklet contained items for two key concepts from the category CK, and for two categories of PK.

To validate the test instrument, we carried out an expert rating in which seven raters were asked to assign the developed items to the categories CK, scientific inquiry, communication, and decision-making. The interrater reliability was $\kappa_{Fleiss} = 0.795$. After a subsequent revision of some items, we administered the test instrument to 760 students (grades 4–8) in a pilot study. To investigate the quality of the test instrument, we conducted separate item response theory (IRT) analyses for CK and PK. Due to the item format, with zero to six achievable points each, we used the rating scale model. We found satisfactory reliabilities (person reliability$_{CK}$ = 0.75; item reliability$_{CK}$ = 0.99; person reliability$_{PK}$ = 0.81; item reliability$_{PK}$ = 0.99). For this reason, we only improved the items in which we identified comprehension difficulties during data collection.

In addition to the newly developed test instrument, we used existing test instruments to assess additional skills of the students as control variables. First, we used the Kognitiver Fähigkeitstest (KFT) (Heller and Perleth, 2000) to measure cognitive skills, and second, the Lesegeschwindigkeits- und Verständnistest (LGVT) (Schneider et al., 2017) to measure reading comprehension.

Data processing (main study)

To evaluate the data, we performed separate IRT analyses for CK and PK, using the rating-scale model, as in the pilot study. We used the estimated person parameters, to compare different subsamples. To analyze whether statistically significant differences exist between groups, we performed either a *t*-test for independent samples, or an analysis of covariance (ANCOVA). The ANCOVA allowed us to include the control variables as covariates,

TABLE 1 Reliabilities, infit, and discrimination estimated in the rating scale model.

	Total	MP 1	MP 2	MP 3
Content knowledge				
Person reliability	0.79	0.74	0.86	0.75
Item reliability	1.00	1.00	1.00	0.99
Infit	M = 1.08 SD = 0.24 *Min*: 0.71 *Max*: 1.55	M = 1.07 SD = 0.24 *Min*: 0.70 *Max*: 1.52	M = 1.08 SD = 0.28 *Min*: 0.75 *Max*: 1.73	M = 1.10 SD = 0.26 *Min*: 0.64 *Max*: 1.58
Discrimination	*Min*: 0.56 *Max*: 1.34	*Min*: 0.67 *Max*: 1.39	*Min*: 0.19 *Max*: 1.35	*Min*: 0.43 *Max*: 1.44
Procedural knowledge				
Person reliability	0.82	0.80	0.89	0.75
Item reliability	1.00	0.99	0.99	0.99
Infit	M = 1.03 SD = 0.21 *Min*: 0.66 *Max*: 1.54	M = 1.02 SD = 0.23 *Min*: 0.65 *Max*: 1.62	M = 1.04 SD = 0.21 *Min*: 0.66 *Max*: 1.58	M = 1.04 SD = 0.20 *Min*: 0.67 *Max*: 1.41
Discrimination	*Min*: 0.37 *Max*: 1.24	*Min*: 0.28 *Max*: 1.29	*Min*: 0.33 *Max*: 1.32	*Min*: 0.48 *Max*: 1.24

TABLE 2 Comparison of person parameters (measurement points 1 and 3) t-test for independent samples.

Group	Sample size	Mean value	Levene-test	t-test
Content knowledge				
Total	n_1 = 972 n_3 = 391	M_1 = −0.25 M_3 = −0.19	p = 0.543	$t(1361)$ = 2.10, p = 0.036, d = 0.126
ST	n_1 = 458 n_3 = 171	M_1 = −0.30 M_3 = −0.37	p = 0.596	$t(627)$ = 1.51, p = 0.132, d = 0.135
MT2	n_1 = 102 n_3 = 142	M_1 = −0.18 M_3 = 0.02	p = 0.730	$t(242)$ = 3.05, p = 0.003, d = 0.396
SC	n_1 = 412 n_3 = 78	M_1 = −0.22 M_3 = −0.15	p = 0.845	$t(488)$ = 0.99, p = 0.324, d = 0.122
Procedural knowledge				
Total	n_1 = 969 n_3 = 391	M_1 = 0.12 M_3 = 0.16	p = 0.318	$t(1358)$ = 0.91, p = 0.365, d = 0.054
ST	n_1 = 457 n_3 = 171	M_1 = 0.04 M_3 = −0.03	p = 0.338	$t(626)$ = 1.27, p = 0.206, d = 0.114
MT2	n_1 = 102 n_3 = 142	M_1 = 0.21 M_3 = 0.38	p = 0.471	$t(242)$ = 2.10, p = 0.037, d = 0.273
SC	n_1 = 410 n_3 = 78	M_1 = 0.20 M_3 = 0.18	p = 0.507	$t(486)$ = 0.29, p = 0.773, d = 0.036

in case the groups differed regarding to those control variables. We used the raw scores of the LGVT, which were identically scaled for all grades, and estimated person parameters for the KFT using IRT analyses, since the items on this test varied across grades, and the raw scores were not comparable.

Results

To check the quality of the test instrument, we first examined the statistical parameters determined in the rating-scale model. Table 1 provides an overview of the statistical parameters. We found good person and item reliabilities for both, CK and PK. In both cases, the person reliability was higher for the second assessment, compared to the first and third assessments. This may indicate that students worked more conscientiously on the test booklets at home, than at school. The infit values and discriminations were almost all satisfactory. Overall, we were able to achieve a good quality of the test instrument. Consequently, we used the person parameters for CK and PK estimated in the rating-scale model across all three assessments, to compare different subsamples in terms of the research questions.

Student performance before and after the school closures

We compared students' competencies before and after the school closures. To do so, we used data from the first and the third assessments. Data from three measurement points (ST, MT2, and SC) were available for the comparison. First, we investigated the comparability of the cohorts from both assessments. We found no significant differences between the two cohorts, in terms of the control variables cognitive skills [$t(1398)$ = 1.59, p = 0.113, d = 0.093] and reading comprehension [$t(1376)$ = 1.04, p = 0.298, d = 0.062]. For this reason, we did not include covariates, but conducted t-tests for independent samples when comparing the person parameters for CK and PK.

We compared the person parameters for both assessments, in general as well as for the subsamples at each measurement point. Table 2 provides an overview of the results of all t-tests for independent samples. For the total sample, we found a significant difference, with a small effect in CK, with lower person parameters before school closures than after them. By contrast, there was no significant difference in PK. At the measurement point ST, there were no significant differences in CK or in PK. At measurement point MT2, by contrast, we were able to find a highly significant difference with a small effect in CK, and a significant difference with a small effect in PK with lower person parameters before the school closures. At measurement point SC, again, we could not find any significant differences in either CK or PK.

Overall, this finding is contrary to expectations. It is not plausible that the competencies of those who had not previously received any face-to-face science instruction were higher only in the middle of the transition period in grade 7, while no difference could be found either at the beginning or at the end of the transition phase. For this reason, we examined person parameters in more detail, at measurement point MT2. We found an anomaly while comparing the mean person parameters of the classes involved in the data collection. In the third assessment, the three classes that achieved the highest mean person parameters were three classes from the same school. The school could not provide any classes from that grade level for data collection in the first assessment. It is likely that these three classes caused the differences between the two assessments, at measurement point MT2. For this reason, we conducted an additional t-test for independent samples for CK and PK, excluding the three classes from the sample. In these t-tests, there were no significant differences in either CK [$t(173) = 0.50$, $p = 0.618$, $d = 0.077$] or PK [$t(173) = 0.07$, $p = 0.942$, $d = 0.011$]. Overall, we concluded that chemistry-related competencies did not differ from each other, at any of the three measurement points. Thus, students' competencies were similar before and after the COVID-19-pandemic school closures. We also showed that the difference in competencies, between the beginning and end of the transition period, was similar before and after the school closures. T-tests showed that for both the first [CK: $t(868) = 2.25$, $p = 0.025$, $d = 0.153$; PC: $t(835.97) = 3.85$, $p < 0.001$, $d = 0.263$] and the third assessments [CK: $t(247) = 3.13$, $p = 0.002$, $d = 0.428$; PC: $t(247) = 2.85$, $p = 0.005$, $d = 0.390$], person parameters were significantly higher at the end of the transition period, than at the beginning of the transition period. Person parameters differed by approximately 0.1–0.2 logits, in both CK and PK, on both assessments.

A possible explanation for these findings is that the quality of face-to-face and distance teaching was comparable, and that students who engaged in face-to-face learning before the pandemic acquired the same competencies as students engaged in distance learning. However, this explanation would contradict Hammerstein et al.'s (2021) finding, that school closures during the COVID-19-pandemic had a negative impact on school performance, particularly for younger students. Another possible explanation is that the measured chemistry-related competencies were largely not acquired in the classroom. It is possible that these were competencies the students acquired outside the classroom, for instance, in a family context. Kähler et al. (2020) showed that some characteristics of the home environment have influences on the science competencies of kindergarten-aged children in Germany. It would be conceivable that science competencies of young students were influenced not only by school, but also by the family and home environments.

Effects of the different assessment situations

In a further step, we compared the students' person parameters from the second and third assessments, to identify the effects of the different assessment situations. In the second assessment, the students processed their test booklets at home after a long period of school closures. At the third assessment, they completed the test booklets at school, but also after a long period of school closures. We compared both, the total sample from the two assessments, as well as the different measurement points, MT1, SC, and EC, separately. Again, we first examined the extent to which the two cohorts were samples with similar characteristics. We did not consider the control variable reading comprehension, because we could observe from the completely filled in LGVT test booklets that many students had worked on it longer than allowed. As the LGVT is a speed test, we did not compare the LGVT results from either of the assessments. However, we compared the control variable cognitive skills between the assessments. It showed a highly significant difference [$t(784) = 2.87$, $p = 0.004$, $d = 0.205$]. We found this difference for the subsample at measurement point MT2 [$t(261.52) = 2.60$, $p = 0.010$, $d = 0.304$]. For the measurement points SC [$t(143) = 1.52$, $p = 0.131$, $d = 0.271$] and SE [$t(312.59) = 0.20$, $p = 0.841$, $d = 0.022$], we found no significant differences between the two assessments. Consequently, we included the control variable cognitive skills as a covariate, in an ANCOVA for the total sample and for the measurement point MT2 in the comparisons

TABLE 3 Comparison of person parameters (measurement points 2 and 3) t-test for independent samples and ANCOVA.

Group	Sample size	Mean value	Levene-test	t-test/ANCOVA
Content knowledge				
Total	$n_2 = 495$ $n_3 = 345$	$M_2 = 0.09$ $M_3 = 0.01$	$p = 0.517$	$t(838) = 1.98$, $p = 0.048$, $d = 0.139$
	$n_2 = 381$ $n_3 = 339$	$M_2 = 0.09$ $M_3 = 0.01$	$p = 0.909$	$F(1, 717) = 4.05$, $p = 0.045$, partial $\eta^2 = 0.006$
MT1	$n_2 = 223$ $n_3 = 131$	$M_2 = 0.02$ $M_3 = -0.20$	$p = 0.344$	$t(352) = 3.60$, $p < 0.001$, $d = 0.396$
	$n_2 = 176$ $n_3 = 126$	$M_2 = 0.03$ $M_3 = -0.20$	$p = 0.565$	$F(1, 299) = 12.10$, $p = 0.001$, partial $\eta^2 = 0.039$
SC	$n_2 = 48$ $n_3 = 78$	$M_2 = -0.07$ $M_3 = -0.15$	$p = 0.442$	$t(124) = 0.68$, $p = 0.498$, $d = 0.125$
EC	$n_2 = 224$ $n_3 = 136$	$M_2 = 0.19$ $M_3 = 0.30$	$p = 0.661$	$t(358) = 1.65$, $p = 1.00$, $d = 0.179$
Procedural knowledge				
Total	$n_2 = 490$ $n_3 = 345$	$M_2 = 0.43$ $M_3 = 0.28$	$p = 0.001$	$t(832.999) = 3.55$, $p = 0.001$, $d = 0.235$
	$n_2 = 378$ $n_3 = 339$	$M_2 = 0.48$ $M_3 = 0.28$	$p < 0.001$	$F(1, 714) = 24.46$, $p < 0.001$, partial $\eta^2 = 0.033$
MT1	$n_2 = 222$ $n_3 = 131$	$M_2 = 0.41$ $M_3 = 0.23$	$p = 0.208$	$t(351) = 2.92$, $p = 0.004$, $d = 0.322$
	$n_2 = 176$ $n_3 = 126$	$M_2 = 0.42$ $M_3 = 0.24$	$p = 0.113$	$F(1, 299) = 8.96$, $p = 0.003$, partial $\eta^2 = 0.029$
SC	$n_2 = 47$ $n_3 = 78$	$M_2 = 0.38$ $M_3 = 0.18$	$p = 0.999$	$t(123) = 1.73$, $p = 0.086$, $d = 0.320$
EC	$n_2 = 221$ $n_3 = 136$	$M_2 = 0.46$ $M_3 = 0.38$	$p < 0.001$	$t(348.732) = 1.22$, $p = 0.224$, $d = 0.116$

of the person parameters in CK and PK. Not all students completed the KFT at home, reducing the sample size of the ANCOVA by 120 cases in CK, and by 118 cases in PK. For this reason, we additionally conducted a *t*-test for independent samples.

Table 3 provides an overview of the results. For the total sample, we found differences in both the *t*-test for independent samples, and the ANCOVA in CK and PK. In both cases, the person parameters were higher for the second assessment compared to the third assessment. We also found this difference for the subsample of the measurement point MT2. For the measurement points SC and EC, we did not find a significant difference in the *t*-tests performed, for either CK or PK. In summary, for younger students, we measured higher competencies with the test booklets worked on at home, rather than with the test booklets worked on at school. We did not find these differences for older students, in their first year of learning chemistry.

A possible reason for these findings is that the students might not have adhered to the instructions, during the data collection at home. It is possible that they disregarded the specified time windows, as they did in the LGVT, and also other rules, such as working on the test items without outside help. As a result, they might have had advantages over students whose data were collected through a face-to-face teaching environment. It is also possible that the younger students were supervised more closely by their parents during the tasks at home and therefore, performed better, while the older students mainly worked by themselves. In contrast to these rather negative effects of working on the test booklets at home, we also have to consider that we found higher person reliabilities, in the rating-scale model for the items worked on at home. This suggests that the students worked more conscientiously at home than at school. Therefore, we assumed that the conditions prevalent at home also had a positive effect on the accuracy of the test results. It is possible that the students were less distracted there than in a data collection at school, could concentrate better on the processing of the items, and consequently, achieved better results.

Discussion

Conclusion

First, the results of the study show that chemistry-related competencies of students at the transition, before the COVID-19 pandemic were similar to those after school closures during the pandemic. The research situation is not conclusive in this context. Some studies report disadvantages in learning due to school closures (e.g., Hammerstein et al., 2021), while other studies demonstrate good learning outcomes with digital materials (e.g., Amilyana et al., 2021; Meeter, 2021). Second, the results show that younger students achieved better test scores when they completed the test booklets at home, rather than at school. Evidence from homework-studies suggests desirable effects on completion only for younger students, because they appear to have less developed study and of self-management habits (Patall et al., 2008). This may also apply to testing situations. From these findings, we can draw implications for science education at this transition, and for data collection.

According to the results, it is not relevant for the competencies in the transition phase whether the instruction previously took place in face-to-face or distance learning. On the one hand, it is conceivable that chemistry-related competencies are taught to a lesser extent than, for example, mathematical or linguistic competencies, so that the difference between face-to-face and distance learning is not as significant. On the other hand, before and after the school closures, there was a difference in competencies between the beginning and the end of the transition phase. Consequently, students acquired competencies during that transition phase, regardless of whether they were taught in face-to-face settings. Therefore, it is possible that the students acquired some of the measured competencies outside of the school context. Consequently, we should take care to ensure that students acquire chemistry-related competencies more systematically in school. For that purpose, we should use the curricula of the subjects involved, and ensure an exchange between the teachers of both school types, in order to promote cumulative learning processes.

The results on the different designs of data collection suggest that younger students were less likely to comply with test-taking instructions at home, or that they were better able to concentrate at home. To verify the reasons for the measured differences, we would need to conduct further data collection, and control for conditions at home.

Limitations

This study has the following limitations. First, some of the subsamples are incomplete, or significantly smaller than other subsamples. This is especially true for the second measurement point. Here, only about half of the test booklets distributed to the students were completed and returned. In addition, we cannot substantiate explanatory approaches relating to the conditions of the testing situation at home, based on this data. They are merely assumptions. Finally, we must consider that the sample used is limited to only a few schools in Germany. Therefore, we cannot generalize these results for the whole of Germany, or for other countries.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Acknowledgments

We would like to thank Editage (www.editage.com) for English language editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

American Association for the Advancement of Science (2001). *Atlas of Science Literacy: Project 2061*. Washington DC: American Association for the Advancement of Science.

American Association for the Advancement of Science (2013). *Project 2061: Benchmarks and Standards as Tools for Science Education Reform*. Washington, DC: American Association for the Advancement of Science.

Amilyana, A. S., Noer, T. K. M. S., and Raharjo. (2021). A teaching material based on science, environment, technology, and society to improve student's critical thinking skills: synchronous and asynchronous learning during Covid-19 Pandemic. *Int. J. Recent Educ. Res.* 2, 372–391. doi: 10.46245/ijorer.v2i4.109

Celik, K. N., and Walpuski, M. (2018). "Development and Validation of Learning Progressions on Chemical Concepts," in *Proceedings of the Electronic Proceedings of the ESERA 2017 Conference: Research, Practice and Collaboration in Science Education*, (Dublin).

Education, Audiovisual and Culture Executive Agency (2011). *Science Education in Europe: National Policies, Practices and Research, Brussels*. Available online at: https://op.europa.eu/o/opportal-service/download-handler?identifier=bae53054-c26c-4c9f-8366-5f95e2187634&format=pdf&language=en&productionSystem=cellar&part= (accessed April 20, 2022).

Hammerstein, S., König, C., Dreisörner, T., and Frey, A. (2021). *Effects of COVID-19-Related School Closures on Student Achievement - A Systematic Review*. Frankfurt: Goethe Universität Frankfurt. doi: 10.31234/osf.io/mcnvk

Heller, K. A., and Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klasse, Revision (KFT 4-12 + R) [Cognitive skills test for grades 4–12, Revision]*. Göttingen: Hogrefe.

Hempel, M. (2010). "Zur Anschlussfähigkeit der Sachfächer an den Sachunterricht – eine Erkundungsstudie' [About the connectivity of natural and social science subjects to Sachunterricht]," in *Anschlussfähige Bildung im Sachunterricht*, eds H. Giest and D. Pech (Bad Heilbrunn: Klinkhardt), 75–82.

Kähler, J., Hahn, I., Ihme, J. M., and Köller, O. (2020). Empirical work: Science literacy of preschool children. Effects of structural and process characteristics of the parental home and the kindergarten on the scientific competency of 4- to 6-year-olds. *Psychol. Erziehung Unterr.* 67, 1–21.

Lee, J. (2012). "Cumulative Learning," in *Encyclopedia of the Sciences of Learning*, ed. N. M. Seel (Boston, MA: Springer), 887–893. doi: 10.1007/978-1-4419-1428-6_1660

Maldonado, J. E., and De Witte, K. (2020). *The effect of school closures on standardised student test outcomes*. Available online at: https://lirias.kuleuven.be/retrieve/588087 (accessed April 19, 2022).

Meeter, M. (2021). Primary school mathematics during the COVID-19 pandemic: No evidence of learning gaps in adaptive practicing results. *Trends Neurosci. Educ.* 25:100163. doi: 10.1016/j.tine.2021.100163

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2021). *Lehrplan für die Primarstufe in Nordrhein-Westfalen: Fach Sachunterricht*. Frechen: Ritterbach.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2008). *Lehrplan Sachunterricht für die Grundschulen des Landes Nordrhein-Westfalen*. Frechen: Ritterbach.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2013). *Kernlehrplan für die Gesamtschule - Sekundarstufe I in Nordrhein-Westfalen: Naturwissenschaften Biologie, Chemie, Physik*. Frechen: Ritterbach.

Möller, K. (2016). "Stufenübergreifendes Lernen fördern - Problem, Ziele, Konzept, Projektaufbau," in *Authentizität und Lernen - das Fach in der Fachdidaktik: Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Berlin 2015*, ed. C. Maurer (Regensburg: Universität Regensburg), 243–245.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., and Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. Available online at https://timssandpirls.bc.edu/timss2019/international-results/(accessed April 19, 2022).

National Research Council (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.

Oludare, B. A., and Alade, I. A. (2018). Influence of home and school environmental factors on cumulative children learning in primary schools. *Eur. J. Educ. Stud.* 4, 155–166.

Organization for Economic Cooperation and Development (2019). *PISA 2018 Ergebnisse (Band 1): Was Schülerinnen und Schüler wissen und können (PISA)*. Bielefeld: WBV media.

Patall, E. A., Cooper, H., and Robinson, J. C. (2008). Parent involvement in homework: a research synthesis. *Rev. Educ. Res.* 78, 1039–1101. doi: 10.3102/0034654308325185

Schneider, W., Schlagmüller, M., and Ennemoser, M. (2017). *LGVT 5-12+: Lesegeschwindigkeits- und -verständnistest für die Klassen 5-12*. Göttingen: Hogrefe.

Shin, N., Choi, S.-Y., Stevens, S. Y., and Krajcik, J. S. (2017). The impact of using coherent curriculum on students' understanding of core ideas in chemistry. *Int. J. Sci. Math. Educ.* 17, 295–315. doi: 10.1007/s10763-017-9861-z

Tomasik, M. J., Helbling, L. A., and Moser, U. (2021). Educational gains of in-person vs. distance learning in primary and secondary schools: a natural experiment during the COVID-19 pandemic school closures in Switzerland. *Int. J. Psychol.* 56, 566–576. doi: 10.1002/ijop.12728

Waldo, J. T. (2014). Application of the test of scientific literacy skills in the assessment of a general education natural science program. *J. Gen. Educ.* 63, 1–14. doi: 10.5325/jgeneeduc.63.1.0001

Wiser, M., Smith, C. L., and Doubler, S. (2012). "Learning progressions as tools for curriculum development: lessons from the inquiry project," in *Learning Progressions in Science: Current Challenges and Future Directions*, eds A. C. Alonzo and A. W. Gotwals (Rotterdam: Sense Publishers), 359–403. doi: 10.1007/978-94-6091-824-7_16

DOI: 10.3389/feduc.2022.928987
URN: urn:nbn:de:hbz:465-20220909-091418-1